Response to Reviewer 1

Thank you for your helpful comments and suggestions. Please note that paper has been updated using data obtained through 16 May, 2018 to give a full 2-year time series. This has not caused any major changes in our conclusions.

Reponses are provided in bold font to each comment in italics.

This manuscript presents the results of a nearly 2-year long test at Summit, Greenland of a commercial device (Snowfox) developed to measure snow accumulation in water equivalents (SWE). The Snowfox sensor measures neutrons produced in the earth's atmosphere by cosmic rays, these neutrons are attenuated by water accumulating above the sensor in the form of snow. While this technique has been applied fairly widely in alpine snowpacks, the authors assert that the test described in the manuscript is the first application in the firn on a polar ice sheet.

We note that "fairly widely" is debatable, as the only existing network is operated by the French utility entity in a relatively small portion of the Alps, with little documentation (only the proceeding and white paper cited in the manuscript, neither of which may have been peer reviewed) and no scientific application we're aware of.

The results are very encouraging (perhaps not surprising), but oversold in my opinion. Specifically, I find the claim that this test confirms better than 0.5% accuracy for SWE < 20 cm and better than 0.7% accuracy for SWE up to 140 cm implausible. I note that previous studies with similar devices (cited in this manuscript) conclude that accuracy in the 3 to 5% range could be achieved, and that data with this quality were useful.

We now clarify that these values are estimates of the sensor's precision, rather than accuracy, as based on the daily standard deviation in hourly measurements as function of depth given in Fig. 4. The accuracy can only be determined through validation, although, as explained in the text, little or no bias is found, indicating that the precision is indeed representative of the error.

Note that previous studies were in alpine settings measuring seasonal snow, where soil moisture and higher spatial variability made their measurements, and their validation, much more uncertain. This is explained in the introduction.

My skepticism is based on a combination of 4 factors: 1) the corrections to account for temporal variations in neutron flux at the snow suface and variations in the column of water vapor in the atmosphere above the buried sensor do not seem likely to be accurate to better than 1% (perhaps considerably less precise), 2) the noise in the estimated accumulation (Figures 4 and 5) appears to be at least +/- 2%, 3) comparison to 2 independent manual techniques to measure accumulation do not show consistent agreement at < 1% across the 10 month comparison period, and 4) over the full month long trial the maximum SWE reached just 42 cm, suggesting

strongly that the claimed accuracy at 140 cm is entirely theoretical, and hinting that the same might be true for SWE <20cm.

Addressing each numbered point:
1) Reasons for neglecting atmospheric pressure are provided where this point is repeated in more detail below.

2) The precision curve is derived from Figure 4 (it's the solid curve). Figure 5. plots the daily difference in accumulation, with 4B giving the percentage change. The mean percentage change is 1.2%, which is close to the 1% error one gets when differencing of two measurements with 0.7% errors.

3) The mean biases between accumulation measured by the cosmic rays sensor and the validation datasets for the ~8 day periods between manual surveys are < 1% of the water thickness. The standard deviations of the differences between cosmic ray and validation dataset estimates are close to or within the expected precision of the cosmic ray sensor and the validation data. This is now explained more clearly in the text.

4) We now make it clear in the abstract and text that these precision estimates are based on the precision curve fit to the daily standard deviations in hourly measurements in up to 56 cm of water shown as the black curve in Figure 4. The loss in precision with thickness is due to the predictable decline in count rate and signal to noise ratio.

Regarding the corrections to raw data, more detail is needed to explain how the neutron monitor at Thule is used to provide No (neutron flux at the surface). It is likely that the Snowfox and the Thule monitors are not measuring neutrons with identical efficiency across the energy range, and I would be surprised if hourly changes in neutron flux were perfectly in sync, given 6 degrees in latitude and nearly 3 km in altitude separation between the 2 sensors.

It is true that our method of correction cannot account for anisotropy of the primary cosmic ray flux. This could lead to some short term error, for example, in accounting for individual Forbush decrease events. However, our method does ensure that on average (i.e. on timescales of multiple days) our solar correction should be unbiased. We employ a factor to account for latitude variation (insignificant here, since both the Thule station and our site are well above the "knee" in neutron intensity, see Hawdon et al. 2014, citation added to paper) and altitude difference (which is significant). Essentially, we estimate that the our site, because of it's higher altitude, is on average 1.19 times as sensitive to solar variations than Thule station. This is based on a regression to monthly neutron monitor data from the global neutron monitor data set, and again, will not work well for individual events. Fortunately, those interested in accumulation and mass balance will be interested in weekly to monthly changes, and short lived solar events will be averaged out.

