Talk: 3(3)

Errors, Biases, and Corrections for Weighing Gauge Precipitation Measurements from the WMO Solid Precipitation Intercomparison Experiment

J. Kochendorfer, R. Nitu, M. Wolff, C. B. Baker, R. Rasmussen, M. Earle, A. Reverdin, K. Wong, C. D. Smith, D. Yang, Y. A. Roulet, S. Buisan, K. Isaksen, T. Laine, J. L. C. Aceituno, R. Brækkan, S. Landolt, and A. Jachcik

Abstract

Although precipitation has been measured for many centuries, precipitation measurements are still beset with significant biases and errors. Solid precipitation is particularly difficult to measure accurately, and biases between winter-time precipitation measurements from different measurement networks or different regions can exceed 100%. Using precipitation gauge results from the WMO Solid Precipitation Intercomparison Experiment (WMO-SPICE), errors in precipitation measurement caused by gauge uncertainty, spatial variability in precipitation, hydrometeor type, and wind are quantified. The methods used to calculate gauge catch efficiency and correct known biases are described briefly. Transfer functions describing catch efficiency as a function of air temperature and wind speed are also presented. In addition, the biases and errors associated with the use of a single transfer function to correct gauge under-catch at multiple sites are discussed.

Introduction

Like many atmospheric measurements, precipitation is subject to the observer effect, whereby the act of observing affects the observation itself. The interference of a precipitation gauge on the air flow around it affects the measurement of precipitation. This is because hydrometeors falling towards a precipitation gauge can be deflected away from the gauge inlet due to changes in the velocity of the air around the gauge that are caused by the gauge itself. The magnitude of this effect varies with wind speed, wind shielding, the shape of the precipitation gauge, and the predominant size, phase, and fall velocity of the hydrometeors. Because all of these factors can significantly affect the amount of undercatch, it is difficult to accurately describe and correct the resultant errors for all gauges in all places in all types of weather. This has been an active area of research for over 100 yrs. (eg. Nipher, 1878;Alter, 1937;Jevons, 1861;Heberden, 1769), with significant findings for manual measurements described in a WMO intercomparison performed in the 1990s (Goodison et al., 1997;Yang et al., 1995;Yang et al., 1998). More recently studies of the magnitude and importance of such measurement errors have also been performed using both analytical (Theriault et al., 2012;Colli et al., 2015;Colli et al., 2016;Nespor and Sevruk, 1999) and observational approaches (eg. Rasmussen et al., 2012;Wolff et al., 2013;Ma et al., 2015;Wolff et al., 2015;Chen et al., 2015).

Due to the importance of precipitation measurements for hydrological, climate, and weather research, and also due to the many outstanding unanswered questions and uncertainties regarding its measurement, beginning in 2010 the WMO began planning an international intercomparison focused on solid precipitation measurements. The goals of this intercomparison included the assessment of new automated gauges and wind shields, and the development of corrections for these gauges and wind shields. The ultimate goal of all of this work is to facilitate the creation of accurate and consistent precipitation records across different climates and different measurement networks, including measurements made using many different precipitation different gauges and shields (eg. Forland and Hanssen-Bauer, 2000;Yang and Ohata, 2001;Scaff et al., 2015).

Results from the WMO Solid Precipitation Intercomparison were used to develop corrections for different types of weighing gauges, within different types of shields. Due to the nature of this unique dataset, which includes many periods of precipitation from many different sites, gauges and shields, new analysis techniques were also developed to accurately develop corrections and describe the errors inherent in applying such corrections. The focus of the work described below is on the most ubiquitous precipitation gauges used both within SPICE and in national networks for the measurement of solid precipitation. Unshielded and single Altar reference weighing gauges were present at all of the sites that had a an automated Double Fence Intercomparison Reference (DFIR), providing a unique opportunity to develop and test wind speed corrections for these gauges using multiple sites varying significantly in their siting, elevation, and climate.

