Response to Referee #1

General comments from Referee #1:

The manuscript "Testing and Development of transfer functions for weighing precipitation gauges in WMO-SPICE" deals with the interesting and relevant topic of how to correct undercatch especially of snowfall for different sensors. This paper is basically an extension of the Kochendorfer et al. 2017 (HESS) that again is based on the (Kochendorfer et al. 2016) study. Here the apply the previously developed correction functions to different sensors and locations and found that the previously developed, more general correction function of Kochendorfer et al. 2017 (HESS) is always recommended to use. In general the paper is mostly very written and I found no crucial error or mistake. My major concern is more related to presentation and the structure of the manuscript. Both section "Methods" and "Results (& Discussion)" are written very detailed, technically and lengthy, giving each sensor, its location, and treatment of the data its place. Figures 4 - 11 display in detail the (mostly non-existing) differences between correction functions. Partly, I found it hard to follow the overall structure/story of the manuscript. In my opinion, this structure would have been justified if the authors would have found several different best-performing equations that need to be presented alongside each sensor and in comparison with the general correction function. But, given that in most cases the general (somehow site and sensor unspecific) correction function is as well performing and recommended by the authors, a more summarizing structure would have been much more feasible and would have cause less redundancies. In its current version, the manuscript unfortunately reads quite lengthy, especially given the final outcome. Many of the interpretations named in the abstract and conclusion are a bit out of blue and should be much more elaborated. Hence, instead of the detailed description, I would recommend to show some summarizing, and comparing analyses/graphics on sensor performance.

Based on this basic concern, and the following further smaller concerns, I recommend to provide either a revised, restructured version that strongly condense the specific sensor part and put more focus on the comparison analysis, or to show the transferability of the Kochendorfer et al. 2017a Equation 3 to other sensors in a HESS-technical note. I am sorry to be that harsh, but I really do think that the level of detail and length of the manuscript does not match your findings.

Authors' response to general comments:

We think this is a fair assessment of the manuscript in some respects. We chose to separate the different types of WMO-SPICE weighing gauge measurements into different manuscripts before discovering that the 'universal' transfer functions (Kochendorfer et al., 2017a) performed well on all of the manufacturer-provided single-Alter shielded and unshielded gauges. The success of the 'universal' transfer functions on these other weighing gauges was a surprise. We had originally assumed that all of the different types of WMO-SPICE weighing gauges would require their own unique transfer function, and it was based in part on this assumption that we decided to separate the WMO-SPICE weighing gauge transfer functions into two separate manuscripts. In addition, the length and amount of detail included in the Kochendorfer et al. (2017a) manuscript would have been unwieldy if all of the weighing gauges from WMO-SPICE were added to it.

We agree that the mode of presenting the results can be simplified and improved, and have made changes accordingly. For example, many of the figure styles were developed for Kochendorfer et al. (2017a), and were chosen to comprehensively compare different types of transfer functions using different wind speeds. In order to achieve a more condensed view, we have changed the figure style by merging gauges into the same plot and excluding some of the transfer functions. As the reviewer pointed out, the majority of the transfer functions evaluated in this manuscript were determined to be obsolete because they were no better than an appropriate 'universal' transfer function or a pre-SPICE transfer function from Kochendorfer et al. (2017b). In addition, the differences between most of the transfer functions evaluated were quite small. Because of this, they do not all merit such a thorough evaluation, the format of the results has been consolidated significantly in response to the reviewer's suggestion. This has helped to improve the presentation of the manuscript and to focus it on the more significant results.

We maintain, however, that the results are significant enough to merit publication as a new manuscript. The manuscript provides independent validation of previously-existing transfer functions, and it also demonstrates that transfer functions derived using one type of weighing precipitation gauge can be used on another type of similar gauge. Many readers (including gauge and site developers, precipitation observers and data users) will find this significant, surprising, and also very useful. The manuscript, based on a well-executed and carefully-designed experiment, includes the evaluation of nine different types of gauges and wind shields.

Reviewer #1 should also note that in addition to the different single-Alter and unshielded gauges that were evaluated, gauges in several other types of windshields were also assessed. Some measurements recorded within these other winds shields were used to validate the results of Kochendorfer et al. (2017b), and others were shown to merit new adjustments. For example, a new set of adjustments for the shielded MRW500 precipitation gauge were derived and recommended. Arguably, such an evaluation on its own could merit publication as a full manuscript, as this is the only transfer function available for this gauge and shield so far, but due to the wealth of new measurements produced by the WMO-SPICE project, it was included with the evaluation of transfer functions for all of the other WMO-SPICE weighing precipitation gauges and wind shields.

The team of WMO-SPICE investigators who authored this manuscript has an obligation to precipitation gauge manufacturers and their customers to recommend transfer functions for each weighing gauge and wind shield tested in the WMO-SPICE project. In addition to the forthcoming WMO report, a journal article like this is the best way to disseminate the results of an international project like WMO-SPICE.

Further general concerns from Referee #1:

1. Several time differences are named significant or non-significant, but I somehow missed the section describing how a significance test was performed.

Authors' response: In response to this t-tests have been performed to more objectively evaluate the significance of differences between the errors associated with different transfer functions and also the unadjusted measurements. The results of these t-tests have been documented in the revised manuscript to replace the subjective determination of significance included in the original manuscript.

2. Given that both the biases of the corrected values and the differences between the corrected functions are mostly rather small (Figure 4- 11, RMSE < 0.5 mm, biases « 0.5 mm, differences far less), I wonder how and if at one can and should interpret these differences, given the measurement accuracy of each sensor.

Authors' response: For the most part, the biases in the corrected measurements and the differences between the different corrections are indeed negligible. The restructured manuscript states this more explicitly.

3. The paper needs to be more independent from Kochendorfer et al. 2017 (HESS). At least the essential equation 3 should also be given in this paper.

Authors' response: This is a good suggestion. The often-cited Eq. 3 and Eq. 4 have been included independently in the paper, and other changes have been made to help the manuscript stand on its own.

Further specific comments:

page 2, line 19-22: I do not understand this sentence. Please check language and try to avoid too long sentences.

Authors' response: The run-on sentence has been divided up for clarification.

page 3ff: I would recomment to use Table 1 much more often to guide the reader through the different configurations.

Authors' response: There are indeed many different gauges and configurations for the reader to familiarize himself or herself with, and the suggestion is a good one. We now refer to the table (which is now Table 2) throughout this section.

page 3, A paragraph that outlines the design of this study would help the reader to follow your following descriptions

Authors' response: We have added an introductory paragraph describing how the intercomparison was designed, with a common automated reference gauge at all sites, along with other shared types of precipitation and meteorological measurements. The goal was to include as many different countries, climates, and gauges as possible while still maintaining some basic standards.

page 3, line 9: typo: "a either a"

Authors' response: Thank you! This has been corrected.

page 3, line 31: CARE, acronym not yet introduced

Authors' response: CARE is "Centre for Atmospheric Research Experiments", has been defined in the manuscript.

page 4, line 4, "have" missing in and a lower porosity

Authors' response: This has been rewritten.

page 4, line 31: 6 s and 1 min data correspond to with sensor / site. Maybe update Table 1

Authors' response: We have now described which data are available with which sensors in a new Table 1.

page 4, line 32: "realistic" and "operational" limits, please define

Authors' response: We have clarified this in the manuscript. The term "realistic" was meant to signify "possible", and was used to remove periods of calibrations/validations and other impossibly high-rate precipitation measurements from the record. The term, "operational" was meant to signify "within the operational specifications of the sensor", such as a weighing gauge depth that was beyond the upper limit of the gauge specifications.

page 5, Maybe I missed it, but where do you define that DFAR is the reference, and please discuss briefly its quality and its deviation from the "truth"

Authors' response: The DFAR was described and defined as the reference at the beginning of the Methods section. The "truth" is indeed difficult to determine for snowfall because pit gauge measurements, which are used as a reference for rain, are subject to blowing snow and capping, and are therefore not appropriate for snowfall. The DFIR is the reference for manual snowfall observations (Goodison et al., 1998.; Yang, 2014). In the beginning of the Methods section, we have added available references to manual DFIR and "bush" gauges, in addition to later comparisons with automated bush gauges if references are available. We also have added a more detailed discussion of how and why the DFAR was defined as the reference for WMO-SPICE.

page 5, line 13ff: How were temperature thresholds defined. If possible refer to citations

Authors' response: The temperature thresholds were defined based on histograms of automated precipitation type and air temperature measurements (Kochendorfer et al., 2017b; Wolff et al., 2015). An explanation and citations have been added to the manuscript.

page 5, line 18: Please provide these equation also here to be more independent from the Kochendorfer et al. (2007a) paper

Authors' response: The equations have been added to the manuscript.

page 7, line 13ff: I do not understand this interpretation. If significant site biases exists, doesn't that mean that you have to develop site-specific empiral correction functions?

Authors' response: This is indeed true, but we do not have a good way to determine such site-specific empirical corrections. These adjustments are designed for use at sites that do not have a good reference, and at such sites there is no good way to determine what the site bias is. For example, in Kochendorfer et al. (2017a) the three high-altitude sites had the largest biases, but at two of these high-altitude sites the biases were negative, and at the other high-altitude site the bias was positive.

page 10, line 31: "1500mm" Geonor ? cp. with page 11, line 1: please be consistent, see also page 11, line 17

Authors' response: These inconsistencies have been corrected.

page 11, line 6: typo: "attributed"

Authors' response: The spelling error has been corrected.

page 11, line 20: does the noise of the Geonor 1500 mm stem from the fact that they have a different sensitivity than the 500 mm ?

Authors' response: Yes that is probably the case. We now suggest this in the manuscript.

page 11, line 20 ff: Are these interpretation valid, given the some deviation and the measurement accuracy?

Authors' response: We may disagree on the definition of 'noise', but it is certainly the case that it is more difficult to measure light snowfall with a 1500 mm Geonor than a 600 mm Geonor. In general, it can be difficult to differentiate signal noise from precipitation with these weighing gauges, and this problem is augmented with the 1500 mm Geonor. This is an important issue for the measurement of snowfall (and light rain over shorter time periods, such as 5 min) because most snowfall is associated with low precipitation rates, particularly in the Arctic.

page 11, line 26, CARE and Marshall *test sites* !

Authors' response: Thank you. This has been corrected.

page 17, Table 1: please add time period of measurements considered

Authors' response: The time period has been documented in the Table 2 caption rather than the actual table. All of the measurements spanned the same time period, so it isn't necessary to describe the time period in the table for each individual site.

References

Kochendorfer, J., Nitu, R., Wolff, M., Mekis, E., Rasmussen, R., Baker, B., Earle, M. E., Reverdin, A., Wong, K., Smith, C. D., Yang, D., Roulet, Y. A., Buisan, S., Laine, T., Lee, G., Aceituno, J. L. C., Alastrué, J., Isaksen, K., Meyers, T., Brækkan, R., Landolt, S., Jachcik, A., and Poikonen, A.: Analysis of single-Alter-shielded and unshielded measurements of mixed and solid precipitation from WMO-SPICE, Hydrol. Earth Syst. Sci., 21, 3525-3542, 2017a.

Kochendorfer, J., Rasmussen, R., Wolff, M., Baker, B., Hall, M. E., Meyers, T., Landolt, S., Jachcik, A., Isaksen, K., Brækkan, R., and Leeper, R.: The quantification and correction of wind-induced precipitation measurement errors, Hydrol. Earth Syst. Sci., 21, 1973-1989, 2017b.

WMO Solid Precipitation Measurement Intercomparison (1998, WMO/TD-No. 872, IOM No. 67, by B.E. Goodison, P.Y.T. Louie and D. Yang)

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, 2015.

Yang, D, 2014: Double Fence Intercomparison Reference (DFIR) vs. Bush Gauge for "true" snowfall measurement, Journal of Hydrology,Vol. 509, p. 94-100

Response to Referee #2

Overview from Referee #2

The paper presents results from tests on the adequacy of (mostly already existing) transfer functions for precipitation gauges. It directly builds on the paper by Kochendorfer et al. 2017, recently published in HESS.

The topic discussed by the Authors is of scientific relevance and timely, and its scope is within the objectives of HESS. The manuscript presents novel findings that may be useful to inform the selection of proper instrumentation for measuring (solid) precipitation. Results and conclusions are clearly outlined; however, I believe the overall presentation should be substantially restructured to better convey the manuscript's findings. Specifically, I think that, in its current form, the manuscript lacks important pieces of information and an overall picture that would enhance its comprehension.

Authors' response: Thank you. We agree with this evaluation and have restructured the manuscript based on the reviewer's suggestions.

In the following, I report a few suggestions for improvement.

General comments:

1. A short introduction should be provided on the reasons why new precipitation gauges are needed, what are the criticalities in measuring precipitation, what we expect the improvements from using alternative measurement systems would be. I understand that many of these aspects were already outlined by the Authors in Kochendorfer et al. 2017; herein, the Authors should focus on measurement equipment alternative to traditional systems. Alternative precipitation gauging systems are blossoming in the hydrological community, and their promise/limitations may be reported to better support the scope of the paper and expand the bibliography.

Authors' response: The Introduction has been augmented with a description of why new precipitation gauges are needed, and areas where the authors see the potential for improvements. By 'alternative' precipitation measurement systems, we assume that Reviewer #2 means non-catchment types of gauges, which include optical devices such as the present weather sensors, present weather detectors, disdrometers, and optical rain gauges. Some of these types of gauges were included in WMO-SPICE, and the results will be detailed in the forthcoming project report. A detailed discussion of these gauges is beyond the scope of the present manuscript, which focusses on weighing gauges ('traditional' systems); however, some references to earlier work have been provided along with a discussion of the advantages and disadvantages of alternative types of measurements.

2. The role of wind in the underestimation of precipitation should be better highlighted through key citations.

Authors' response: We have added more citations of the effects of wind on gauge catch.

3. Why were these specific gauging systems selected? I think that the description of gauges can be improved by providing further details on how they work, what their features contribute to, and what we should be expecting in terms of performance and limitations. I also suggest that Figure 2 is improved and key features are highlighted for each of the gauges.

Authors' response: All of the weighing precipitation gauges tested in WMO-SPICE were included in the manuscript. Many of the gauges were provided by the manufacturers that chose to participate in WMO-SPICE, and others were provided (and selected) by site hosts for their own national and scientific interests. Such an explanation has been added to the manuscript. Some description of the individual gauges was provided in the appropriate Results sections, but these have been expanded and moved to the Methods section. This also helped with the restructuring of the manuscript.
Figure 2 was included to provide readers with visual examples of field installations, and to help familiarize readers with the different types of gauges and windshields discussed in the manuscript. Such images can also be used to visually assess the effects of wind on the different types of gauges and shields. These examples provide valuable context for the interpretation of results and discussion regarding wind effects on different gauge/shield combinations. The technical features of the gauges are now described in a new table, added to the revised manuscript. However visual depictions of the specific transducers for each gauge type, and of other gauge elements such as heaters and buckets, are not critical to the interpretation of results, and will not be included.

