Author's Response to Anonymous Referee #1

We would like to thank Referee #1 for taking the time to review and offer feedback on this manuscript. Our responses to the comments and actions taken are listed below in red.

RC: P2, L15: In this paragraph, a short comment of how these stations are maintained after snowfall, or when they might get ice around, would be beneficial.

AC: This paragraph is more of an introduction to the SPICE sites and intercomparison period so a discussion of site maintenance may not be appropriate here. However, to answer the question, site maintenance was the responsibility of the site host and many sites kept service logs that were made available to improve data quality control. These logs usually note gauge servicing and instrument malfunctions.

Action: The following sentence has been added to the Methods sections to clarify how site maintenance issues were identified and handled in data quality control: "Where available, service logs were provided by site hosts to assist in data quality control and the identification of outliers due to servicing (e.g. rapid drops or increases in precipitation gauge bucket weights) or maintenance (e.g. instrument malfunctions or other human interventions that may impact the data)."

RC: P2, L28: What is a Precipitation detector?

AC: For SPICE, this was typically an optical disdrometer that is capable of detecting the occurrence of even light precipitation events at relative high temporal resolution (1-min in this case).

Action: Added the statement to describe "precipitation detector" as "typically an optical disdrometer capable of identifying the occurrence of even light precipitation.

RC: P3, L6: A reference after "demonstrating that the unshielded catch from both SUT types were very similar." would help in this sentence.

AC: This statement appears in the same sentence that references Kochendorfer et al. (2017b). The same reference applies to the latter statement in the sentence.

Action: To clarify, the sentence was changed to read "The transfer functions presented in K2017b were developed…, after the authors demonstrated that the unshielded catch…"

RC: P4, L9-11: I think that a more detailed explanation of this sentence would help to complete the idea. For example, answering what makes these evaluations "marginally independent"?

AC: We will modify this paragraph to clarify.

Action: The paragraph now states: "As discussed above, the data from all eight sites in Table 1, including data from multiple gauges of the same configuration at each of the sites, were combined to fit the transfer function models. This model, developed by pooling the data from multiple sites, was then applied to each individual gauge at each site. By applying a model developed from data collected at

multiple sites to individual gauges at each individual site, the authors maintained some independence between the model development and the evaluation."

RC: P4, L17: I would include this sentence as the main result in the abstract.

Action: The relevant sentence in the abstract now reads: "Due to the short intercomparison period, the dataset was not sufficiently large to develop and evaluate transfer functions using independent precipitation measurements, although on average the adjustments were effective at reducing the bias in unshielded gauges from -33.4% to 1.1%."

RC: P4, L29: What do you mean with "more natural circumstances"?

AC: We meant in situations more typical of those experienced by a user collecting data in the field. We will revise this sentence to reflect this.

Action: Sentence was changed to say: "The methodology used during SPICE for developing and evaluating the transfer functions used only a subset of the observed data (the SEDS), and although this was a robust methodology for developing transfer functions, it did not provide a comprehensive evaluation of the adjustments under circumstances more typical of users collecting precipitation data in the field where the data is less filtered to remove smaller amounts."

RC: P4, L36: This has to be consistent with P4, L9-11.

AC: We think that the revision on P4, L9-11 now sufficiently explains the issue such that no further action is required here.

Action: none taken

RC: P5, L3: I think the word "continued" should be in present tense.

AC: agreed

Action: "continued" change to "continues"

RC: P5, L7: I suggest completing the end of the sentence as: "...it will also isolate the performance to adjust snow measurements"

AC: agreed

Action: sentence modified as suggested

RC: P9, L15: I have a question here, why not using a standardized RMSE value, as it highly depends on the precipitation rates?

AC: There is advantages and disadvantaged to either. Standardizing the RMSE by using precipitation rate would allow better intercomparison between sites but information regarding the absolute magnitude of the error at each individual site would be lost. The most important interpretation here is the change (or

lack of) in RMSE with adjustment. Therefore, we think that using RMSE and cautioning the interpretation between sites, combined with the use of the other metrics, is the appropriate action.

Action: none taken

RC: P10, L1: Could you please add a number or a percentage to quantify the adjective“insubstantial”?

Action: We quantified this as < 0.005 mm for the single-Alter and as 0.05 mm for the unshielded gauges.

RC: P9, L11-15; P10, L5 L7-8; P10, L28-32: these sentences describe the statistical parameters used to analyze the data. I suggest moving them to the methodology section.

AC: Agreed

Action: The description and justification of the metrics have been moved to section 2.2 Performance metrics under Methods

RC: P12, L10-11, The number of events that are not captured by the single-Alter shielded in XBK is over 50%, do you have any comments to propose a better monitoring system in this location? Or how we could possibly integrate these cases to the adjusted precipitation?

AC: Fortunately, many of those event are small events which reduces the impact on total precipitation. However, they are still significant given the percentage of the reference precipitation amount not measured by the non-reference gauges. Currently, the most effective way to reduce this error is to increase the catch efficiency of the non-reference gauge by adding shielding. This is captured in the sentence: “Since more shielding (e.g. double-Alter) generally means a higher catch (Watson et al., 2008; Smith, 2009; Rasmussen et al., 2012; Kochendorfer et al., 2017b), more shielding would also reduce the number of unmeasured events.” The use of disdrometers could also increase detection of small events in cold and windy conditons.

Action: Added the following sentence: “Perhaps another option to increase detection of small events in cold and windy locations would be the use of optical disdrometers paired with the conventional accumulating gauges.”

RC: Table 1; Please, specify the meaning of “gh” in the table caption, or as a footnote.

Action: done

RC: Figures 3 and 4, I suggest increasing the panel size and/or font size.

AC: That’s a good suggestion and we will make sure those are larger for final production. For review, we wanted to try to get them all on the same page.