Methods

Eight sites, each of which had an automated reference DFIR-shielded gauge, were included in this analysis. They include the Canadian CARE site (CARE), the Norwegian Haukeliseter site (Hauk), the Swiss Weissfluhjoch site (Weis), the Finnish Sodankyla site (Sod), the Canadian Caribou Creek site (CaCr), the Spanish Formigal site (For), the US Marshall site (Ma), and the Canadian Bratt's Lake site (BrLa). These sites are described in more detail in the WMO-SPICE commissioning reports (available here: http://www.wmo.int/pages/prog/www/IMOP/intercomparisons/SPICE/SPICE.html). Results from the winters of 2013-2014 and 2014-2015 were used, with the exception of Formigal, Spain, where only one season of DFIR measurements was available. The reference weighing precipitation gauge types were designated by the SPICE International Organizing Committee as the Pluvio2 (OTT Hydromet, Kempten, Germany) and the T200B (3-wire T200B, Geonor Inc., Oslo, Norway). Some of the sites used here relied upon the Pluvio^2 gauge (Sod, Weis, and For) and others used the T200B (CARE, Hauk, CaCr, Ma, and BrLa). Some of these sites had both types of gauges, with preliminary analysis indicating no significant differences between them, so we hypothesized that the effects of shielding, siting, and climate were much more important than the type of gauge used. The results of this study will confirm this hypothesis.

The 30 min precipitation measurements created using the WMO-SPICE smoothing, QA/QC, and event selection criteria were used for all analyses presented here. These methods include the application of a Gaussian filter for minimizing high-frequency noise, the use of a separate precipitation detector for determination of periods with precipitation, and a minimum threshold of 0.25 mm from the reference DFIR-shielded precipitation gauge.

Additional QA/QC was performed on the resultant datasets for the purpose of developing accurate transfer functions. For example, at several sites wind directions associated with compromised wind speed and precipitation measurements due to wind-shadowing from towers, shields, and other obstructions were removed from the record. For the sake of developing and testing transfer functions, minimum thresholds were also used for the gauges under test.

The use of a minimum threshold was necessary because even the reference DFIR precipitation measurements were subject to random variability. Tests performed using identical gauge/shield combinations likewise revealed that the application of a minimum threshold to only one gauge arbitrarily included some events near the threshold and excluded others, and thereby biased the results towards the gauge used for the event selection. Because the results were sensitive to the magnitude of the threshold of the gauge under test, a conservative minimum threshold for the gauge under test was estimated using Eq. 1.

$$THOLD_{UT} = median\left(\frac{P_{UT}}{P_{DFIR}}\right) 0.25\, mm \tag{1}$$

where $THOLD_{UT}$ is the threshold of the gauge under test, P_{UT} is the 30-min precipitation from the gauge under test, and P_{DFIR} is the 30-min DFIR precipitation. Only solid precipitation measurements (T_{air} < -2 °C) with relatively high winds (U_{10m} > 5 $\mathrm{ms^{-1}}$, and U_{10m} < 9 $\mathrm{ms^{-1}}$) were used for the determination of the median catch ratio used to determine the $THOLD_{UT}$. When all available measurements were used for the determination of the minimum threshold for the gauge under test, the inclusion of rain and low wind speed measurements resulted in a higher minimum threshold, which may have erroneously excluded valid low-rate, low-catch-efficiency solid precipitation measurements from the analysis.

A minimum threshold was thus calculated for the reference unshielded ($THOLD_{UT}$ = 0.06 mm) and single Alter gauges ($THOLD_{UT}$ = 0.10 mm), and all respective measurements from the gauges under test that were below this threshold were excluded from the analysis. Unreasonably large accumulations were also removed using the measured catch efficiency ($CE = P_{UT}/P_{DFIR}$), with large outliers from the gauge under test identified and excluded from the analysis when the catch efficiency was more than three standard deviations greater than one: $CE > [3 \times stdev(CE) + 1]$. The resultant maximum unshielded (UN) threshold was 1.93 x P_{DFIR}, and the maximum single Altar (SA) threshold was 1.98 x P_{DFIR}.