4. The Discussion and Conclusions should clearly state what research findings are and recommend best practice for measuring solid precipitation. I suggest the Authors include a Table in the "Synthesis" section where each gauge is coupled with the recommended transfer functions and comments are provided on eventual limitations.

Authors' response: We have modified and restructured the manuscript accordingly, with a larger Synthesis Section, and a smaller Section describing the results of the individual gauges. A table has also been added referencing the recommended transfer functions available either in the Supplemental Material or earlier manuscripts.

Specific comments:

1. Abstract: I think the Abstract should be simplified (it is not necessary to list the names of all gauges) and the paper objectives and results clearly outlined.

Authors' response: We have removed the list of gauges and further simplified the Abstract by focusing on the general objectives and results.

2. Introduction: I believe including a synthesis Table on previous experiments would help the reader to frame the work within previous studies. I also recommend the Authors expand the last paragraph by (i) justifying the selection of specific gauges; (ii) clearly stating hypotheses; (iii) and identifying key objectives.

Authors' response: The Introduction has been expanded to describe in greater detail how the present work relies upon and supports previous studies. Clarification of the inclusion of all available WMO-SPICE weighing gauges has been included. Hypothesis and key objectives have also been described.

3. Methods: Many parameters/terms were not properly defined. I believe the Authors should devote a paragraph to re-state what catch efficiency is, what are key variables influencing the response of the gauges, and to report previously developed transfer functions. Please also clarify the data structure (sentence on Page 4 line 31 is out of the blue).

Authors' response: Thank you. The Introduction has been expanded to more thoroughly introduce the reader to new terminology, precipitation gauge undercatch, and transfer functions. The Methods section has been augmented with definitions of key terms, measurements, and the data structure.

4. Results and Discussion: Since many of the tested gauges had a similar behavior, I do not think separate sections and Figures 4 to 12 are necessary. I suggest the Authors consolidate results in a Table. I would also move the transfer function coefficients in the Supplementary Material.

Authors' response: These are good suggestions. We developed a more succinct way to present the main results, and moved the transfer function coefficients to the Supplement.

5. I think the presentation quality of the paper is sufficient; however, the number of references could be extended. I also suggest the Authors double check the English for minor typos. I herein list some of them:

Authors' response: The number of references has been increased, and several of the authors have checked the revised manuscript for typos

- Page 3 line 9: "consisted of a either a"has been

Authors' response: This has been corrected.

- Page 10 line 6: "This result are"

Authors' response: This has been corrected.

- Page 11 line 6: "measurements are attributed"

Authors' response: This has been corrected.

- Page 13 line 17: "3-dimensional" (please clarify what this means)

Authors' response: By '3-dimensional' we meant as a function of both wind speed and air temperature. The term has been removed, as it is unnecessary.

- Page 15 line 4: This sentence is unclear, please elaborate.

Authors' response: The following sentence clarifies: "At higher wind speeds, where such measurements require doubling or even tripling, the uncertainty in the measurements was also doubled or tripled, accordingly." But the sentence is indeed confusing, and has been rewritten.

I also suggest the Authors pay special attention in defining all acronyms

Authors' response: The manuscript has been reviewed carefully to ensure that all acronyms have been defined properly.

Testing and Development of Transfer Functions for Weighing Precipitation Gauges in WMO-SPICE

John Kochendorfer[1], Rodica Nitu[2,3], Mareile Wolff[4], Eva Mekis[2], Roy Rasmussen[5], Bruce Baker[1], Michael E. Earle[6], Audrey Reverdin[7], Kai Wong[2], Craig D. Smith[8], Daqing Yang[8], Yves-Alain Roulet[7], Tilden Meyers[1], Samuel Buisan[9], Ketil Isaksen[4], Ragnar Brækkan[4], Scott Landolt[5], and Al Jachcik[5]

[1]Atmospheric Turbulence and Diffusion Division, ARL, National Oceanic and Atmospheric Administration, Oak Ridge, TN, 37830, US
[2]Meteorological Service of Canada, Environment and Climate Change Canada, Toronto, ON, M3H 5T4, Canada
[3]World Meteorological Organization, Geneva, CH-1211, Switzerland
[4]Norwegian Meteorological Institute, Oslo, 0313, Norway
[5]National Center for Atmospheric Research, Boulder, CO, 80305, US
[6] Meteorological Service of Canada, Environment and Climate Change Canada, Dartmouth, NS, B2Y 4N6, Canada
[7]Meteoswiss, Payerne, CH-1530, Switzerland
[8]Climate Research Division, Environment and Climate Change Canada, Saskatoon, SK, S7N 3H5, Canada
[9]Delegación Territorial de AEMET (Spanish National Meteorological Agency) en Aragón, Zaragoza, 50007 Spain

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

Abstract.

Weighing precipitation gauges are widely used for the measurement of all forms of precipitation, and are typically more accurate than tipping-bucket precipitation gauges. This is especially true for the measurement of solid precipitation;~~.~~ h~~H~~owever, weighing precipitation gauge measurements must still be adjusted for undercatch in snowy, windy conditions. The WMO-SPICE (World Meteorological Organization Solid Precipitation Intercomparison Experiment) compared different types of weighing precipitation gauges and shields and determined adjustments ~~Adjustments~~ for the undercatch of solid precipitation caused by wind.~~ were developed for different weighing gauge/wind shield combinations tested in WMO-SPICE. These include several different manufacturer-provided unshielded and single-Alter shielded weighing gauges, a MRW500 precipitation gauge within a small, manufacturer-provided shield, and host-provided precipitation gauges within double-Alter, Belfort double-Alter, and small Double-Fence Intercomparison Reference (SDFIR) shields. Previously-derived adjustments were also tested on measurements from each weighing gauge/wind shield combination.~~ For the various combinations of gauges and shields, adjustments using both new and previously existing transfer functions were evaluated. For most of the gauge and shield combinations, previously derived transfer functions were found to perform as well as those ~~The transfer functions developed specifically for each of the different types of unshielded and single-Alter shielded weighing gauges~~more recently derived~~ did not perform significantly better than the more generic and universal transfer functions~~

~~developed previously using measurements from eight different WMO-SPICE sites~~. This indicates that wind shield type (or lack thereof) is more important in determining the magnitude of wind-induced undercatch than the type of weighing precipitation gauge. It also demonstrates the potential for widespread use of the previously ~~-~~developed~~, multi-site single-Alter shielded and unshielded~~ transfer functions. ~~. In addition, corrections for the lower-porosity Belfort double-Alter shield and a standard double-Alter shield were developed and tested using measurements from two separate sites for the first time. Among all of the manufacturer-provided shields tested, with an average undercatch of about 5%, the Belfort double Alter shield required the least amount of correction, and caught ~ 80% of the reference amount of precipitation even in snowy conditions with wind speeds greater than 5 m s^{-1}. The SDFIR-shielded gauge accumulated 98% of the Double-Fence Automated Reference (DFAR) precipitation amount on average, accumulated 90% of the DFAR accumulation in high winds, and was almost indistinguishable from the full-sized DFAR used as a reference.~~ Another ~~interesting,~~ overarching result was that, ~~I~~in general, the more effective ~~wind~~ shields, ~~,~~ ~~that~~ which were associated with smaller unadjusted errors, also produced more accurate measurements after adjustment. This indicates that although transfer functions can effectively reduce measurement biases, effective wind shielding is still required for the most accurate measurement of solid precipitation.

1 Introduction

Precipitation measurements are frequently underestimated due to the interactions among wind, the precipitation gauge, and hydrometeors in the air around the gauge. Wind deflected over and around a precipitation gauge can alter the trajectory of hydrometeors falling toward the gauge inlet, diverting them away from the inlet of the gauge, and causing the gauge to underestimate the actual precipitation rate. The magnitude of this underestimation is affected by the wind speed and the phase, size, density and crystal habit of precipitation, with the problem being particularly severe for snowy, windy conditions.~~.~~ Many past observational (eg. Rasmussen et al., 2012;Wolff et al., 2013;Ma et al., 2015;Wolff et al., 2015;Chen et al., 2015) and theoretical (Theriault et al., 2012;Colli et al., 2015;Colli et al., 2016;Nespor and Sevruk, 1999;Sevruk et al., 1991;Baghapour et al., 2017) studies support this finding, including the first World Meteorological Organization (WMO) Solid Precipitation Measurement Intercomparison performed in the 1990s (Goodison et al., 1998;Yang et al., 1998b;Yang et al., 1995). To address this issue, adjustments (or transfer functions) were developed to correct the undercatch for manual precipitation measurements. These adjustments were typically a function of wind speed and precipitation type, because manual precipitation measurements are typically accompanied with manual precipitation type observations (e.g. Goodison, 1978;Groisman et al., 1991;Yang et al., 1998a;Yang et al., 1999;Yang et al., 2005). More recently, transfer functions for automated precipitation measurements have been derived as a function of wind speed and air temperature (e.g. Kochendorfer et al., 2017b;Wolff et al., 2015). Transfer functions developed for automated measurements are also applied over shorter time periods, such as 30 – 60 min, whereas manual measurements are mainly adjusted per observation, with a typical observation period of either 12 or 24 hours.

~~The previous WMO Solid Precipitation Intercomparison was performed in the 1990s (Goodison et al., 1998).~~ Since ~~then~~the previous WMO Solid Precipitation Intercomparison (Goodison et al., 1998), many new automated sensors designed to measure solid precipitation have become available. In addition to not requiring a human observer, which allow~~sing~~ them to be deployed in remote locations, automated measurements can be recorded at higher frequencies and can be used to monitor precipitation continuously throughout a storm. Instead of recording precipitation every 12 or 24 h, precipitation accumulation and rate can be accurately monitored in real time. The primary types of automated precipitation gauges available today are heated tipping-bucket gauges, weighing type gauges, and non-catchment gauges. Non-catchment type precipitation sensors, which typically monitor the intensity of precipitation using optical sensors, can also be used to record precipitation type. They can detect very low precipitation rates, but can also suffer from inaccuracies in measuring the intensity of solid precipitation over shorter time intervals (e.g. 30 minutes) due to variability in hydrometeor size, fall velocity, and density (Roulet et al., 2016). In addition, it is difficult to validate or calibrate a non-catchment type precipitation gauge. Heated tipping-bucket precipitation gauges are tipping bucket gauges equipped with heaters to melt solid precipitation collected in the gauge inlet, allowing the tipping mechanism to measure liquid precipitation as discrete tips. Weighing precipitation gauges are also used to monitor all phases of precipitation, and they function by monitoring the total mass of precipitation collected below an inlet of known area. Weighing gauges should be serviced with antifreeze and oil to melt collected precipitation and inhibit the evaporation of antifreeze and precipitation. Unlike heated tipping bucket precipitation gauges, precipitation collected in a weighing gauge does not need to be heated and melted prior to measurement, and weighing gauges are typically measure very light and very heavy precipitation more accurately than tipping bucket gauges. However, as the capacity, resolution, and measurement frequency increase, all weighing gauges are limited in the smallest amount of precipitation they are able to resolve, and their ability to discern precipitation from measurement noise. This limitation is especially important for snowfall, which is often associated with very low precipitation rates. Another shortcoming of weighing gauges is their limited capacity, and the need to replace and dispose of oil and antifreeze every time they are serviced.

To update the findings of the previous WMO Solid Precipitation Intercomparison, which was focussed mainly on manual observations (Goodison et al., 1998), and evaluate many of the automated precipitation measurements that are now in use, the~~A new~~ WMO Solid Precipitation Intercomparison Experiment (WMO-SPICE) was initiated in 2010. The goal of WMO-SPICE was to study and correct the effects of wind-induced errors on automated solid precipitation measurements, and also to evaluate new and existing precipitation and snow depth sensors in different configurations and climate regimes. Because snowfall and precipitation measurement methods vary from one region or country to another, and also because errors and biases in these measurements vary as a function of climate, meteorology, and local topography (e.g. Kochendorfer et al., 2017a, hereafter K2017a), with a given sensor potentially well-suited for one site but poorly for another, one of the goals of WMO-SPICE was to include as many countries and testbeds in the intercomparison as possible. In addition to the preferred national precipitation measurement systems evaluated at most testbeds, many of the WMO-SPICE testbeds included a common set of measurements; this included a reference weighing precipitation gauge shielded in a double fence (the Double Fence Automated Reference, or DFAR), air temperature measurements, wind speed measurements, and optical precipitation monitors. The DFAR was modelled after the Double Fence Intercomparison Reference (DFIR), which was the manual reference measurement used in the ~~last~~previous WMO Solid Precipitation Intercomparison (Goodison et al., 1998), utilizing the same design for the two large, concentric wood shields that surrounded the gauge and inner shield. The DFAR is described in more detail in Ryu et al. (2016;2012).

Past WMO-SPICE related work included ~~Measurements from the Spanish WMO-SPICE site were used to~~ the development of adjustments for tipping bucket gauge measurements from the Spanish WMO-SPICE site (Buisán et al., 2017). Measurements from two WMO-SPICE test sites that pre-date the ~~recent intercomparison~~actual WMO-SPICE period ~~have been~~were also used to describe and correct wind-induced undercatch for different types of wind shields (Kochendorfer et al., 2017b;Wolff et al., 2015). In addition, measurements from eight different WMO-SPICE sites ~~during the intercomparison~~ were used to derive multi-site adjustments for single-Alter shielded and unshielded weighing gauges (K2017a)~~(Kochendorfer et al., 2017a)~~. The~~se~~ results of K2017a indicate that despite some climate- or site-specific biases, multi-site adjustments (or transfer functions) can be used to effectively minimize the wind-induced undercatch of solid precipitation. ~~In addition to the h~~Host-provided measurements were used to derive the K2017a ~~Kochendorfer et al.(2017a) m~~multi-site single-Alter shielded and unshielded precipitation transfer functions.~~,~~ WMO-SPICE also included several manufacturer-provided weighing precipitation gauges for evaluation, ~~and~~ in addition to ~~also~~ weighing gauges tested within other types of shields for more specific national and scientific interests. These previously unevaluated measurements ~~were~~ have been processed ~~-~~using standardized methods developed and implemented by WMO-SPICE, allowing for both the creation of new transfer functions and the evaluation of existing transfer functions derived from independent measurements.