To describe this is more detail, we have added the following to section 3:

"The unitless scaling parameter **b** accounts for differences in the magnitudes of solar-induced variations between the reference and local sensors due to differences in latitude and elevation, only the latter of which will be significant in this case (Hawdon et al. 2014). We use a value of **b** =1.19 based on regressions to the global neutron monitor dataset. While this single value produces unbiased corrections on weekly or longer timescales, we expect some errors associated with short term variation, such as individual Forbush decrease events."

(Are the Snowfox devices inexpensive enough to use one mounted above the surface to provide No (near a single buried Snowfox, or in the middle of a regional array of buried sensors in a future study)).

A local reference station would certainly reduce uncertainty in corrections, especially for measuring short-term variability as described above. However, it's not obvious to what extent uncertainty would be reduced over longer-time scales, and if this reduction would outweigh the risk of relying on a reference sensor on the ice sheet as opposed to a global-network monitor like THUL.
We have added the following sentence to section 5:
"Finally, additional accuracy on short time scales (< 1 week) may be obtained by deploying a local reference sensor above the surface, removing uncertainties in the incoming cosmic ray corrections due to differences in elevation and latitude."

Atmospheric water vapor is not a simple linear function of atmospheric pressure, varying depending on synoptic conditions in addition to pressure. Assuming solely pressure dependence has to introduce more than 0.5 or 0.7% uncertainty in the derived SWE. Note that there are several data sets on water vapor above Summit that might allow more precise treatment of its impact on neutron flux, or at least provide an estimate of the magnitude of uncertainty introduced by neglecting changes in water vapor that are not just a function of pressure.

The only mechanisms by which atmospheric water vapor would affect our measurements would be (a) attenuating high energy "source" neutrons or (b) moderating fast neutrons produced locally. In regard to (a), we note that neutron monitor stations--which require a much higher precision than our instrument-- do not necessarily correct for water vapor. This is probably because water vapor accounts for only ~1% of the mass of the atmosphere (and less over greenland) and has attenuation properties similar to dry air (attenuation length = 95 g cm^-2 for water, as opposed to 130 g cm^-2 for air).

With regard to (b), the work of Rosolem et al. (2013) do imply some atmospheric vapor effect for a subaerially exposed detector. This effect occurs due to the moderation of fast neutrons as they scatter through the atmosphere. However, we emphasize that our

detector is buried in the snow, and that most of the recorded neutrons are produced in the snow. For this we add the statement at the beginning of section 3:

"We do not apply a correction for atmospheric water variability because nearly all fast neutrons impacting the sensor are produced in the snowpack, rather than the atmosphere."

Rosolem, R., W.J. Shuttleworth, M. Zreda, T.E. Franz, X. Zeng, and S.A. Kurc, 2013: The Effect of Atmospheric Water Vapor on Neutron Count in the Cosmic-Ray Soil Moisture Observing System. J. Hydrometeor., 14, 1659–1671, https://doi.org/10.1175/JHM-D-12-0120.1

I also find that the manuscript is a little sloppy, particularly in describing the corrections applied to convert measured neutron counts to SWE above the sensor. For example, in the discussion of equation 1 used to calculate the relative count rate N_r is defined as the reference sensor count rate while N_s is the reference count rate??? If $N_r = N_s$ then the relative count rate from this equation is always 1. Next page in discussion of equation 3 N_o is defined as the reference count rate (3rd ref ct rate) at the surface obtained before burial of the sensor (at time = 0). The term N/N_o in Equation 3 suggests N_o should be the count rate at SWE = 0 (i.e., flux reaching the snow surface), both N and N_o should be measured (estimated) at every time (it does not make sense to ratio N at each time to N_o measured just once, given time variations in both cosmic ray flux and water vapor/pressure).

We agree that Section 3 "Count Rate Correction and Conversion" was poorly presented and we have re-assigned several variables and rewritten the equations to be consistent with review given in Andreasen et al., 2017 (cited in revised paper). The specific points of confusion above have been addressed. We note that only a single "snow free" reference count is needed because the reference is corrected for solar and pressure variations. This is made more clear in the revised text.

Below are listed a variety of additional editorial comments (some are additional examples of sloppiness, a few more substantive), referenced by page/line #.

1/5 "background cosmic ray intensity" is probably not the correct term. What is really needed is variation in the neutron flux reaching the surface above the sensor at Summit, which could vary widely due to solar events (likely to dwarf changes in the flux of "cosmic ray background" impacting the solar system)

Replaced with "incoming cosmic radiation"

1/21 I would be very hesitant to claim that accumulation at Summit is "consistently low in June/July" based on less than 2 year record (not even considering prior results that find different results)

Replaced with "with the lowest accumulation in July and highest accumulation in the autumn of both years."