To develop transfer functions for both 10 m height and gauge-height wind speeds, the best available wind speed sensor at every site was used to estimate both the 10 m and the gauge height wind speeds. The exact methods used to do this varied site-by-site based on the available wind speed measurements, but generally the log-profile law was used to predict the change in wind speed with height.

$$U_z \approx \ln\left[{(z-d)}/{z_0}\right] \tag{2}$$

where U_z is the wind speed (U) at a height z (U_z). Using 30-min mean wind speeds uncompromised by obstacles, the roughness length and displacement height were estimated for sites with wind profile measurements (Thom, 1975), and for sites without wind speed measurements at multiple heights a generic roughness length (z_0 = 0.01 m) and displacement height (d = 0.4 m) were used.

The available measurements from all eight sites were combined to create individual transfer functions for both the UN and SA gauges. For example, the unshielded reference weighing gauge measurements at all eight sites were pooled together and used to create a single universal UN transfer function. Individual site errors and biases were assessed by applying the resultant universal transfer function and comparing the results to the DFIR measurements at each site. These errors were used to calculate RMSE and biases site by site, for all of the sites. This approach was chosen because it produced transfer functions that best represented all of the sites within WMO-SPICE while simultaneously providing realistic estimates of the magnitude of site biases that can occur based on local variations in climate.

A single transfer function of T_{air} and U was created using all the like reference precipitation gauge measurements. For example, all of the unshielded reference precipitation measurements were grouped together irrespective of whether they were recorded using a Pluvio^2 or a T200B. This transfer function was then used to estimate the reference DFIR precipitation, and the RMSE and bias were calculated at each site individually using the difference between the corrected single Alter or unshielded gauge and the actual DFIR precipitation. Equation 3 describes the form of transfer function used.

$$CE = e^{-a(U)\left(1 - \left[\tan^{-1}\left(b(T_{air})\right)+c\right]\right)}, \tag{3}$$

where U is wind speed, T_{air} is the air temperature, and a, b, and c are coefficients fit to the data. The sigmoid transfer function (Wolff et al., 2013) was also tested with these data, but did not perform significantly better than eq. 1, so the simpler eq. 1 was used.

Without explicitly including T_{air}, transfer functions for mixed and solid precipitation were also created separately as an exponential function of U using Eq. 4.

$$CE = (a)e^{-b(U)} + c \tag{4}$$

where a, b, and c are coefficients fit to the data. This was done for comparison with past studies that used similar techniques, to make such corrections available to users that prefer them, and also to evaluate the advantages and disadvantages of explicitly including T_{air} in the transfer functions.

Due to the prevalence of air temperature measurements in precipitation measurement networks, and the fact that not all of the WMO-SPICE sites included precipitation type measurements, T_{air} was used to determine precipitation type. Mixed precipitation was defined as $T_{air} \geq -2$ °C and $T_{air} \leq 2$ °C, and solid precipitation was defined as $T_{air} < -2$ °C. Liquid precipitation ($T_{air} > 2$ °C) data were also evaluated, but due to the limited quantify of warm-season measurements and the negligible magnitude of the liquid precipitation correction, no rain transfer functions were created. For comparison with the Eq. 3 results, the resultant transfer functions were used to correct the precipitation measurements, with no correction applied to the liquid precipitation measurements.

The maximum wind speed for every transfer function developed was chosen by visually assessing the availability of high wind speed results for all air temperatures, after plotting CE as a function of T_{air} and U. Furthermore, for the validation of the transfer functions, wind speeds exceeding the maximum wind speed were replaced by the maximum wind speed, which is recommended for the application of these transfer functions. Failure to do this resulted in large RMSE for some sites due to the inaccuracy of transfer functions at very high wind speeds, where the data available to constrain the fit were scarce. Although all data were used in the creation of the transfer functions, the high wind speed data were more accurately corrected using the maximum wind speed rather than the measured wind speed. Prospective users of these equations should take this as evidence of the undesirable effects of applying these transfer functions using wind speeds greater than the maximum wind speed.