The goal of this work is to test and recommend transfer functions for all of the previously unevaluated WMO-SPICE weighing gauges. In the present study, transfer functions were developed and tested using WMO-SPICE measurements from several types of weighing gauges and shields. These~~,~~ includ~~e~~~~ing~~ gauge types that have never been intercompared before, and for which no other adjustments are currently available. Previously ~~-~~derived adjustments were also applied to these measurements to test their applicability and efficacy for each of the gauge/shield types under evaluation, and also to test the hypothesis that for a given shield type, the same ~~corrections~~ adjustments can be used to minimize wind-induced errors for different types of ~~heated~~ weighing precipitation gauges of a similar size and shape. Our hypothesis is that the type of shield (or the lack of a shield) is the primary determinant of undercatch. The alternative hypothesis is that every type of weighing precipitation gauge requires its own transfer function. ~~Assuming that the type of shield (or the lack of a shield) is the~~

primary determinant of undercatch, wWe used these new WMO-SPICE measurements to test both hypotheses by comparinge transfer functions derived specifically for these each gauge and shields with more generic transfer functions derived previously using other gauges and measurements. WMO-SPICE included several manufacturer-provided unshielded and single-Alter shielded weighing gauges for which no transfer functions had previously been derived or tested. For these unshielded and single-Alter shielded gauges, the performance of newly derived transfer functions were compared to the performance of transfer functions from K2017a(Kochendorfer et al., 2017a), which were derived using host-provided WMO-SPICE measurements from eight different testbeds. WMO-SPICE also included weighing gauges within the larger double-Alter, Belfort double-Alter, and Small Double Fence Intercomparison Reference (SDFIR) shields. For these larger shields, new transfer functions derived from these WMO-SPICE measurements were compared to transfer functions derived from measurements that predate WMO-SPICE (Kochendorfer et al., 2017b, hereafter K2017b). Based on all these evaluations, specific transfer functions are rRecommended ations for all of the weighing gauges and wind shields included in WMO-SPICEare made based on these evaluations.

2 Methods

2.1 Intercomparison overview

Measurements from all of the WMO-SPICE sites wherefor which weighing precipitation gauge measurements were available for evaluation are included in this study. All of the sites included in this evaluation hadused a DFAR as the, which was used to record the reference precipitationconfiguration. The reference precipitation measurements recorded at these WMO-SPICE sites were compared to simultaneous measurements from the gauges under evaluation. Catch efficiencies, defined as the ratio of accumulated precipitation reported betweenby a gauge under test andto that reported by the DFARreference configuration, were calculated. Using thesethe computed catch efficienciesey measurements and concurrent measurements of air temperature and wind speed, transfer functions were created for the weighing gauges under evaluation. The same set of measurements was also used to evaluate independently derived transfer functions from K2017a and K2017b. Based on the results of these evaluations, transfer functions are recommended for all of the weighing gauges included in the study.

-2.21 Precipitation mMeasurements

Weighing gauges and shield configurations tested at six of the WMO-SPICE testbeds (Fig. 1) are included in this study. These sites are described in more detail in K2017a. Theyy gauges were all evaluated during 2013-2015, with measurements during the two winter seasons (Oct 1 – April 30) from this period considered in the present analysis. Measurements were recorded at either 1-min or 6 s intervals, and transferred to a central database at the National Center for Atmospheric Research (NCAR) in Boulder, CO. The Double Fence Automated Reference (DFAR), which was defined as the ~~working~~ automated reference for WMO-SPICE, consisted of ~~a~~ either a OTT Pluvio^2 or a Geonor T-200B3 within a DFIR shield, and was used as the reference for all of the measurements evaluated here (Nitu, 2012). The DFIR shield ~~has been~~ as described ~~in detail~~ by ~~Kochendorfer et al. (2017b) and~~ Goodi~~n~~son et al. (1998), ~~and~~ comprises two concentric, octagonal fences constructed out of 1.5 m long wooden lath, with the outer shield having a diameter of 12 m, and the inner shield having a diameter of 4 m. For the DFAR, a~~A~~t the centre of the inner shield~~,~~ the weighing gauge is installed in a single-Alter shield. Typically the top of the single-Alter shield and the inlet of the weighing gauge within the DFAR are at a height of 3 m, but at Weissfluhjoch and Haukeliseter they were installed higher than this (at 3.5 and 4.05 m, respectively) to prohibit drifting snow from burying the shield and gauge.

The weighing gauge models included in this study are detailed in Table 1, and are the Sutron TPG, MRW500, TRwS 405, T-200-MD3W 1500 mm Geonor, T-200-B 600 mm Geonor, and Pluvio^2. Shielding configurations of these gauges are listed in Table 2. The TRwS 405 (MPS Systém, TRwS 405), provided by the manufacturer, has a heated 400 cm^2 orifice, a 750 mm capacity, and used~~s~~ a strain gauge type load cell to measure the mass of precipitation accumulated within its bucket (Table 1). It was provided by the manufacturer without a wind shield and tested at both the Haukeliseter and Marshall testbeds (Table 2, Fig. 2a). The MRW500 (Meteoservis, MRW500, Czech Republic), with a heated 500 cm^2 orifice and a~~n~~ -9~~14~~00 or 1800 mm capacity (with or without antifreeze), also employs strain gauges for weight measurement (Table 1), and was tested at ~~both~~ the Marshall and Bratt's Lake testbeds. Both an unshielded (Fig. 2b) and a shielded gauge (Fig. 2c) were installed at each site (Table 2), -with the small, manufacturer-provided shield constructed out of fixed metal slats ~~(similar to a Tretyakov shield)~~ and attached to the same base as the gauge. An unshielded and a single-Alter shielded (Fig. 2d) Total Precipitation Gauge (TPG) provided by Sutron (Sutron, TPG, Virginia~~A~~, USA) were tested at the Marshall testbed (Table 2). The Sutron TPG uses a load cell to quantify the amount of accumulated precipitation, has a 914 mm capacity, an 8" diameter inlet (20.32 cm diameter, 324.3 cm^2 area), and was provided with a heater (Table 1). The T-200-MD3W (1500 mm Geonor) from Geonor (Geonor, Oslo, Norway) was tested at Marshall, Bratt's Lake, Weissfluhjoch, and Caribou Creek (Table 2), ~~and~~ but only the gauges at Marshall and Weissfluhjoch were~~was~~ provided with a heater (Table 1). -The 1500 mm Geonor is based on the same design as the 600 mm and 1000 mm Geonor T-200B3 3-wire, vibrating-wire gauges, but it has a taller cover, taller bucket with increased capacity, and different vibrating wire transducers.

The double-Alter shield (Fig. 2d), which consists of a single-Alter shield surrounded by a second, larger (2.4 m in diameter) row of 40 cm long slats, was tested at CARE (Centre for Atmospheric Research Experiments) with an OTT Pluvio^2 (OTT Hydromet, Pluvio^2, Kempten, Germany; Tables 1 and 2) and at Marshall with a Geonor T-200B3 (3-wire, 600 mm capacity, T200B, Geonor Inc., Oslo, Norway; Tables 1 and 2). Both gauges included inlet heating, with the Pluvio^2 at CARE using the manufacturer-provided heater and the Geonor at Marshall using the US Climate Reference Network heater (described in

NOAA Technical Note NCDC No. USCRN-04-01). Based on the additional shielding provided by the double-Alter shield and also past studies (K2017b)~~(Kochendorfer et al., 2017b)~~, we expect the double-Alter shield to perform better than the single-Alter shield, and to accumulate greater than 50% of the DFAR even in snowy and windy conditions. The Belfort double-Alter shield (Fig. 2f), which has the same sized footprint as the standard double-Alter shield, but with a lower porosity (30% vs the standard double-Alter porosity of 50%) and longer slats (46 cm long for the inner shield, and 61 cm long for the outer shield) that do not taper like the double-Alter, ~~and a lower porosity (30% vs the standard double-Alter porosity of 50%),~~ was also tested at both CARE and Marshall (Table 2). ~~. Another important distinction is that t~~The standard double-Alter slats also rotate freely, while the Belfort double-Alter slats employ springs to limit their travel within ~~45~~30° of the vertical. Like the standard double-Alter, the Belfort double-Alter was tested with a Pluvio^2 at CARE and a 600 mm Geonor T-200B3 at Marshall (Table 2). Previous to this study, this shield had only been evaluated at Marshall (K2017b). Based on that ~~past~~ study, we expect gauges within this shield to accumulate almost as much precipitation as the DFAR, with typical catch efficiencies greater than 0.8 in snowy and windy conditions.

The ~~small DFIR (~~SDFIR~~)~~ shield (Fig. 2g), which is 2/3 the size of standard DFIR shield and was designed to be more easily constructed out of commonly ~~-~~sized North American lumber, was tested only at the Marshall site (Table 2). Like the standard DFIR shield, the SDFIR ~~configuration~~ comprises three concentric ~~wind~~ shields. The wooden laths on the two-outermost concentric shields were 1.2 meters long, and the diameters of the two outer shields were 8.0 meters and 2.6 meters. The height of the inner wooden shield was 10 cm lower than the outer shield. A standard single-Alter shield, mounted at the same height as the gauge inlet and 10 cm lower than the inner wooden shield, was mounted around the gauge. Based on its design and a past study (K2017b), we expect this shield to be almost indistinguishable from the DFAR, except in high winds and snowy conditions, where it will typically accumulate about 90% as much precipitation as the DFAR. ~~Table 1 summarizes the different gauges and shields that were tested, and includes some statistics describing the available measurements.~~

2.3~~2~~ Wind speed and direction

Wind speed measurements from each site were used to create 30-minute-average wind speeds. Because transfer functions developed from both the gauge height and the 10 m height wind speed were desired, and not every site included wind measurements at both heights, the available 30-min measurements and the logarithmic wind profile were used to determine either the gauge height or 10 m height wind speed, when necessary. These methods ~~used to do this~~ are described in detail in K2017a.~~Kochendorfer et al.(2017a).~~

~~Quality control of the wind speed measurements included removal of wind speeds equal to zero, removal of 'stuck' wind speeds at the Haukeliseter site, where the wind speed would occasionally remain unchanged for several hours at a time, and removal of 10 m height wind speeds at the CARE site that were less than the gauge height wind speed. At the Marshall site,~~

a composite 10 m height wind speed was created using two separate 10 m height wind speed sensors, as the sensors were observed to shadow each other from either due north or due south. All of these steps are described in detail in Kochendorfer et al.. For the TRwS 405 precipitation gauge at Haukeliseter, an additional screening for wind direction was performed based on the gauge's position relative to the DFAR; p.; Pprecipitation measurements with wind directions between 115° and 140° were excluded from the analysis due to wind shielding by the DFAR.

2.34 Quality cControl and sSelection of 30-min pPeriods of pPrecipitation

The methods applied to thePrecipitation measurements were recorded at either 6-s or 1-min intervals (Table 1). These available 6-s and 1-min data includewere subject to the following quality controls: a range filter, to remove values that exceeded the capacitywere above or below realistic limits, or outside of the operational limits for of a giventhe gauge; a 'jump' filter, to remove sudden changes in accumulation exceeding a specified threshold; and a Gaussian filter, to remove high-frequency noise. For use in developing transfer function developments, the resultant 1-min (all 6-s data were aggregated to 1-min), quality-controlled measurements were then used to create 30-min datasets that included only periods of precipitation. To exclude noise and light precipitationensure that precipitation was occurring, a minimum threshold of 0.25 mm of precipitation as measured by the DFAR was used. In addition, based on independent optical precipitation detector measurements, precipitation had to occur for at least 60% of every 30-min period (18 min). Methods created within the WMO-SPICE project to smooth and quality control precipitation measurements are also detailed elsewhere (K2017a, Reverdin, 2016).

Following K2017aochendorfer et al. , a minimum precipitation threshold was identified for every gauge or shield under evaluation to help create an unbiased pool of measurements available for analysis. All precipitation measurements below the minimum threshold determined for each specific gauge/wind shield configuration were excluded from the analysis. The minimum precipitation thresholds were calculated by multiplying the minimum DFAR precipitation of 0.25 mm by the median catch efficiency of the gauge under test, using only solid precipitation measurements (mean T_{air} < -2 °C) with high winds (5 m s^{-1} < U_{10m} < 9 m s^{-1}). Also following K2017aKochendorfer et al. , a maximum catch efficiency threshold was calculated as 1.0 plus three times the standard deviation of the catch efficiency of the gauge under test, and all measurements exceeding the relevant maximum catch efficiency threshold were excluded from the analysis.

2.54 Transfer Function Models

The Kochendorfer et al. EEquations 1 3 (hereafter KOC-Eq. 3) and 4 (hereafter KOC-Eq. 4) wasere fit to the resultant 30-minute weighing gauge measurements, following K2017a.

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

Where U is the mean wind speed, T_{air} is the mean air temperature, and a, b, and c are coefficients fit to the data. Equation 2 was also fit to the precipitation measurements, following K2017a.

$$CE = (a)e^{-b(U)} + c \qquad (2)$$

Where a, b, and c are coefficients fit to the data. ~~KOC Eq. 3~~Equation 1 was fit as a function of wind speed ~~(U)~~ and air temperature ~~(T_{air})~~, while ~~KOC Eq. 4~~Equation 2 was fit separately to solid and mixed precipitation measurements as a function of wind speed only. In the latter case, precipitation type was determined using air temperature ~~T_{air}~~, with solid precipitation defined as T_{air} < -2 °C, and mixed defined as 2 °C $\geq T_{air} \geq$ -2 °C. These specific temperature thresholds were selected to estimate precipitation type based on past evaluations of precipitation type and air temperature (K2012b, Wolff et al., 2015). For some of the gauges examined here, ~~KOC Eq. 4~~Eq. 2 unrealistically over-predicted catch efficiency at low wind speeds when insufficiently constrained by the available measurements, and in these cases a more constrained function was used to describe realistic corrections for sparser or noisier results, especially for gauges with fewer low wind speed measurements:

$$CE = (a)e^{-b(U)} + (1 - a), \qquad (3~~1~~)$$

W~~w~~here ~~U is wind speed, and~~ a, and b are coefficients fit to the data. Following K2017a~~ochendorfer et al.~~, ~~both~~ transfer functions were developed for both gauge height and~~wind~~ ~~speed corrections and~~ 10 m height wind speeds ~~speed corrections were developed for all of the transfer functions tested~~.

2.6~~5~~ Maximum w~~W~~ind s~~S~~peed t~~T~~hreshold

For the application of transfer functions, a maximum wind speed threshold (U_{thresh}) above which the transfer function should not be applied was determined based on a visual assessment of the ~~KOC Eq. 3~~Eq. 1 transfer function fit to the available measurements. This was done by viewing the catch efficiency function of wind speed and air temperature superimposed over the actual measurements, and identifying the wind speed above which all temperature ranges below 2 ˚C were not generally well represented by the available measurements. The same threshold was applied to the ~~Eq. 1~~Eq. 2 and ~~the KOC Eq. 4~~Eq. 3 transfer functions ~~as well~~. In practice, when the wind speed is above the maximum wind speed threshold, the wind speed should be forced down to the maximum wind speed threshold to adjust the precipitation. A diagram describing the effects of the maximum wind speed threshold on an example transfer function is shown in Figure 3, using the unshielded ~~KOC Eq. 4~~Eq. 2 type transfer function developed in K2017a~~ochendorfer et al.~~.

2.7~~6~~ Testing of t~~T~~ransfer f~~F~~unctions

When measurements from more than one site were available for a specific gauge or shield combination, all of the available measurements were ~~used~~ merged to create a common transfer function, and the transfer function was then tested on data from each site independently to determine the magnitude of site biases and the appropriateness of the transfer function for each individual site. For gauges that were only tested at one site, a 10-fold cross validation was relied upon to maintain some independence between the measurements used to produce and test the transfer functions. The 10-fold cross validation was performed in 10 separate iterations, using 90% of the measurements to determine the transfer function and the remaining 10% to test the transfer function. The resulting error statistics were based on the average of all ten iterations.