1/28 measuring the volume of accumulation (delete "of" before volume)

Corrected

2/14 mˆ2 (superscript)

Corrected

2/29 the statement here that "neutron counts increase with altitude and latitude" (more specifically geomagnetic latitude) demands that more be said later regarding how well a monitor at Thule can constrain neutron flux at Summit

Addressed above

3/4 "calibration data sets" suggest that there will be calibration data presented later. Turns out that all of the (critical) parameters in Table 1 appear to be taken from specs provided by Snowfox vendor.

Replaced with "validation".

3/12 and 13 Juxtaposing statement attributed to Alley, 1993 that Summit snow has "average surface density of 0.35 gm cmˆ-3" and citation of Dibb and Fahnestock, 2004 is sloppy. Latter paper presents density profiles from 22 "monthly" snowpits sampled at Summit over 2 years and shows that the mean density in the top 99 cm never exceeded 0.336 g cmˆ3 and averaged 0.305 g cmˆ3. This is also relevant to the Snowfox "validation" presented later. (Also note that the "-3" in manuscript should be superscript.)

We now use a mean top 10-cm density of 28 g cm^3 referenced to the recent SUMup datasets described in Montgomery et al. (2018). We also use the SUMup densities for the stake conversion described below. Superscript typo corrected.

3/18 MSF is an acronym for the "Mobile Science Facility". Until summer 2017 the main science facility at Summit was TAWO.

Corrected

3-4/25-30 and 1-5 (equations 1-3) see comments above. Also, N in Eq 3 is never defined (think this is the actual measured neutron count, at a given time T, from the buried Snowfox)

We now clearly define N as raw counts at the sensor.

4/12 Fig 2 does not show any time series, rather a curve based on assumed performance of the Snowfox sensor.

Corrected

Also confusing to introduce N*/No here, since Eq 3 defines N* to be a function of N/No.

Corrected and clarified

4/16-17 such that the resolution is (delete "that" before resolution)

Corrected

4/30 "42 observations" (snow cores)? ? Earlier in this paragraph it is stated that cores were sample every 10 days. 42 x 10 is 420. 13 Mar 17 to 17 Jan 18 is 310 days. Plot in Fig 7 seems to show 36 cores. These are not all consistent (sloppy)

Corrected.

5/1 Differences in the values of hw derived from any single core by weighing and by measuring the volume of melted snow are not due to "unconstrained errors in the sampling procedure." These are 2 different measurements of the same sample, so the errors have to be in the measurements.

Replaced with "errors in the measurement procedure".

5/3 Given estimate of the mean density in the snowpack from surface down to depth of the Snowfox from 42 (or 36 or 31, whatever may be the actual number) cores over 10 months, what can you say about 1) whether constant value of 0.35 g/cmˆ3 is reasonable, 2) is any variation in the measured density seasonal, 3) does it look like what Dibb and Fahnestock saw, 4)why not use these measured values to convert the stake measurements rather than a constant, loosely defined "surface" value from the literature?

We have revised our approach, as described in section 4.
The reason for the high density value used was that this was supposed to also account for compaction. To make this more clear, we have reformulated the conversion to separately account for the density (using the mean monthly SUMup values) and the compaction rate (as a tuning parameter qualitatively compared to Dibbs and Fahnestocks observations and Zwally and Li 2002's model). The result is very similar to that obtained from the constant density.

5/6 No good justification to use constant value for density, given that you have measured it at fairly frequent intervals, and that Dibb and Fahnestock showed that it is not constant (and was always lower than the assumed constant value used here).

See response to previous point. Also, the surface core densities may not be applicable to the snow stake conversion because 1) the depth of the sample varies with time 2) the base board may impact compaction and 3) this is a point measurement that may not apply over a large area. This is now explained in section 4.

5/19-20 There is an overall decline

Corrected

5/26 0.4 g cmˆ3 is a pretty high value for the density of a wind slab at Summit, also note that it is sloppy to change the units to kg/mˆ3 here

We clarify that these values here are approximate. Units are changed to g cm^{-3}

6/1 should "0.013 cm + 0.007" just be 0.02, or 0.013 +/- 0.007?

Replaced with "The best fit line predicts a standard deviation of 0.005 cm at h_w=0, increasing by 0.007 cm per cm of h_w…"

6/1-9 this paragraph does not support the very high accuracy claims made for Snowfox in the astract.

The abstract quotes exactly these values, which are now clarified to indicate precision, not accuracy.

6/19 "that much of the" (delete extra "the")

Corrected

6/21 seems that the overall rate should refer to 16 May '16 to 18 Jan '18 (not Jan '17)

Corrected to read "2018"

6/28 given that the Snowfox estimated SON accumulation differed by more than a factor of 2 between 2016 and 2017, how confident can one be that JJ are consistently low accumulation months based on the same almost 2 years of record.