<u>Results</u>

Combining precipitation measurements from all available sites, eight sites contributed both unshielded and single Altar references gauge measurements. The associated RMSE and bias calculated from the difference between the corrected SA and UN gauges and the DFIR were estimated. For the sake of comparison and to demonstrate the efficacy and limitations of the transfer functions, the uncorrected SA/UN RMSE and bias are also included. The SA and UN results were similar, with an example of the UN, uncorrected results shown in Fig. 1 and 2.

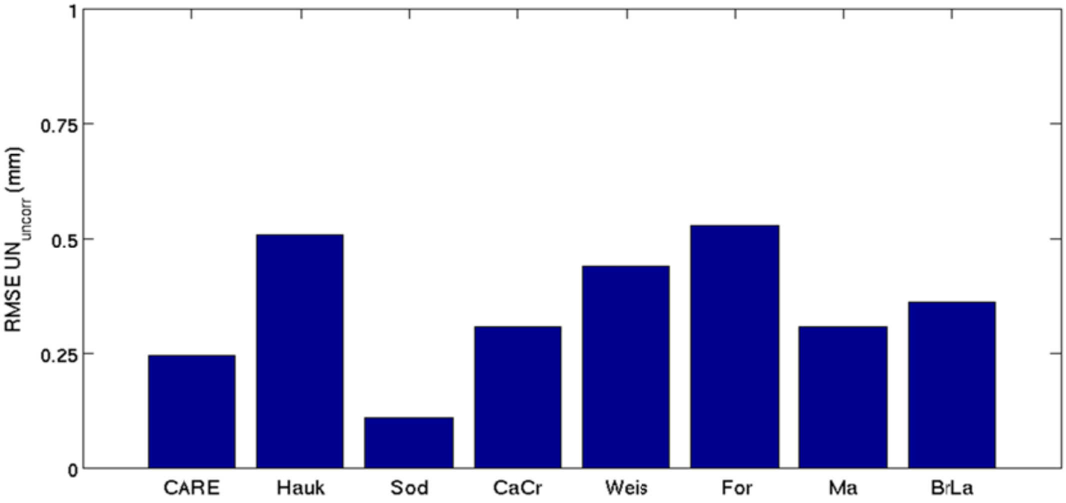

Fig. 1. RMSE from all 8 sites calculated from the difference between the DFIR and uncorrected, unshielded gauges.

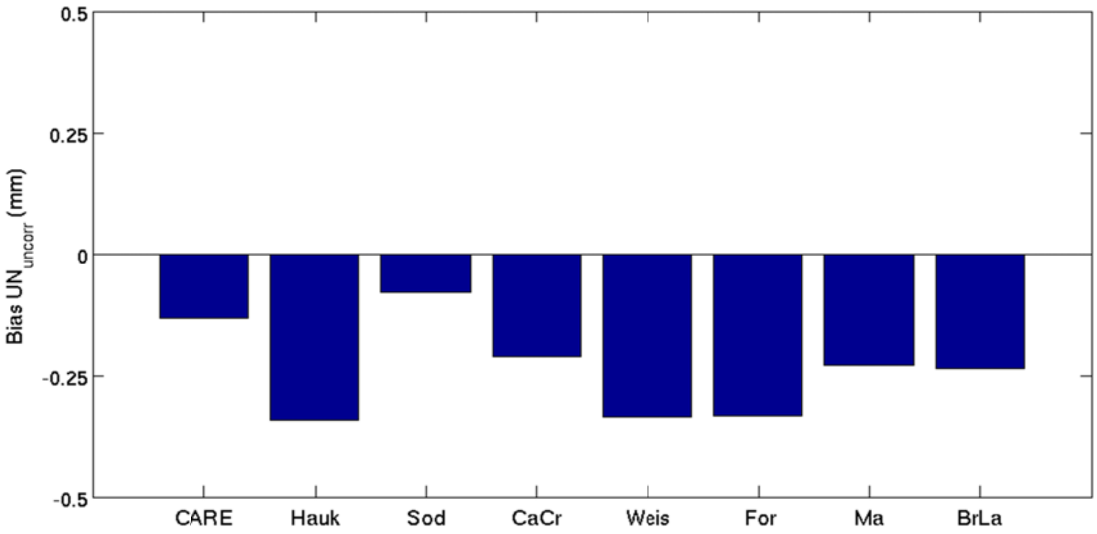

Fig. 2. Bias from all 8 sites calculated from the difference between the DFIR and the uncorrected, unshielded gauges.