Errors in the adjusted measurements were estimated by applying the appropriate transfer function and comparing the results to the corresponding DFAR measurements. The errors were then used to calculate the root mean square error (RMSE), the mean bias, the correlation coefficient (r), and the percentage of 30 min events with errors less than 0.1 mm ($PE_{0.1\ mm}$). These statistics were estimated for the ~~KOC-Eq. 3~~Eq. 1 transfer functions using all of the available precipitation measurements. Following K2017a,~~or~~ the Eq. 2~~1~~ and ~~KOC-Eq. 4~~Eq. 3 transfer function~~s, the~~ error statistics were estimated by separating the datasets using the mean air temperature into liquid (T_{air} > 2 ˚C), mixed~~-~~ (2 °C ≥ T_{air} ≥ -2 °C.), and solid (T_{air} < -2 ˚C) precipitation, correcting the mixed and solid precipitation using the appropriate transfer functions, combining these results with the uncorrected liquid precipitation measurements, and comparing the results to the corresponding DFAR measurements.

A t-test with a significance level of 5% was used to evaluate the significance of differences based on the different transfer functions. When a given precipitation gauge/shield configuration was tested at more than one site, the adjusted measurements from all of the available sites were pooled together before determining the significance of measurement differences. The same test was also used to compare unadjusted measurements to adjusted measurements. ~~The liquid precipitation measurements were not corrected because warm-season precipitation measurements were not included in the WMO-SPICE dataset, and the resultant liquid precipitation wind speed catch efficiencies were not significantly different than one. However some liquid precipitation measurements were necessary for the successful creation of KOC-Eq. 3 type transfer functions, so the available liquid precipitation measurements were included both in the derivation and the evaluation of the KOC-Eq. 3 type transfer functions.~~

2.8~~7~~ Evaluation of i~~I~~ndependent t~~T~~ransfer f~~F~~unctions

In addition to developing new adjustments for all of the weighing gauges tested within WMO-SPICE, the WMO-SPICE weighing gauge measurements were also used to evaluate other independently ~~-~~derived transfer functions that ~~are~~ were available. These include the WMO-SPICE single-Alter shielded and unshielded transfer functions~~(Kochendorfer et al., 2017a)~~ derived using eight different testbeds (K2017a), and transfer functions determined from pre-SPICE measurements recorded at the Marshall testbed (K2017b)~~(Kochendorfer et al., 2017b)~~.

Single-Alter and unshielded transfer functions developed previously by K2017a ~~Kochendorfer et al.~~ ~~from~~ for WMO-SPICE host-provided weighing gauges (either Geonor T-200B3 or OTT Pluvio^2) were tested on measurements from all of the manufacturer-provided unshielded and single-Alter shielded gauges evaluated within WMO-SPICE. The hypothesis behind this testing is that the response of a gauge to wind speed and air temperature is more sensitive to wind-shielding (or a lack thereof) than to the gauge type. Although the transfer functions from K2017a ~~ochendorfer et al.~~ were not developed for these specific gauges, they include measurements from eight different sites.~~,~~ ~~and~~ T~~t~~herefore, they may be more robust and universally applicable than transfer functions developed from measurements at the limited number of sites where a specific manufacturer-provided gauge was tested. A robust transfer function should arguably be developed from measurements representing a wide variety of precipitation types and wind speeds, as any transfer function is only valid for the range of conditions represented during its development. In addition, as shown in K2017a~~ochendorfer et al.~~ , ~~significant~~ site biases do exist, indicating that the use of data from several sites for the creation of a transfer function is preferable to the use of data from just one or two sites. Because of this, it is possible that a generic transfer function developed using data from several sites may be more universally applicable to a manufacturer-provided weighing gauge than the gauge-specific transfer function. For the double-Alter, Belfort double-Alter, and the SDFIR shielded gauges, which were not tested as broadly within WMO-SPICE as the single-Alter shielded and unshielded gauges, the WMO-SPICE measurements recorded at Marshall and CARE were used to evaluate transfer functions created by K2017b ~~ochendorfer et al. (2017b)~~ using measurements recorded at the Marshall testbed before WMO-SPICE began.

~~For the sake of simplicity and ease of presentation, for the testing of previously-derived transfer functions, only results from the KOC Eq. 3 version of the 'universal' transfer function with the gauge height wind speed are presented here. The different types of corrections tested in Kochendorfer et al. were all shown to produce similar results. For example, the Kochendorfer et al. corrections were developed as functions of both the gauge height wind speed and the 10 m wind speed. In addition, one set of transfer functions explicitly included T_{air} (KOC Eq. 3), and the other set included different functions for mixed and solid precipitation (KOC Eq. 4). However, the resultant errors were similar for all of these different types of transfer functions, and the performance of one version of the Kochendorfer et al. universal transfer function tested on these new gauges can be considered representative of the performance of all the different types of transfer functions from Kochendorfer et al. .~~

2.9~~8~~ T~~Uncertainty of T~~ransfer f~~F~~unction uncertainty~~s~~ ~~with~~ and w~~W~~ind s~~S~~peed

Errors in the transfer functions were ~~further~~ evaluated further as a function of wind speed~~. This was done~~ by ~~calculating~~ computing RMSE values from the available catch efficiencies after binning by wind speed (1 m s^{-1} bins). For each wind speed bin, the RMSE values of the ~~transfer function~~ adjustments were calculated from differences between the transfer function catch efficiencies and the measured catch efficiencies. In addition, RMSE values of the adjusted catch efficiencies were estimated by applying the appropriate transfer function to the measurements and calculating the resultant error in the catch efficiency~~; ideally the adjusted catch efficiency should be equal to one, so the difference between the adjusted catch efficiency and one was used to quantify errors in the adjusted catch efficiency~~. This evaluation was limited to solid precipitation (T_{air} < -2 °C) measurements for ease of presentation, and the ~~KOC Eq. 3~~Eq. 1 type transfer functions were used for all of the ~~different~~ gauge configurations examined.

3 Results and Discussion

3.1 Transfer function types

Using the methods described above, transfer functions were created for the weighing gauges and ~~wind~~shields ~~included~~ tested in WMO-SPICE (Table 2). For all of the gauges and shields tested, the ~~KOC Eq. 3~~Eq. 1 transfer function was fit to the catch efficiency measurements using the measured wind speed and air temperature. In addition, either the ~~KOC Eq. 4~~Eq. 2 or the ~~Eq. 1~~Eq. 3 type adjustment was fit to the mixed (2 °C ≥ T_{air} ≥ -2 °C~~.~~) and solid (T_{air} < -2 ˚C) precipitation measurements as a function of wind speed. These transfer functions were created for both the gauge height wind speed and the 10 m height wind speed. Statistics describing the transfer function errors were calculated based ~~from~~ on the differences between the adjusted precipitation measurements and the DFAR precipitation measurements. In addition, transfer functions available from other studies were tested on these precipitation measurements, and errors in the uncorrected measurements were also described.

Based on the results of t-tests used to compare different types of transfer functions, no significant differences were found between measurements adjusted using the Eq. 1, Eq. 2, or Eq. 3 transfer functions. In addition, there were no significant differences between measurements adjusted using the 10 m height wind speeds or the gauge height wind speeds. For the sake of simplicity, and to focus on more significant results, a comparison of results for all transfer functions for each gauge tested is not included here. Examples of such comparisons are available in K2017a, and like the present study, these examples demonstrate that the different types of transfer functions fitted to the same data performed similarly. Newly developed and previously existing transfer functions are, therefore, compared using Eq. 1 transfer functions with gauge height winds, with these results found to be representative of all of the different types of transfer functions tested.

3.2~~1~~ ~~Sutron TPG~~Unshielded gauges

Unshielded Sutron, MRW500, and TRwS 405 weighing gauges were tested at several different WMO-SPICE testbeds. Separate transfer functions were developed for all of the different gauge~~s~~ types. In addition, an independent pre-existing ‘universal’ transfer function (K2017a)~~(Kochendorfer et al., 2017a)~~ was tested on all of the unshielded gauge measurements. Statistics describing errors in the adjusted and unadjusted measurements are shown in Fig. 4. Improvements in the adjusted bias (Fig. 4b) and $PE_{0.1\ mm}$ (Fig. 4d) values were more notable than improvements in the RMSE (Fig. 4a) or correlation coefficients (Fig. 4c). In addition, the ‘universal’ transfer function performed as well as the transfer functions fitted specifically to each gauge type. This indicates that all of the unshielded gauges tested, including those from K2017a, will accumulate similar amounts of precipitation when exposed to the same weather. It also indicates that small differences in the shapes of these gauges do not affect the relationship between their catch efficiency and wind speed. Using t-tests, the measurements adjusted with the 'universal' equation were compared to measurements adjusted with each custom equation. For the unshielded Sutron, MRW500, and TRwS gauges the t-tests indicated that there were no significant differences between measurements adjusted using gauge-specific transfer functions and the unshielded K2017a transfer function. Two Sutron TPG gauges were tested in WMO-SPICE, and both were installed at the Marshall testbed. One was unshielded and the other was provided with a single-Alter shield. Separate transfer functions were developed using measurements from each configuration. In addition, the more universal Kochendorfer et al. transfer functions for weighing gauges in unshielded and single-Alter configurations were tested using measurement data from the Sutron gauge in the corresponding configuration.

3.1.1 Unshielded Sutron TPG

Transfer functions were developed from the 30-min measurements as a function of wind speed for both mixed and solid precipitation (KOC-Eq. 4), and also as function of both wind speed and air temperature (KOC-Eq. 3). Because this gauge type was only available at one site, the RMSE (Fig. 4a), bias (Fig. 4b), correlation coefficient (Fig. 4c), and $PE_{0.1\ mm}$ (Fig. 4d) were estimated using 10-fold cross-validation (Section 2.6). One of the 'universal' unshielded transfer functions developed by Kochendorfer et al. was also tested on these independent measurements, and performed as well as the corrections developed specifically for this gauge-shield combination (Fig. 4). As a result of this, our recommendation is to use the more 'universal' corrections from Kochendorfer et al. for the unshielded Sutron TPG; the resultant transfer function coefficients developed for this specific gauge are therefore not included here.

3.1.2 Single Alter Sutron TPG

The different transfer functions developed and tested for the single-Alter (SA) shielded Sutron TPG performed similarly (Fig. 5). Like the unshielded Sutron TPG, the custom transfer function developed for this gauge did not perform significantly better than the 'universal' transfer function from Kochendorfer et al. . Because the 'universal' functions from Kochendorfer et al. were developed from single-Alter gauges at eight separate sites and included more measurements and a wider range of wind speeds and precipitation types, they are more widely applicable than the transfer functions developed specifically for the Sutron with the single-Alter gauge using measurements from a single site.

3.2 Meteoservis MRW500

Meteoservis MRW500 weighing gauges were tested at the Bratt's Lake and Marshall testbeds, with unshielded and shielded gauges tested at each site. Error statistics were calculated for different forms of the transfer functions. The 'universal' unshielded transfer function was applied to the unshielded gauge measurements. Although the shielded MRW500 weighing gauge was provided with a custom Tretyakov-type shield, which was smaller than a standard single Alter shield and constructed out of metal slats mounted at a fixed angle (Fig. 1e), the 'universal' single-Alter shielded adjustment was tested on this configuration as previously-derived adjustments for an automated gauge within a comparable shield were available.

3.2.1 Unshielded MRW500

The resultant error statistics were similar for all the transfer functions developed from these data, with no significant differences between the KOC-Eq. 3 and KOC-Eq. 4 type adjustments for the gauge height wind speed and the 10 m wind speed (Fig. 6). All of the error statistics also showed significant improvement in the corrected measurements compared to the uncorrected measurements (Fig. 6). The universal unshielded function derived in Kochendorfer et al. performed as well as the unshielded functions developed specifically for this gauge. We therefore recommend using the universal transfer function derived in Kochendorfer et al. (2017a), rather than the transfer function fit to these specific gauge measurements, as there appears to be no significant advantages to using the transfer function coefficients derived specifically for this type of unshielded weighing gauge.

3.2.2 Shielded MRW500

All of the different transfer functions developed and tested using the Marshall and Bratt's Lake measurements effectively reduced the RMSE and bias values, and increased the correlation coefficients and $PE_{0.1\ mm}$ (Fig. 7). In addition, the 'universal' SA transfer function was tested with this gauge. Although the shielded MRW500 weighing gauge was provided with a custom Tretyakov-type shield, which was smaller than a standard single Alter shield and constructed out of metal slats mounted at a fixed angle (Fig. 1c), the 'universal' single-Alter shielded adjustment was tested on this configuration as previously-derived adjustments for an automated gauge within a comparable shield were available.The resultant RMSE values were similar to the RMSEs for the transfer function developed by Kochendorfer et al. (Fig. 7a), and the $PE_{0.1}$ values were actually improved by the use of the SA transfer function (Fig 7d). However, the negative bias resultant from the application of the universal single-Alter correction indicates that this gauge was generally under-corrected by the single Alter transfer function, particularly at Bratt's Lake (Fig. 7c). For this reason, the new transfer function coefficients provided in Table 2 should be used to correct this gauge-shield combination, rather than the 'universal' single Alter correction from Kochendorfer et al. .

For the exponential wind speed transfer functions developed separately for solid and mixed precipitation, Eq. 1 was developed and used for the shielded MRW500 because KOC-Eq. 4 resulted in unreasonably high transfer function results as the wind speed approached 0 m s^{-1}. This result are probably more closely linked with the scarcity of low wind speed events from these two sites and random errors in the low wind speed catch efficiencies rather than the gauge configuration. In addition, although an exponential fit was used for these data because it is more realistic at high wind speeds (where a linear fit would predict negative catch efficiencies), for the data available, the shielded MRW500 catch efficiency responded quite linearly to wind speed. Unfortunately there were insufficient high-wind data available from the Bratt's Lake and Marshall sites to evaluate this shield at higher winds, where the catch efficiency would presumably asymptote at minimum value that was greater than zero.

3.3 Unshielded MPS TRwS 405

The MPS TRwS 405 gauge relies on a strain gauge to measure the mass of water accumulated within it. It was provided without a shield, and was tested at the Haukeliseter and Marshall testbeds. Transfer functions were developed specifically for this gauge by combining the Marshall and Haukeliseter measurements together, and the efficacy of the transfer functions was tested by applying the corrections to the measurements and calculating statistics using the resultant errors at each site. The RMSE, bias, correlation coefficients, and $PE_{0.1\ mm}$ values (Fig. 8) indicated that the Kochendorfer et al. 'universal' unshielded correction performed as equally well as the unshielded correction fit specifically to this gauge, and because of this the gauge-specific transfer function coefficients developed here are neither recommended or shown.