The corrected unshielded RMSE and bias are also shown (Fig. 4 and 4), showing that in most cases the RMSE and biases were normally improved by the application of the transfer function. The change in bias is most noteworthy, as before correction all of the gauges exhibited a significant negative bias indicating undercatch, and after correction the biases were generally much closer to zero and more variable in sign, with the mean corrected UN measurements either slightly larger or smaller than the DFIR measurements at the different sites.

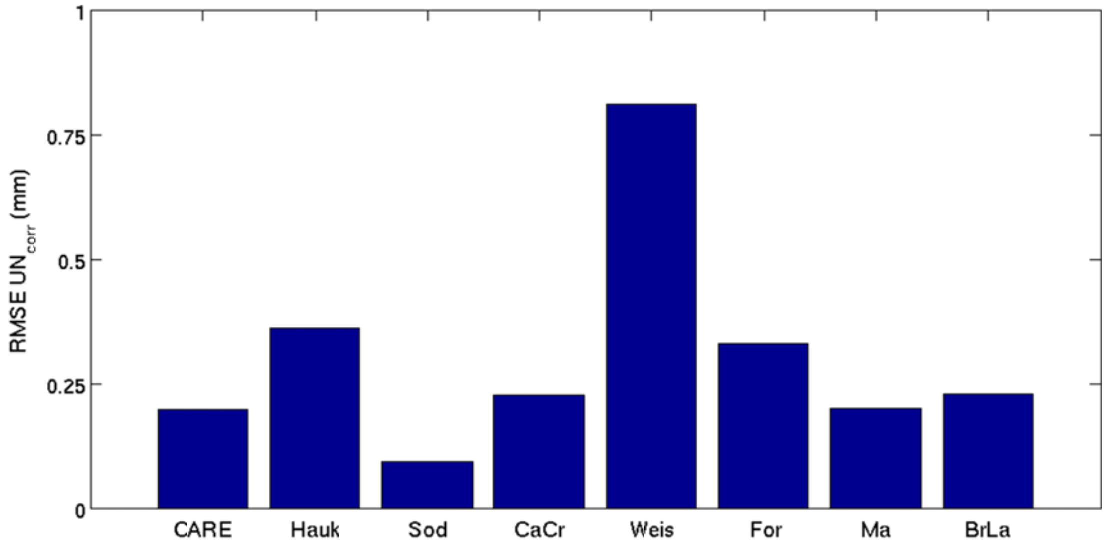

Fig. 3. RMSE from all 8 sites calculated from the difference between the DFIR and the corrected, unshielded gauges.

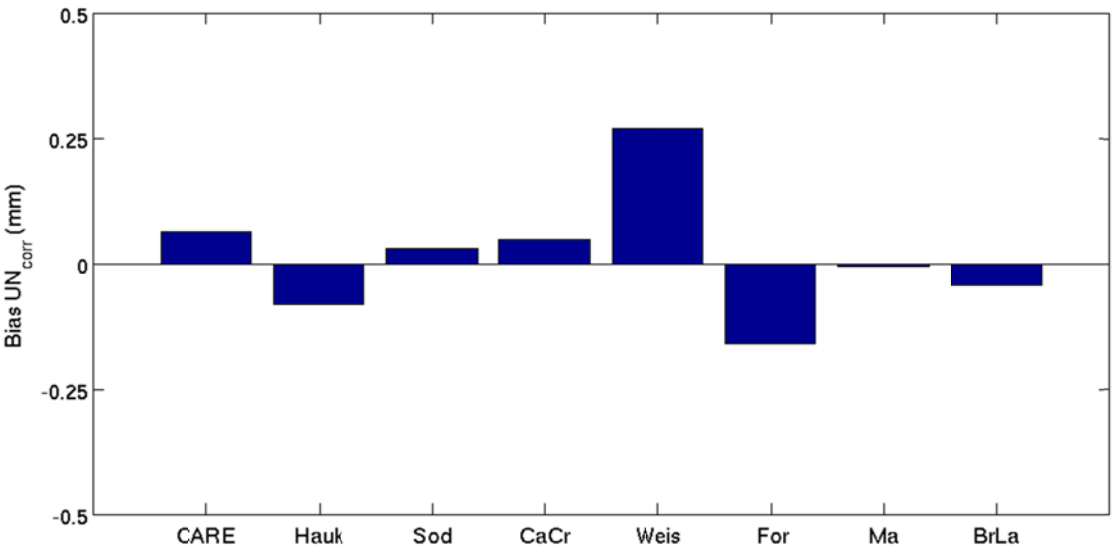

Fig. 4. Bias from all 8 sites calculated from the difference between the DFIR and the corrected, unshielded gauges.

After differentiating the 30 min periods into different precipitation types using the air temperature, a simple exponential function of wind speed (Eq. 4) was used to describe the relationship between catch efficiency and air temperature. These results were produced for the single Alter and unshielded gauges and also for the other weighing gauges under test, and were generally quite similar to the Eq. 3 results.

Generally the transfer functions were able to reduce the bias significantly, but as indicated by the RMSE, significant and seemingly random differences between the corrected gauge measurements and the DFIR measurements still persisted. This was presumably due to a combination of factors such as random spatial variability of precipitation across an individual site, sensor-induced noise in the precipitation measurements, the effects of variability of crystal type on catch efficiency, and also the effects of the transfer function corrections multiplying small measurement errors by two ore even three at high wind speeds.

For the UN and SA gauges that were present at all of the sites, the Weissfluhjoch results were significantly different than the other sites'. This is because at Weissfluhjoch the catch efficiency decreased at wind speeds above 5 ms^{-1} much less than at the other sites. The transfer functions developed for all sites worked fairly well at Weissfluhjoch at low wind speeds, but greatly over-corrected the unshielded and SA gauges at high wind speed (Fig. 5), creating large mean biases and RMSE. For comparison with a more typical, high wind speed site, results from Haukeliseter are also shown (Fig. 6). We hypothesize that the complex topography at this site and perhaps compromised wind speed measurements contributed to this problem. To more carefully examine the effects of the Weissfluhjoch measurements on the resultant transfer functions, a sensitivity test performed using the UN and SA results indicated that exclusion of the Weissfluhjoch measurements did not significantly affect the resultant 'universal' transfer functions. However the general trend found for the Weissfluhjoch errors generally were valid for all sites, with the RMSE and bias driven by the high wind speed results. This is at least in part because for high wind speeds in cold, snowy conditions the transfer function correction can more than triples the magnitude of the precipitation. This large correction significantly enhanced the differences between the corrected CE and the actual measured CE, especially when the measured CE was higher than typical; at high wind speeds a relatively small difference or error in the measured precipitation is tripled, resulting in errors in the corrected precipitation of same order of magnitude as the DFIR measurement itself. For this reason errors in corrected precipitation or CE look significantly different than errors in the measured CE, highlighting the need for determination of errors and biases in the corrected precipitation, rather than the CE transfer function; to accurately represent errors in the transfer function, the transfer function should be applied to real data and compared to the standard. Likewise, changing the max U and thereby changing the magnitude of the applied transfer function at high wind speeds has a large effect on the resultant site-specific errors and biases. Changing the U_{gh} max U from 7.2 to 5, as an extreme example, improved the UN Weissfluhjoch bias from 35% to 16%, while simultaneously worsening the Haukeliseter bias from -7% to -23%, with similarly significant changes to the RMSE at each site.