3.3 Single-Alter shielded gauges and the shielded MRW500 gauge

Transfer functions for the single-Alter shielded Sutron and 1500 mm Geonor measurements were developed and tested, in addition to transfer functions for the MRW500 gauge within the small MRW500 shield (Fig. 5). These results are discussed in more detail below.

3.3.1 Single-Alter shielded gauges

At Marshall, the Sutron gauge was equally well adjusted using both the custom and the 'universal' transfer functions (Fig. 5 a-d), with a t-test confirming that measurements adjusted using both methods did not differ significantly from each other. This indicates that undercatch in the Sutron TPG responds to wind and air temperature in the same manner as the single-Alter shielded gauges used in K2017a.

The 1500 mm Geonor was tested at four different sites, and results, significant differences were found between the transfer functions fit to these measurements and the K2017a 'universal' transfer function. Figure 5 a-d supports this finding, showing differences between error statistics for measurements adjusted using the Eq. 1 and Univ. Eq. 1 transfer functions. These differences are most apparent in the bias and RMSE values estimated at Weissfluhjoch. 3.4 Single-Alter shielded Geonor T-200-MD3W (1500 mm)

The Geonor T-200-MD3W is based on the same design as the 600 mm and 1000 mm Geonor T-200B3 3-wire, vibrating-wire gauges, but it has a taller cover, taller bucket with increased capacity (1500 mm), and different vibrating wire transducers. This gauge was tested at the Marshall, Bratt's Lake, Weissfluhjoch, and Caribou Creek sites. For reasons that are not well understood, the catch efficiency of the single-Alter shielded 1500 mm Geonor 1500 mm gauges at Weissfluhjoch and Caribou Creek did not did not decrease significantlydecrease with wind speed as much as at the other sites. The 1500 mm Geonor at Caribou Creek was installed near low trees, which may have sheltered the gauge from the wind from some directions; however, a lack of wind direction measurements available in the WMO-SPICE event dataset from this site prohibits confirmation evaluation of this hypothesis. The Weissfluhjoch site whas previously been demonstrated to be less sensitive to wind than other sites (K2017a)(Kochendorfer et al., 2017a) , for reasons that are were also difficult to understand or confirm. The relative insensitivity of the catch efficiency to wind speed at these two sites is apparent from the fact that even using the custom 1500 mm transfer functions to correct the measurements, there was no improvement in the RMSE at either of these sites (Fig. 9a), and the bias (Fig. 9b) showed that the 1500 mm Geonor transfer functions (derived from the measurement data from this specific gauge configuration) over-corrected the measurements at Caribou Creek and Weissfluhjoch. Using the universal transfer function, the 1500 mm Geonor measurements from Weissfluhjoch and Caribou Creek werewere further over-corrected (Fig. 59b), whereas using a custom 1500 mm Geonor Eq. 1 adjustment, the resultant bias was much smaller.

The significant differences between the 'universal' transfer function and the transfer functions fit to the 1500 mm Geonor measurements are can be attirbutedattributed mainly to differences in the available measurements used to create the universal transfer function and the 1500 mm Geonor transfer functions.; Thirty-five percent35% of the available 1500 mm Geonor measurements were recorded at Weissfluhjoch, where the catch efficiency for all the gauges did not drop offdecrease with wind speed to the same degree as most of the other sites (Kochendorfer et al., 2017a). When the number of Weissfluhjoch measurements contributing to the 1500 mm transfer functions was artificially reduced, for example, the resultant 1500 mm transfer function was similar to the universal single-Alter adjustment. At the Marshall and Bratt's Lake sites, where the catch efficiency decreased with wind speed as expected, the universal transfer function performed better than the gauge-specific transfer function.

In general, the different error statistics generated from the 1500 mm Geonor measurements indicate that this gauge was subject to more noise than the host-provided gauges used to develop the universal single-Alter transfer functions in K2017a. For example, at Marshall, the 1500 mm single-Alter Geonor RMSE values were about 0.25 mm and the $PE_{0.1\ mm}$ values were about 60%, while for the 600 mm Geonor at this same site, the RMSE values were about 0.15 mm, and the $PE_{0.1\ mm}$ values were about 70% (K2017a). The increased capacity of the 1500 mm Geonor gauge appears to be associated with a decrease in sensitivity, which may make it more difficult to accurately measure snowfall events and distinguish between gauge noise and light precipitation.Because there is no obvious physical explanation for why the Geonor 1500 mm gauge would have a higher catch efficiency than the 600 mm or 1000 mm Geonor gauges (the collecting area is the same for each configuration), and the reasons for the relatively poor performance of the universal function may be due more to the specific population of 1500 mm Geonor measurements available within this intercomparison, the universal transfer function is still recommend over the custom transfer functions derived from the 1500 mm Geonor measurements. At the Marshall site, where the catch efficiency decreased with wind speed as expected, the universal transfer function performed better than the gauge-specific transfer functions, and at Caribou Creek, the differences between the custom 1500 mm adjustment and the universal single-Alter adjustment were small.

3.3.2 Shielded MRW500In general, the different error statistics generated from the 1500 mm Geonor measurements indicate that this gauge was subject to more noise than the host-provided gauges used to develop the universal single-Alter transfer functions in Kochendorfer et al. . For example, at Marshall, the 1500 mm single-Alter Geonor RMSE values were about 0.25 mm and the $PE_{0.1\ mm}$ values were about 60%, while for the 600 mm Geonor at this same site, the RMSE values were about 0.15 mm, and the $PE_{0.1\ mm}$ values were about 70%.

The shielded MRW500 weighing gauge was provided with a custom shield, which was smaller than a standard single-Alter shield and was constructed out of metal slats mounted at a fixed angle (Fig. 2c). Because no previously derived transfer function was available for this specific shield, the ‘universal’ single-Alter shielded adjustment was tested on this configuration. The resultant RMSE values were similar to each other (Fig. 5e), and the $PE_{0.1}$ values were improved by the use of the ‘universal’ SA transfer function (Fig 5h). However, the negative bias resultant from the application of the universal single-Alter adjustment indicates that this gauge was generally under-corrected, particularly at Bratt’s Lake (Fig. 5f).

Exponential wind speed transfer functions were also developed separately for solid and mixed precipitation. After analysis, Eq. 3 was developed and used for the shielded MRW500 gauge because Eq. 2 predicted unreasonably high catch efficiencies as the wind speed approached 0 m s^{-1}. This result is probably more closely linked with the scarcity of low wind speed events and random errors in the low wind speed catch efficiencies from these two sites, rather than with the specific gauge configuration. In addition, although an exponential fit was used for these data because it was more realistic at high wind speeds (where a linear fit would predict negative catch efficiencies), for the data available, the shielded MRW500 catch efficiency responded quite linearly to wind speed. Unfortunately, there were insufficient high-wind data available from the Bratt’s Lake and Marshall sites to evaluate this shield at higher winds, where the catch efficiency would presumably reach an asymptote at a minimum value greater than zero. Transfer function coefficients for the MRW500 are available in the Supplement (Table S1).

3.45 Double-Alter, Belfort double-Alter, and SDFIR shields

Overall, measurements recorded within these larger shields were subject to smaller uncorrected biases, and also smaller errors after adjustment. For example, the RMSE values of the adjusted measurements shown in Fig. 6 were smaller than those for the single-Alter and unshielded gauges tested (Fig. 4 and 5), and the $PE_{0.1\ mm}$ values were larger. More specific results are discussed for individual shields below.

3.4.1 Double-Alter shield

The double-Alter shield was tested at both the CARE and Marshall testbeds; the CARE site had an OTT Pluvio[2] in a double-Alter shield, and the Marshall site had a 600 mm Geonor T-200B3 in a double-Alter shield. The pre-SPICE double-Alter transfer function (K2017b) (Kochendorfer et al., 2017b) performed about as well as the WMO-SPICE transfer function (Fig. 6 a-d). One thousand, three hundred and ninety-two , and the different types of transfer functions developed from the WMO-SPICE measurements all performed similarly (Fig. 10). 1392 measurements were available from the pre-SPICE Marshall measurements (Kochendorfer et al., 2017b), and only 723 measurements were available from the WMO-SPICE measurements (Table 21). However, this new WMO-SPICE correction is arguably more defensibly applicable to all double-Alter measurements than the pre-SPICE transfer function, because it was developed using measurements from two sites. Table 3 includes the resultant transfer function coefficients. However, due to the limited warm season, liquid precipitation measurements included in the SPICE datasets, use of the KOC Eq. 3 transfer functions presented here is not recommended when T_{air} is > 5 °C, as they produce unrealistically high warm-temperature precipitation catch efficiencies. If a KOC Eq. 3 type function is needed that is also applicable to warm-season measurements, we recommend using the pre-SPICE function, which was demonstrated in the testing performed here to be quite similar to the WMO-SPICE functions (Fig. 10).

3.4.26 Belfort double-Alter

The Belfort double-Alter shield, which has the same size footprint as the standard double-Alter shield, but with longer slats and a decreased porosity of about 30% relative to that of the standard double-Alter (~ 50%), was more effective at reducing undercatch than the standard double-Alter. This is demonstrated by the generally small RMSE improvements improvements of the corrected measurements over the uncorrected measurements (Fig. 6e11a), and also by the near-zero uncorrected biases for the gauges at both Marshall and CARE (Fig. 6f11b). These measurements, recorded at two separate sites, confirm the efficacy of the Belfort double-Alter shield documented by K2017b ochendorfer et al. (2017b) using measurements from a single site. In terms of data availability,Nine hundred and nineteen 919 30-min measurements were included in the present WMO-SPICE measurements transfer function (Table 21), and 1204 30-min measurements were available for the Pre-SPICE Marshall transfer function development (K2017b). Although the two datasets resulted in similar transfer functions, (as evidenced by the equivalent performance of the pre-SPICE transfer function when tested on these new measurements in Figure 11), we recommend using the transfer functions determined from the WMO-SPICE measurements, in this work, because they include measurements from two sites, and are, therefore, expected to be more broadly applicable. Like the double-Alter WMO-SPICE transfer function, the KOC Eq. 3 Belfort double-Alter transfer functions did not include many liquid precipitation events, but in this case, the resultant transfer functions were more realistic at warm temperatures, and can therefore be recommended for use in all conditions. The associated transfer function coefficients are provided in Table 4.

3.4.27 Small Double Fence Intercomparison ReferenceDFIR

Tested only at the Marshall testbed, the SDFIR was the largest wind shield testedevaluated in our study, and the uncorrected and corrected SDFIR measurements were associated with the lowest RMSE and bias values (Fig. 6i and 6j12a and 12b), and the highest correlation coefficient and $PE_{0.1\ mm}$ values (Fig. 612kc and 16l2d) among the wind shields tested. The WMO-SPICE catch efficiencies were similar to the Kochendorfer et al. (2017b)2017b SDFIR results, and provided independent validation of these pre-SPICE Kochendorfer et al. (2017b) transfer function (Fig. 612). The Kochendorfer et al. (2017b) SDFIR transfer function was developed using five years of measurement data (1508 30-min precipitation events), whereas only two winter seasons (410 30-min events) were available for the development of the WMO-SPICE transfer function. Perhaps most notably, the lack of rain events in the WMO-SPICE dataset (Table 1) resulted in unrealistically large SDFIR shielded catch efficiencies predicted by the KOC-Eq. 3 type transfer function for warm temperatures and high wind speeds. The Kochendorfer et al. (2017b) KOC-Eq. 3 type catch efficiencies were more realistic for all temperature/wind speed regimes. For this reason, the KOC-Eq. 3 type transfer function coefficients determined from the WMO-SPICE measurements are not included here, as the pre-SPICE transfer function was developed using measurements from the same gauge/shield, at the same site, over a much longer period. However, because Kochendorfer el al. (2017b) did not include KOC-Eq. 4 coefficients, the KOC-Eq. 4 transfer functions determined from the WMO-SPICE measurements, which were unaffected by the lack of warm-temperature precipitation in the WMO-SPICE measurements, are included in Table 5.

In general, the necessity of transfer function adjustments for SDFIR-shielded measurements is disputable, as the corrected measurements were only marginally not significantly better than the uncorrected measurements (Fig. 12), with only verya small improvements observable in the mean bias values (Fig. 6j12b) and the $PE_{0.1\ mm}$. values (Fig. 6l). -However, errors in the uncorrected SDFIR measurements were determined to be significantly different than zero. In addition, the different correction types all performed similarly. These corrected and uncorrected SDFIR-shielded measurements results are still also interesting, however, because they provide a good indication of the magnitude of errors when comparing well-shielded gauges. As such, tThese measurements are a good representation of the current limits in accuracy for precipitation measurements recorded using two different well-shielded gauges at the same site, both of which are well-shielded. The inferences that can be drawn from such well-shielded measurements are further emphasized below in the following section comparisonng of the different shields and adjustments. .

3.5~~8~~ Synthesis

3.5.1 Recommended transfer functions

As shown in Table 3, we recommend using the appropriate K017a transfer functions for all of the unshielded and single-Alter shielded WMO-SPICE gauges. Although the 'universal' single-Alter transfer function performed poorly on the 1500 mm Geonor gauges at Caribou Creek and Weissfluhjoch, we still recommend using it on single-Alter shielded 1500 mm Geonor gauges. This is because there is no obvious physical explanation for a higher catch efficiency for the Geonor 1500 mm gauge than the 600 mm or 1000 mm Geonor gauges (the collecting area and inlet shape are the same for each configuration). In addition, the relatively poor performance of the 'universal' single-Alter function may be due to the specific population of 1500 mm Geonor measurements available within this intercomparison. As also indicated in Table 3, the new transfer function coefficients provided in the Supplement (Table S1) should be used to correct the MRW500 with the MRW500 shield, rather than the K2017a 'universal' single Alter correction that was tested on this gauge/shield combination.

The double-Alter shield transfer function coefficients in Table S2 in the Supplement are recommended for use with the double-Alter shield. However, due to the limited warm season, liquid precipitation measurements included in the SPICE datasets, the use of the Eq. 1 transfer functions presented there is not recommended when T_{air} is > 5 ˚C, as they produce unrealistically high warm-temperature precipitation catch efficiencies. If an Eq. 1 type function applicable to warm-season measurements is needed, we recommend using the pre-SPICE function, which performed quite similar to the WMO-SPICE functions (Fig. 6 a-d). Like the double-Alter WMO-SPICE transfer function, the Eq. 1 Belfort double-Alter shield transfer functions did not include many liquid precipitation events, but in this case, the resultant transfer functions were more realistic at warm temperatures, and can, therefore, be recommended for use in all seasons. The associated transfer function coefficients are provided in the Supplement (Table S3).

For the SDFIR, the pre-SPICE and the SPICE transfer functions performed quite similarly (Fig. 6 i-l). The K2017b SDFIR transfer function was developed using five years of measurement data (1508 30-min precipitation events), whereas only two winter seasons (410 30-min events) were available for the development of the WMO-SPICE transfer function. For warm temperature and high wind speed conditions, the lack of rain events in the WMO-SPICE dataset (76 in Table 2) resulted in unrealistically large SDFIR-shielded catch efficiencies predicted by the Eq. 1 type transfer function. However, the K2017b Eq. 1 type catch efficiencies were more realistic for all temperature/wind speed regimes. For this reason, and because the pre-SPICE transfer function was developed using measurements from the same gauge and shield, at the same site, over a much longer period, we recommend using the Eq. 1 type transfer function from K2017b. Because K2017b did not include Eq. 2 coefficients, the Eq. 2 transfer function coefficients determined from the WMO-SPICE measurements are included in

Table S4 in the Supplement. The Eq. 2 transfer functions, which are only for solid and mixed precipitation, were unaffected by the lack of warm-temperature precipitation in the WMO-SPICE measurements.

3.5.2 Comparison of shield types

Examples of all of the recommended KOC-Eq. 3Eq. 1 type adjustments for solid precipitation are included in Fig. 713a, with the 3-dimensionsal transfer functions plotted against the gauge height wind speed with at T_{air} = -5 ˚C. The T_{air} value of -5 ˚C was selected because it was fairly representative of the solid precipitation events included in this analysis, which had a median T_{air} of -5.2 ˚C. The unshielded and single Alter 'universal' multi-site KOC-Eq.3Eq. 1 transfer functions from K2017aochendorfer et al. (2017) are also included, as these were generally recommended over the gauge-specific unshielded or single-Alter transfer functions developed here. These resultsFigure 7a demonstrates the relative magnitudes of the adjustments for different wind shields, with the more effective shields (SDFIR, Belfort double-Alter) resulting in much higher adjusted catch efficiencies than less effectively shielded gaugess (single-Alter, MRW500 shield) or unshielded gauges.

The uncertainty in each transfer function was also estimated for different wind speeds. Errors in the resultant tTransfer functions errors were calculated from differences between the measured catch efficiency efficiencies and the adjustment (or transfer function) fit to the appropriate catch efficiencies. The resultant RMSE values errors were relatively insensitive to wind speed (Fig. 137b). This is significant, because catch efficiency was presumably less affected by the interaction of snow crystals and wind at low wind speeds. This, so it suggests indicates that variability in snowflake habit, which affects hydrometeor drag and fall velocity, may not be the primary cause of uncertainty in the relationship between catch efficiency and wind speed.the uncertainty in Other these adjustments may have other more important causes of uncertainty, such as random variability in the precipitation gauge measurements and the natural spatial variability in precipitation, may in fact be more important.

In addition, errors in the adjusted catch efficiencies were calculated by applying the appropriate adjustments to the measurements, and calculating RMSE values from the resultant catch efficiencies (Fig. 13c). After adjustment, the catch efficiency should be equal to approximately 1.0, so the RMSE of the adjusted catch efficiency was quantified using the difference between the adjusted catch efficiency and 1.0 (Fig. 7c). The relationship between the magnitudes of the adjustments and the uncertainties in the adjusted catch efficiencies is apparent from comparison of Figs. 137a and 137c. Measurements that required larger adjustments experienced larger errors in the adjusted catch efficiencies. This is due, at least in part, to basic arithmetic; for example, a precipitation measurement associated with a predicted catch efficiency of 50% would be doubled by adjustment, and any errors in the measured catch efficiency would likewise be doubled by the adjustment. At a given wind speed, the errors in the adjusted catch efficiencies (Fig. 137c) are approximately equal to the errors in the catch efficiency (Fig. ~~13~~7b) divided by the adjustment (Fig. ~~13~~7a). Errors in the measured catch efficiency (shown in Fig. ~~13~~7b) were enhanced by the appropriate adjustments (shown in Fig. ~~13~~7a).

This indicates that despite the necessity and utility of transfer functions, effective wind shielding is ~~still~~recommended and beneficial for the measurement of solid precipitation in windy conditions. Many sites and networks use either unshielded or single-Alter shielded precipitation gauges due to the cost of purchasing, transporting, installing, and maintaining larger shields. For example, the DFAR used at the Weissfluhjoch site was built in eight sections and flown in piece-by-piece using a helicopter. This level of expense is necessary at an intercomparison site, but may not be feasible on a larger scale. Many areas where snowfall is monitored are remote and cannot be accessed via road. Sufficient space must also be available to install a large wind shield. In addition, wooden shields require maintenance, and will eventually become weathered and require replacement. Such limitations affect meteorological, climate, and hydrological networks, and must be taken into consideration before selecting a wind shield. For this reason, the performance of the Belfort double-Alter wind shield is notable, as it is much smaller than the SDFIR and DFAR. A stainless steel wind shield like the Belfort double-Alter can also be designed to be relatively low maintenance, as compared to a wooden shield. The efficacy of the Belfort double-Alter shield also indicates that further improvements in wind shield design are possible by altering the geometry and porosity of the wind shield.

4 Conclusions

New transfer functions were developed using precipitation measurements from both host- and manufacturer-provided WMO-SPICE weighing gauges, and were tested alongside existing transfer functions. The resultant errors in corrected precipitation measurements were presented, and recommendations for the correction of different types of weighing gauges were made. These transfer functions were demonstrated to reduce the mean bias of weighing gauge measurements relative to the DFAR, and the remaining uncertainty in the corrected measurements was described using different statistics.

For the unshielded and single-Alter shielded weighing gauges provided by different manufactures for testing in WMO-SPICE, the multi-site transfer function developed ~~by~~ in K~~ochendorfer et al.~~ 2017a typically worked as well as the gauge-specific transfer functions developed in this study. Therefore the more universal unshielded and single-Alter multi-site transfer functions from K2017a~~ochendorfer et al.~~ are recommended for adjusting measurements from all the ~~the~~ unshielded and single-Alter-shielded weighing gauges tested.

The low-porosity double-Alter shield ~~developed~~ manufactured by Belfort performed well relative to the DFAR, with an average uncorrected bias of only – 0.04 mm, or – 5.4%. The Belfort double-Alter ~~Comparison of results with~~shielded gauges performed better than weighing gauges in traditional single- and double-Alter shields~~ indicated better performance for the~~

~~Belfort double-Alter~~, suggesting that it is a viable, high-efficacy option for networks or sites that do not have the resources to build, site, and maintain a large wooden shield like the SDFIR or DFIR.~~ The performance of the Belfort double-Alter shield also indicates that future improvements in shield design may yet be possible, especially considering the significant resources required to site, install, and maintain large wooden shields like the SDFIR or DFIR.~~

Precipitation measurements from weighing gauges in higher-efficacy shields, such as the SDFIR and the Belfort double-Alter, showed not only much smaller uncorrected biases relative to the corresponding reference configurations, but also smaller ~~corrected~~ adjusted RMSE and higher ~~corrected~~ adjusted $PE_{0.1\ mm}$. M~~Further, m~~easurements from these gauge/shield configurations required less ~~correction~~ adjustment than the unshielded gauges tested, and the resultant errors estimated by comparing the ~~corrected~~ adjusted measurements to the DFAR measurements were also much smaller. The errors that remained after ~~correcting~~ adjusting the unshielded and single-Alter shielded measurements were much larger than the errors experienced by the more effectively shielded gauges. Upon closer inspection and bin-averaging by wind speed, the magnitude of ~~it appears that the~~ uncorrectable errors in the adjusted measurements increased with the size of the required adjustment~~ that less effectively shielded measurements are subject to are enhanced by the adjustments required to remove the undercatch;. A~~ at higher wind speeds, where ~~such~~ less-effectively shielded measurements required doubling or even tripling, the uncertainty in the measurements was also doubled or tripled, accordingly. This suggests that there is a limit to the amount of uncertainty that can be removed by such adjustments, and the transfer functions presented here may already be approaching this limit. These results also suggests that although adjusted unshielded and single-Alter shielded gauge measurements can be used to ~~effectively~~ measure the total amount of precipitation without a large bias, the ~~only~~ best way to ~~significantly~~ reduce the uncertainty of the measurement is to shield the gauge~~it~~ more effectively using a shield such as the DFIR, SDFIR, or Belfort double-Alter.

Acknowledgements

The authors thank Hagop Mouradian from Environment and Climate Change Canada for contributing the mapped site locations (Fig. 1). We thank the manufacturers that provided many of the sensors used to produce these results. We also thank the World Meteorological Organization for supporting this intercomparison.

Disclaimers

Many of the results presented in this work were obtained as part of ~~the Solid Precipitation Intercomparison Experiment (SPICE)~~SPICE, conducted on behalf of the World Meteorological Organization (WMO) Commission for Instruments and Methods of Observation (CIMO). The analysis and views described herein are those of the authors at this time, and do not necessarily represent the official outcome of WMO-SPICE. Mention of commercial companies or products is solely for the purposes of information and assessment within the scope of the present work, and does not constitute a commercial endorsement of any instrument or instrument manufacturer by the authors or the WMO.

References

Baghapour, B., Wei, C., and Sullivan, P. E.: Numerical simulation of wind-induced turbulence over precipitation gauges, Atmospheric Research, 189, 82-98, http://dx.doi.org/10.1016/j.atmosres.2017.01.016, 2017.

Buisán, S. T., Earle, M. E., Collado, J. L., Kochendorfer, J., Alastrué, J., Wolff, M., Smith, C. D., and López-Moreno, J. I.: Assessment of snowfall accumulation underestimation by tipping bucket gauges in the Spanish operational network, Atmos. Meas. Tech., 10, 1079-1091, 10.5194/amt-10-1079-2017, 2017.

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.

Goodison, B., Louie, P., and Yang, D.: The WMO solid precipitation measurement intercomparison, World Meteorological Organization-Publications-WMO TD, 65-70, 1998.

Goodison, B. E.: Accuracy of Canadian snow gauge measurements, Journal of Applied Meteorology, 17, 1542-1548, 10.1175/1520-0450(1978)017<1542:aocsgm>2.0.co;2, 1978.

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.

Kochendorfer, J., Nitu, R., Wolff, M., Mekis, E., Rasmussen, R., Baker, B., Earle, M. E., Reverdin, A., Wong, K., Smith, C. D., Yang, D., Roulet, Y. A., Buisan, S., Laine, T., Lee, G., Aceituno, J. L. C., Alastrué, J., Isaksen, K., Meyers, T., Brækkan, R., Landolt, S., Jachcik, A., and Poikonen, A.: Analysis of single-Alter-shielded and unshielded measurements of mixed and solid precipitation from WMO-SPICE, Hydrol. Earth Syst. Sci., 21, 3525-3542, 10.5194/hess-21-3525-2017, 2017a.

Kochendorfer, J., Rasmussen, R., Wolff, M., Baker, B., Hall, M. E., Meyers, T., Landolt, S., Jachcik, A., Isaksen, K., Brækkan, R., and Leeper, R.: The quantification and correction of wind-induced precipitation measurement errors, Hydrol. Earth Syst. Sci., 21, 1973-1989, 10.5194/hess-21-1973-2017, 2017b.

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.

Nitu, R.: Proposed configuration of intercomparison sites and of the field references, Second session of the international organization committee for the WMO solid precipitation intercomparison experiment, World Meteorological Organization, Boulder, CO, 2012.

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.

Reverdin, A.: Description of the Quality Control and Event Selection Procedures used within the WMO-SPICE Project, Proceedings of the WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation, Madrid, Spain, 2016.

Roulet, Y.-A., Reverdin, A., Buisan, S., and Nitu, R.: Non-catchment type instruments for snowfall measurement and reporting of precipitation type: General considerations and issues encountered during the WMO CIMO SPICE experiment, and derived recommendations, Proceedings of the WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation, Madrid, Spain, 2016.

Ryu, S., Lee, G., Nitu, R., Smith, C., Lim, E., and Kin, H.-L.: Automatic Double Fence Reference (DFAR) for measuring solid precipitation: Gauge Based Characterization, Second session of the international organization committee for the WMO solid precipitation intercomparison experiment, World Meteorological Organization, Boulder, CO, 2012.

Ryu, S., Lee, G., Nitu, R., Smith, C., Lim, E., and Kim, H. L.: Quantitative comparison of two types of working field reference systems for the measurement of precipitation amount (R2), Proceedings of the WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation Madrid, Spain, 2016,

Sevruk, B., Hertig, J.-A., and Spiess, R.: The effect of a precipitation gauge orifice rim on the wind field deformation as investigated in a wind tunnel, Atmospheric Environment. Part A. General Topics, 25, 1173-1179, 1991.

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.

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., Ishida, S., and Benson, C. S.: Adjustment of daily precipitation data at 10 climate stations in Alaska: Application of World Meteorological Organization intercomparison results, Water Resources Research, 34, 241-256, 10.1029/97wr02681, 1998a.

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, 1998b.

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.

Gauge	Measurement	Heater	Output	Capacity	Recorded rate
MPS Systém TRwS 405	Load cell	15 W	Digital	750 mm	1 min (Ma, Hauk)
Meteoservis MRW500	Load cell	350 W	Digital	900 mm	1 min (Ma, BrLa)
Sutron TPG	Load cell	200 W	Digital	914 mm	6 s (Ma)
1500 mm Geonor	Vibrating wire	60 W (Ma, Weis), N/A* (BrLa, CaCr)	Analog	1500 mm	6 s (Ma, Weis), 1 min (BrLa, CaCr)
600 mm Geonor	Vibrating wire	60 W	Analog	600 mm	6 s (Ma)
OTT Pluvio[2]	Load cell	50 W	Digital	1500 mm	6 s (CARE)

Table 1. Gauge type, measurement type, heater power, output type, gauge capacity, and the recorded rate of the different types of gauges included in this evaluation. The recorded rate was determined by the logging system at a given site, and was not intrinsic to the gauges. Testbeds included in the study include the Centre for Atmospheric Research Experiments (Canada, abbr. CARE), Caribou Creek (Canada, abbr. CaCr), Bratt's Lake (Canada, abbr. BrLa), Marshall (United States, Ma), Haukeliseter (Norway, abbr. Hauk), and Weissfluhjoch (Switzerland, abbr. Weis). *The 1500 mm Geonor gauges at Bratt's Lake and Caribou Creek were tested without heaters.