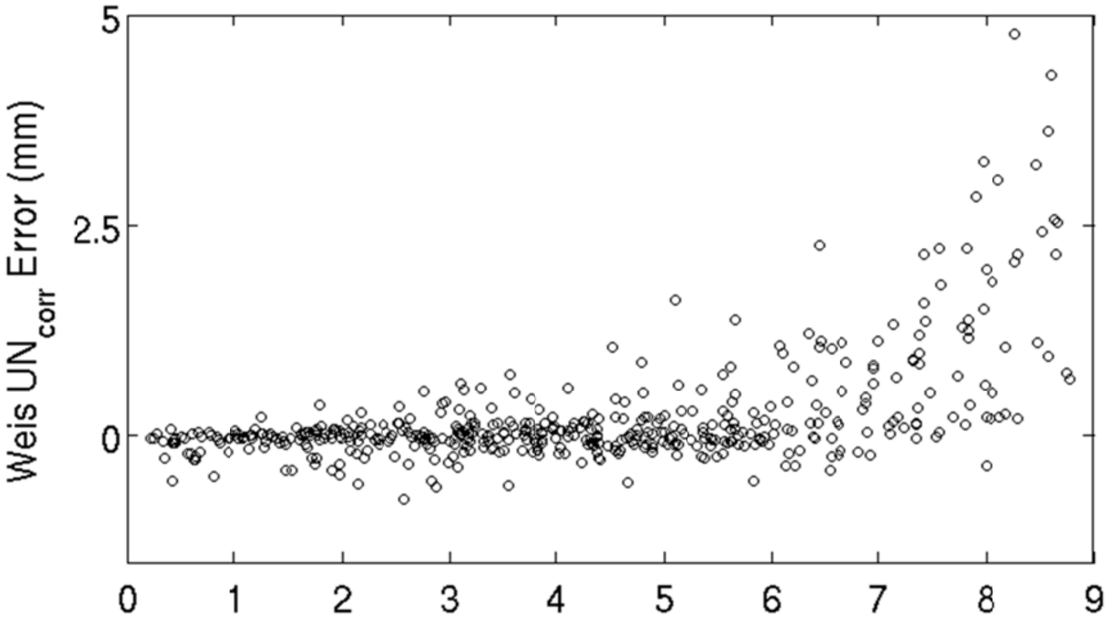

Fig. 5. Errors in the 30-min precipitation from Weissfluhjoch, estimated from the difference between the DFIR and the corrected, unshielded gauge.