Shield	Gauge	Testbed	Mean U_{gh}	Max U_{gh}	N_{liquid}	N_{mixd}	N_{solid}
Unshielded	Sutron TPG	Ma	2.8 m s^{-1}	10.2 m s^{-1}	77	166	208
Unshielded	MRW500	Ma, BrLa	3.2 m s^{-1}	10.2 m s^{-1}	121	214	230
Unshielded	TRwS 405	Ma, Hauk	4.0 m s^{-1}	17.0 m s^{-1}	128	250	250
MRW500 shield~~Unshielded~~	MRW500~~1500 mm Geonor~~	Ma, BrLa~~Ma, BrLa, Weis, CaCr~~	3.2 m s^{-1}~~3.3 m s^{-1}~~	10.2 m s^{-1}~~11.6 m s^{-1}~~	121~~172~~	214~~374~~	230~~996~~
Single Alter~~MRW500 shield~~	Sutron TPG~~MRW500~~	Ma~~Ma, BrLa~~	2.8 m s^{-1}~~3.2 m s^{-1}~~	10.2 m s^{-1}~~10.2 m s^{-1}~~	78~~121~~	172~~214~~	201~~230~~
Single Alter ~~Single Alter~~	1500 mm Geonor~~Sutron~~	Ma, BrLa, Weis, CaCr~~Ma~~	3.3 m s^{-1}~~2.8 m s^{-1}~~	11.6 m s^{-1}~~10.2 m s^{-1}~~	172~~78~~	374~~172~~	996~~201~~
Double Alter	Geonor/Pluvio[2]	Ma, CARE	3.0 m s^{-1}	10.2 m s^{-1}	147	173	403
Belfort double Alter	Geonor/Pluvio[2]	Ma, CARE	3.0 m s^{-1}	10.2 m s^{-1}	206	244	496
SDFIR	Geonor	Ma	2.9 m s^{-1}	10.2 m s^{-1}	76	173	161

Table 2~~1~~. Shield type, gauge type, testbed, mean and maximum (Max) of the measured gauge height wind speed (U_{gh}), and the number of liquid (N_{liquid}), mixed (N_{mixd}), and solid (N_{solid}) 30-min precipitation measurements included in this study. The measurements for all gauges were recorded during the two winter seasons (Oct 1 – April 30) of 2013-2015. Wind speed statistics only describe periods of precipitation.

Testbeds included in the study include the Centre for Atmospheric Research Experiments (Canada, abbr. CARE), Caribou Creek (Canada, abbr. CaCr), Bratt's Lake (Canada, abbr. BrLa), Marshall (United States, Ma), Haukeliseter (Norway, abbr. Hauk), and Weissfluhjoch (Switzerland, abbr. Weis).

Shield	Gauge	Recommendation
Unshielded	Sutron TPG	Tables 2 – 3 in K2017a
Unshielded	MRW500	Tables 2 – 3 in K2017a
Unshielded	TRwS 405	Tables 2 - 3 in K2017a
Unshielded	1500 mm Geonor	Tables 2 - 3 in K2017a
MRW500 shield	MRW500	Table S1 in Supplement
Single Alter	Sutron TPG	Tables 2 - 3 in K2017a
Double Alter	Geonor/Pluvio^2	Table S2 in Supplement
Belfort double Alter	Geonor/Pluvio^2	Table S3 in Supplement
SDFIR	Geonor	Table 2 in K2017b; Table S4 in Supplement

Table 3. Recommended transfer functions for different weighing precipitation gauges and shields.

Table 2. Shielded MRW500 transfer function coefficients for gauge-height (GH) and 10 m wind speeds. Coefficients for Eq. 1 and KOC-Eq. 3 are provided, along with the maximum wind speed threshold (U_{thresh}), above which the transfer function should be applied by forcing the wind speed down to U_{thresh}.

Wind Speed	**KOC-Eq. 4, *f(U, mixd)***			**KOC-Eq. 4, *f(U, solid)***			**KOC-Eq. 3 *f(U, T_{air})***			**U_{thresh}**
	a	***b***	***c***	***a***	***b***	***c***	***a***	***b***	***c***	**(m s^{-1})**
$f(U_{GH})$	0.982	0.0201	0.00	0.829	0.213	0.32	0.0404	0.4247	0.00	6.1
$f(U_{10m})$	0.98	0.0137	0.00	0.886	0.143	0.25	0.0312	0.427	0.00	8

Table 3. Double-Alter transfer function coefficients for gauge-height (GH) and 10 m wind speeds. Coefficients for KOC-Eq. 4 and KOC-Eq. 3 are provided, along with the maximum wind speed threshold (U_{thresh}), above which the transfer function should be applied by forcing the wind speed down to U_{thresh}.

Wind Speed	KOC-Eq. 4, *f(U, mixd)*			KOC-Eq. 4, *f(U, solid)*			KOC-Eq. 3 $f(U, T_{air})$			U_{thresh}
	a	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*	($m\ s^{-1}$)
$f(U_{GH})$	0.260	1.512	0.95	1.068	0.049	0.00	0.0146	0.27	0.05	6.1
$f(U_{10m})$	0.254	1.052	0.95	1.075	0.039	0.00	0.0110	0.29	0.08	8

Table 4. Belfort double-Alter transfer function coefficients for gauge-height (GH) and 10 m wind speeds. Coefficients for KOC-Eq. 4 and KOC-Eq. 3 are provided, along with the maximum wind speed threshold (U_{thresh}), above which the transfer function should be applied by forcing the wind speed down to U_{thresh}.

Wind Speed	KOC-Eq. 4, *f(U, mixd)*			KOC-Eq. 4, *f(U, solid)*			U_{thresh}
	a	*b*	*c*	*a*	*b*	*c*	($m\ s^{-1}$)
$f(U_{GH})$	0.492	0.00	0.492	0.187	0.410	0.875	6.1
$f(U_{10m})$	0.492	0.00	0.492	0.186	0.230	0.875	8

Table 5. Small DFIR (SDFIR) transfer function coefficients for gauge-height (GH) and 10 m wind speeds. Coefficients for KOC-Eq. 4 are provided, along with the maximum wind speed threshold (U_{thresh}), above which the transfer function should be applied by forcing the wind speed down to U_{thresh}.

Figure 1. Map of WMO-SPICE test sites with weighing gauges considered in this study.

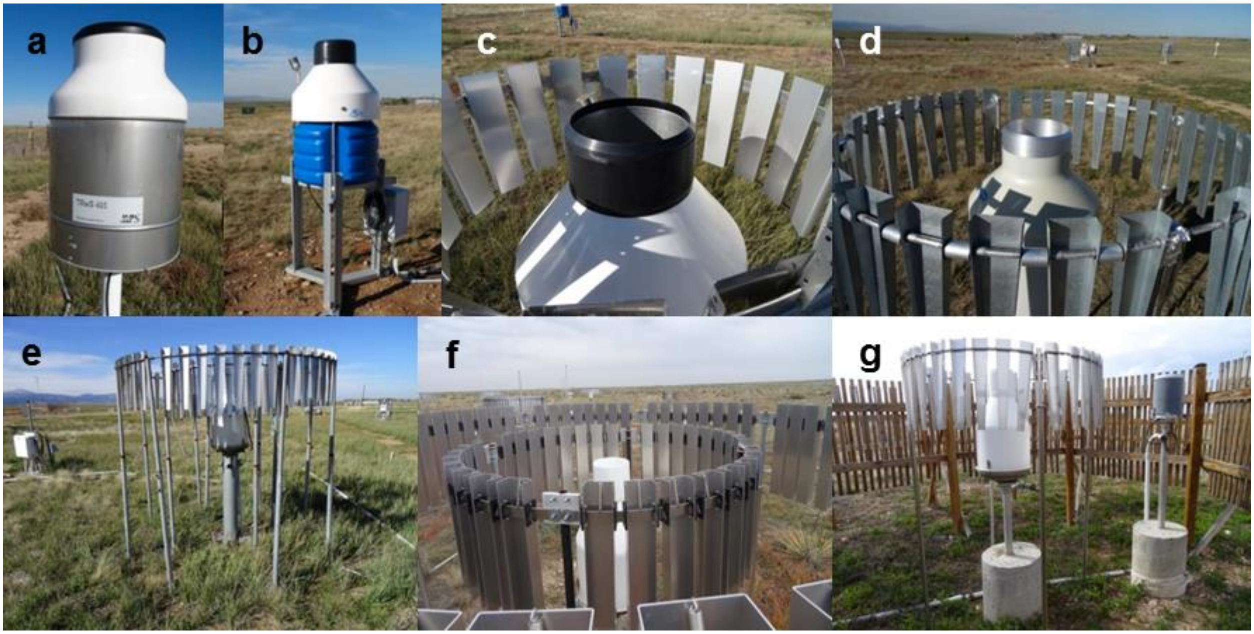


Figure 2. Photos of the (a) TRwS 405, (b) unshielded MRW500, (c) shielded MRW500, (d) single-Alter shielded Sutron, (e) double-Alter shielded Geonor, (f) Belfort double-Alter shielded Geonor, and (g) SDFIR-shielded Geonor at the Marshall, CO, US testbed.

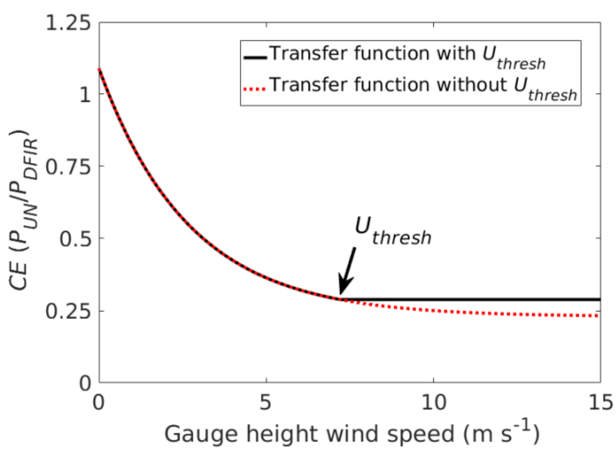


Figure 3. Example application of the maximum wind speed threshold (U_{thresh}), using an ~~KOC Eq. 4~~Eq. 2 type transfer function describing unshielded (UN) solid precipitation catch efficiency (*CE*) from K2017a~~Kochendorfer et al.~~. At wind speeds exceeding U_{thresh} (7.2 m s^{-1} in this case) the catch efficiency is fixed at the value determined at U_{thresh}.

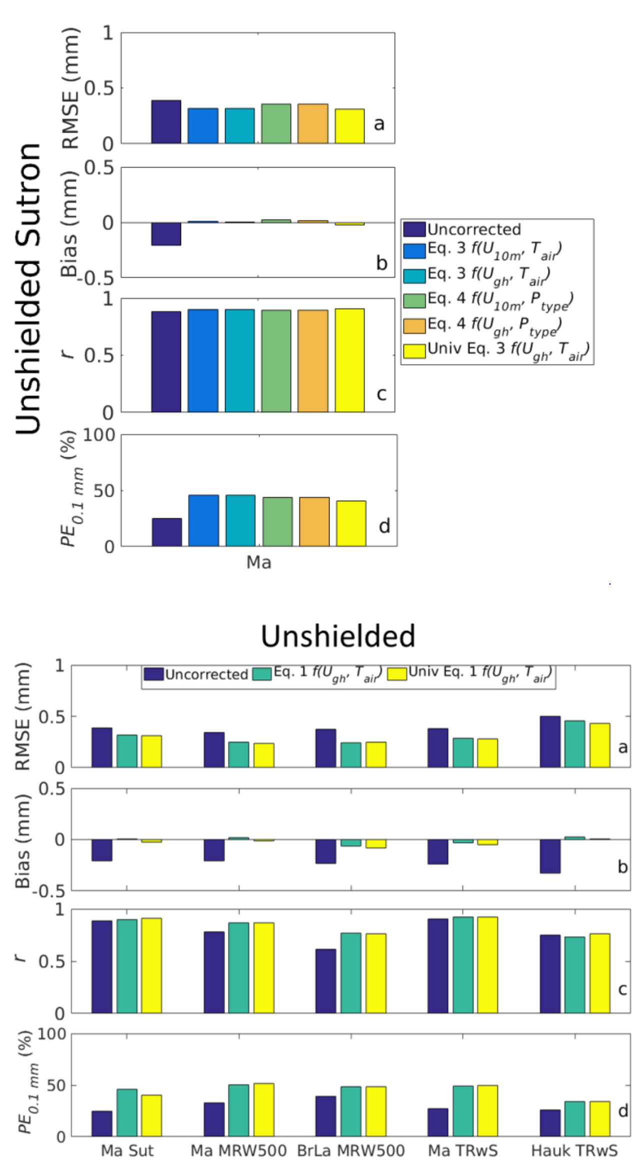


Figure 4. (a) Root mean square error (RMSE), (b) bias, (c) correlation coefficient (*r*), and (d) the percentage of events with errors less than 0.1 mm ($PE_{0.1\ mm}$) calculated from the difference between the 30 min precipitation measurements from the corrected, unshielded Sutron gauges and the DFARDFIR-shielded reference automated precipitation gauge at the Marshall testbed (Ma)corresponding test sites. The Sutron from Marshall (Ma Sut), the MRW500 from Marshall (Ma MRW500) and Bratt's Lake (BrLa MRW500), and the TRwS 405 from Marshall (Ma TRWS) and Haukeliseter (Hauk TRWS) are included. The uUncorrected measurements are also shown (Uncorrected, dark blue). Measurements corrected adjusted using several different types of an example transfer functions (Eq. 1) are also shown, both as defined by KOC Eq. 3 (Eq. 3) and KOC Eq. 4 (Eq. 4), fit to precipitation, air temperature (T_{air}) and as a function of the gauge height wind speed (U_{gh}) and the 10 m height wind speed (U_{10m}).are shown in green. Statistics describing these measurements adjusted using The results of the 'universal' unshielded transfer function (Univ. Eq. 13, yellow), which were based on independent Eq. 1 coefficients derived in KKochendorfer et al. 2017a, are also shown.

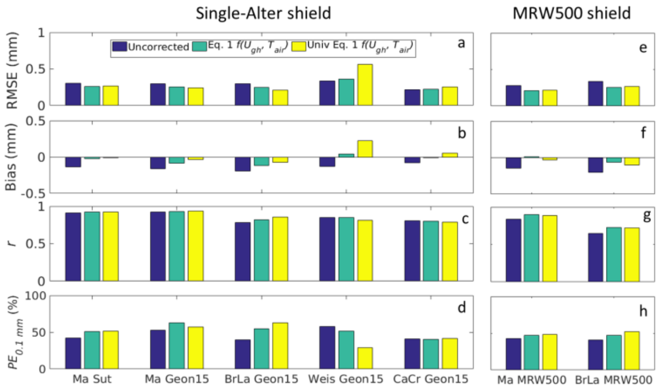


Figure 5. Error statistics calculated for the single-Alter shielded Marshall Sutron gauge (Ma Sut) and single-Alter shielded 1500 mm Geonor gauges (Geon15) at the Marshall (Ma), Bratt's Lake (BrLa), Weissfluhjoch (Weis), and Caribou Creek (CaCr) testbeds (a – d). Error statistics from the shielded MRW500 gauges from Marshall (Ma MRW500) and Bratt's Lake (BrLa MRW500) are also shown (e – h). The different statistics and correction types are described in the Fig. 4 caption. The results of the 'universal' single-Altear transfer function (Univ. Eq. 1, yellow), which were based on independent Eq. 1 coefficients derived in K2017a, are shown for all of the single-Alter shielded gauges and also for the gauges within the MRW500 shield.

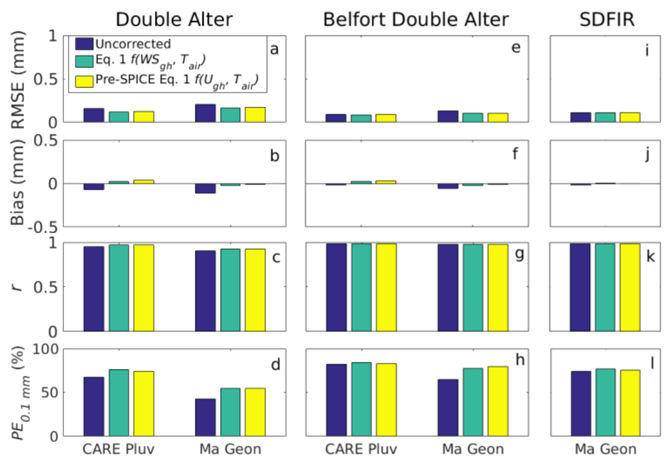
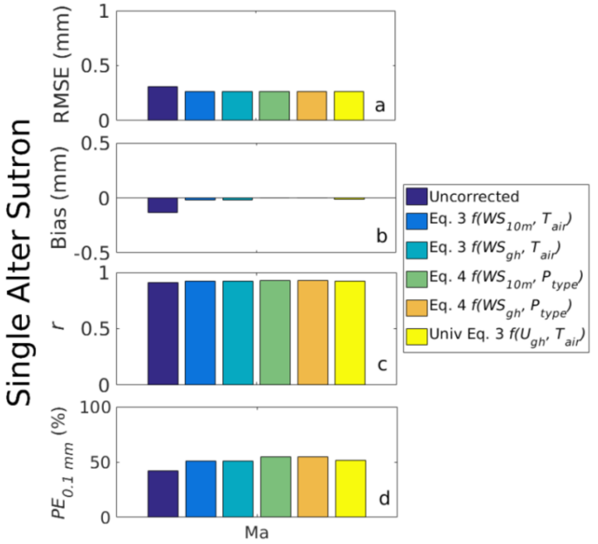
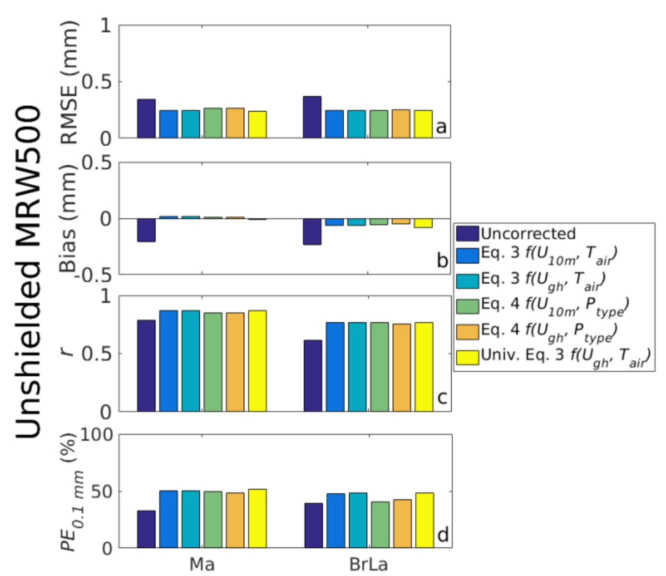


Figure 6. Error statistics calculated from the double-Alter shielded Pluvio^2 gauge at CARE (Care Pluv, a - d), the double-Alter shielded Geonor at Marshall (Ma Geon, a - d), the Belfort double-Alter shielded Pluvio^2 at CARE (CARE Pluv, e – f), the Belfort double-Alter shielded Geonor at Marshall (Ma Geon, e – h), and the SDFIR shielded Geonor at Marshall (Ma Geon, i – l). The different adjustment types are described in the Fig. 4 caption. Statistics describing these measurements adjusted using the

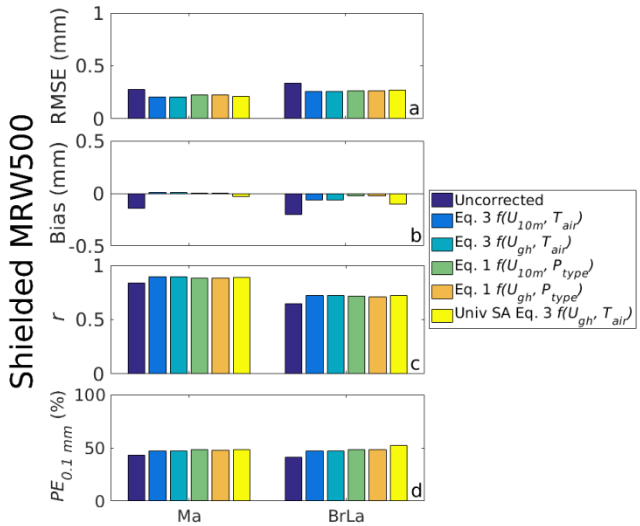
appropriate pre-SPICE transfer functions from K2017b are also shown (Pre-SPICE. Eq. 1, yellow).



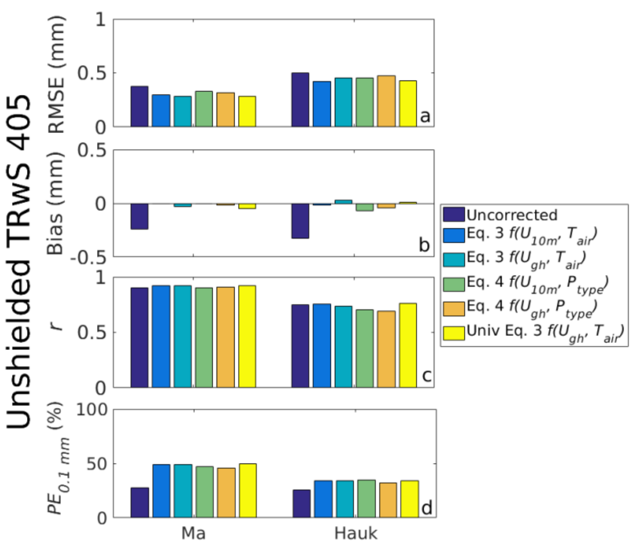
~~Figure 5. Error statistics calculated from the difference between the 30 min precipitation measurements from the single-Alter shielded (SA) Sutron gauge and the DFIR-shielded reference automated precipitation gauge at the Marshall testbed (Ma). The different statistics and correction types are described in the Fig. 4 caption.~~



~~Figure 6. Error statistics calculated from the difference between the 30 min precipitation measurements from the unshielded (UN) MRW500 gauge and the DFIR-shielded reference automated precipitation gauge at the Marshall (Ma) and Bratt's Lake (BrLa) testbeds. The different statistics and correction types are described in the Fig. 4 caption.~~



~~**Figure 7. Error statistics calculated from the difference between the 30 min precipitation measurements from the shielded MRW500 gauge and the DFIR-shielded reference automated precipitation gauge at the Marshall testbed (Ma) and Bratt's Lake (BrLa) testbeds. The different statistics and correction types differ from those described in the Fig. 4 caption in the respect that an Eq. 1 type adjustment was used for this gauge configuration in place of KOC Eq. 4.**~~



~~**Figure 8. Error statistics calculated from the difference between the 30 min precipitation from the unshielded TRwS 405 gauge and the DFIR-shielded reference automated precipitation gauge at the Marshall (Ma) and Haukeliseter (Hauk) testbeds. The different statistics and correction types are described in the Fig. 4 caption.**~~

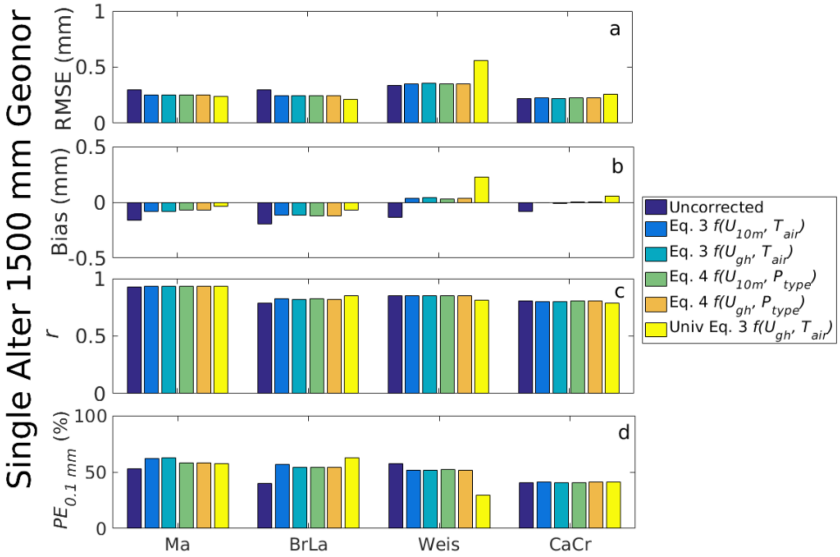


Figure 9. Error statistics calculated from the difference between the 30 min precipitation measurements from the single-Alter shielded 1500 mm Geonor and the DFIR-shielded reference automated precipitation gauge at the Marshall (Ma), Bratt's Lake (BrLa), Weissfluhjoch (Weis), and Haukeliseter (Hauk) testbeds. The different statistics and correction types are described in the Fig. 4 caption.

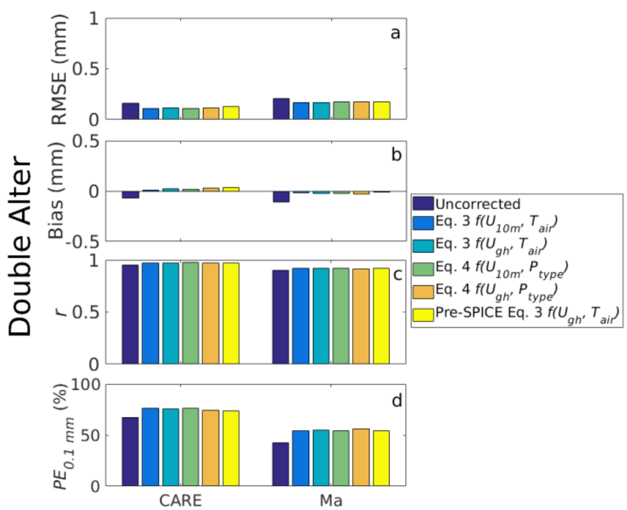
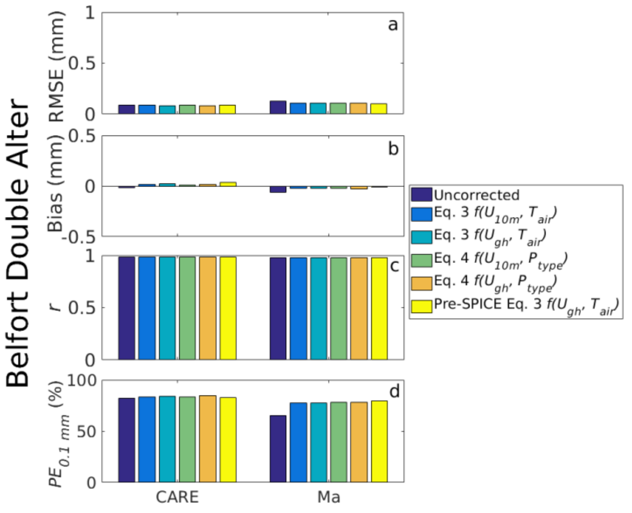
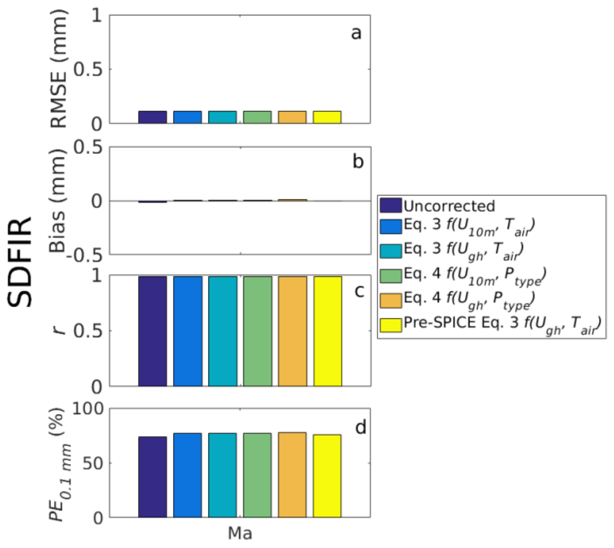


Figure 10. Error statistics calculated from the difference between the 30 min precipitation measurements from the double-Alter shielded gauges and the DFIR-shielded reference automated precipitation gauge at the CARE and Marshall (Ma) testbeds. The different correction types are described in the Fig. 4 caption.



~~**Figure 11. Error statistics calculated from the difference between the 30 min precipitation measurements from the Belfort double-Alter shielded gauges and the DFIR-shielded reference automated precipitation gauge at the CARE and Marshall (Ma) testbeds. The different correction types are described in the Fig. 4 caption.**~~



~~**Figure 12. Error statistics calculated from the difference between the 30 min precipitation measurements from the SDFIR-shielded gauge and the DFIR-shielded reference automated precipitation gauge at the Marshall (Ma) testbed. The different correction types are described in the Fig. 4 caption.**~~

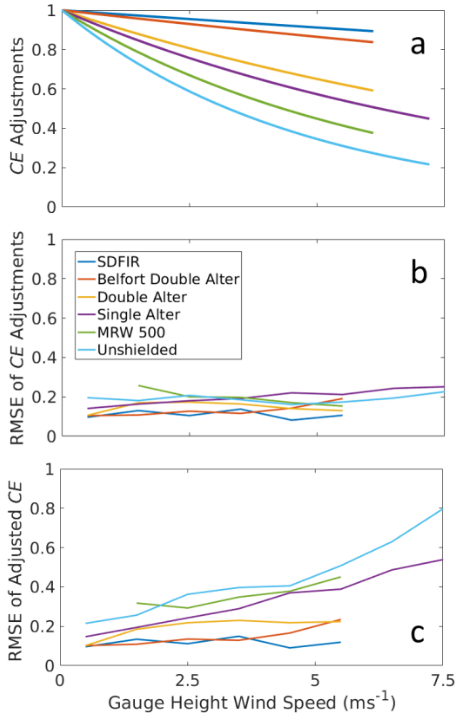


Figure 7~~13~~. (a) Comparison of recommended ~~KOC Eq. 3~~Eq. 1 catch efficiency (*CE*) adjustments at T_{air} = -5 °C. (b) RMSE of the *CE* adjustments, estimated by bin-averaging for every 1 m s^{-1} of wind speed. (c) RMSE of the adjusted *CE* bin-averaged by wind speed, estimated by adjusting the measurements and calculating RMSE values from the resultant catch efficiencies. The SDFIR, Belfort double- Alter, standard double-Alter and MRW500 transfer function coefficients were determined from the results presented here. The single-Alter shielded and unshielded results are from K~~ochendorfer et al.~~ 2017a. Bin-averaged RMSE values are only shown (b and c) when more than 30 values were available within a given 1 m s^{-1} wind speed range.