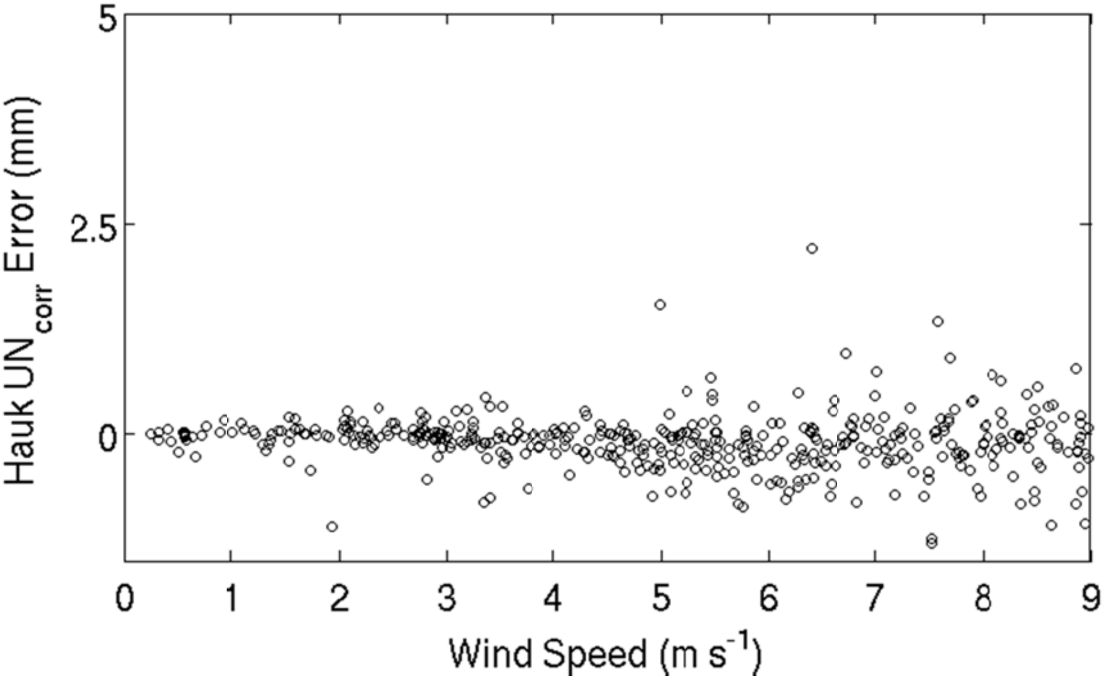

Fig. 6. Errors in the 30-min precipitation from Haukeliseter, estimated from the difference between the DFIR and the corrected, unshielded gauge.

Discussion

Transfer functions were developed and tested on eight separate sites, and with the exception of a complex mountainous site in Switzerland, they showed great efficacy in reducing the mean differences between unshielded and single Altar weighing gauge precipitation measurements. In general the more complex sites were more difficult to correct, with the transfer functions performing worst at the mountainous sites, which were Haukeliseter, Formigal, and Weissfluhjoch. Excluding these sites and only considering the flat sites, the bias in the unshielded measurements was reduced to below 0.06 mm, or about 10%. As indicated by the RMSE however, significant differences between corrected measurements and the standard DFIR measurements persisted even after correction. For example, excluding these same three mountainous sites, the RMSE were still about 0.2 mm, or 30% of the mean 30-min precipitation. This suggests that to produce more accurate measurements we must develop a better understanding of the physics and microphysics affecting these errors. Better gauge shielding for may also be necessary to reduce the uncertainty in the measurements.

The results from the Sodankyla site support the conventional wisdom that putting your gauge in a clearing in the forest effectively removes the wind speed errors and biases.

These results are preliminary, and in addition to the actual transfer coefficients and a more through description of the results, additional weighing gauges and shields from WMO-SPICE will be included in future analysis.

References

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.

Chen, R., Liu, J., Kang, E., Yang, Y., Han, C., Liu, Z., Song, Y., Qing, W., and Zhu, P.: Precipitation measurement intercomparison in the Qilian Mountains, north-eastern Tibetan Plateau, Cryosphere, 9, 1995-2008, 10.5194/tc-9-1995-2015, 2015.

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.

Colli, M., Lanza, L. G., Rasmussen, R., and Theriault, J. M.: The Collection Efficiency of Shielded and Unshielded Precipitation Gauges. Part II: Modeling Particle Trajectories, Journal of Hydrometeorology, 17, 245-255, 10.1175/jhm-d-15-0011.1, 2016.

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.

Heberden, W.: Of the different auantities of rain, which appear to fall, at different heights, over the same spot of ground., Philosophical Transactions (1683-1775), 59, 359-262, 1769.

Jevons, W. S.: On the deficiency of rain in an elevated raingauge as caused by wind, Philosophical Magazine, 22, 421-433, 1861.

Ma, Y., Zhang, Y., Yang, D., and Bin Farhan, S.: Precipitation bias variability versus various gauges under different climatic conditions over the Third Pole Environment (TPE) region, International Journal of Climatology, 35, 1201-1211, 10.1002/joc.4045, 2015.

Nespor, V., and Sevruk, B.: Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation, Journal of Atmospheric and Oceanic Technology, 16, 450-464, 10.1175/1520-0426(1999)016<0450:eowieo>2.0.co;2, 1999.

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.

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.

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

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.

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.

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., Golubev, V. S., Bates, R., Pangburn, T., and Hanson, C. L.: Accuracy of NWS 8" standard nonrecording precipitation gauge: Results and application of WMO intercomparison, Journal of Atmospheric and Oceanic Technology, 15, 54-68, 10.1175/1520-0426(1998)015<0054:aonsnp>2.0.co;2, 1998.
Yang, D. Q., and Ohata, T.: A bias-corrected Siberian regional precipitation climatology, Journal of Hydrometeorology, 2, 122-139, 10.1175/1525-7541(2001)002<0122:abcsrp>2.0.co;2, 2001.