Since we do not make a statement about "consistency" here, we assume the reference is to the statement in the abstract, which we have edited as described above.

6/30 "change in water equivalent" (add in)

Corrected

7/1-8 Make it clear what is signified by the "mean difference" (i.e., is Snowfox biased high or low by mean of 0.77 cm vs cores and 0.22 cm vs stakes).

Clarified.

Also consider redoing the stake comparison using a better estimate of density, with seasonal variation (from measured density of the cores in this study and/or values from Dibb and Fahnestock.) (Note that the agreement with Snowfox is likely to be worse using more realistic, lower, density for snow in the top 42 cm of SWE.

Redone as described above and in the text (section 4). The results have not significantly changed.

Also, why not show the comparison to stakes for the entire 20 months? It is unfortunate that the validation cores started almost a year late, but the stakes were measured monthly for a different project since 2003.

The entire time series comparison is now shown in Figure 7B.

Response to Reviewer 2

Thank you for your helpful comments and suggestions. Please note that paper has been updated using data obtained through 16 May, 2018 to give a full 2-year time series. This has not caused any major changes in our conclusions.

Reponses are provided in bold font to each comment in italics.

General Comments
This paper describes the application of a cosmic ray neutron-sensing instruments to measure snow accumulation in m water equivalents on the Greenland ice sheet. The manuscript is reasonably well written, though a bit sloppy. I appreciate the potential of this method for glacier applications, though the results are not very surprising or new since the method is already successfully used for several years in Alpine terrain. The manuscript should make is much more clear what the added value of this study is.

We have added text to the introduction to make the added value more clear.

Specific comments
My main comments regarding this manuscript are the estimated uncertainty, and the sloppy writing style.

The sloppy writing style results in ambiguities. To give some examples:
What is the difference between the 'reference sensor count rate Nr' and the 'reference count rate Ns',

We have rewritten the section and the equations to make these meanings more clear and to correct an error in the reference to N*. Our equations and variables now conform to Andreasen et al., (2017).

What is plotted in figure 2a? In the caption it is stated: corrected relative count rate, followed by N/N0, but in the text N* is the corrected relative count rate.

The label has been corrected. This is the corrected relative count rate N*.

In the list of corrections asked for below, more of these errors are noted.

Throughout the text when you refer to mass balance the units do not make it clear that you are talking about mass in m water. Since often mass balance observations are measured in m snow or ice, this is confusing.

We now consistently refer to cm of water.

Another point is the provided final uncertainty in the method. Only using daily scatter in hourly estimates to estimate the uncertainty does not account for uncertainties introduced by the background signal correction and the atmospheric moisture content correction, or the absolute uncertainty compared to the other two methods.

We have clarified that only the precision is estimated from the scatter in hourly estimated, with accuracy assessed from the validation datasets.

I also found the abstract not easy to comprehend without having read the manuscript. Since this method is not new, the introduction should be much more clear in what the additional values of this study is. The conclusions are much clearer about that. What I am also missing in the introduction is a comment to the fact that although this adds a method providing mass balance in m water, this still is a point or local measurement. It does not resolve the problem of obtaining a mass balance estimation for the total ice sheet. For that satellite methods and/or models are still the best option.

Added the following statement to the introduction:
"Finally, while cosmic ray sensing only provides local, in situ measurements of accumulation, continuous measurements from networks of sensors may be used to correct and validate ice-sheet wide mass balance estimates from atmospheric models and satellite altimetry."

Technical corrections
Abstract
P1 L10: Replace 'a reliable remote sensing method' by 'reliable remote sensing methods'.

Not changed - there is not a single reliable remote sensing method to obtain mass accumulation, so "methods" is redundant.

P1 L12: Replace 'four decades ago' with exact time period. In a couple of years, this is not correct anymore.

Replaced with "in the 1970's "

P1 L20: Is 'greater' correct English, perhaps 'larger' is better?

Changed to "to monthly"

P1 L21: Note that 2 years of observations is not enough to make any statement about seasonality.

Revised.

Introduction
P1 L28: Remove 'of' between 'the' and 'volume'.

Corrected

P2 L7: I am missing a sentence about the pro's and con's of modelling. For example that it still needs direct observations for evaluation of the results.

Statement added to introduction regarding potential application of cosmic ray measurement to constrain model and altimetry mass balance estimates.

P2 L23: 'soil moisture' in snow? Rephrase or explain.

Replaced with "in the moisture of the underlying soil"

P2 L23: Introduce abbreviation 'swe' and also use it throughout the text! It is not clear when you refer to m water or m snow/ice.
Snow water equivalent (swe), while common terminology in seasonal snow studies, is not common in glaciology/ice sheet studies. This is partly due to the ambiguity between snow and multi-year firn. We use snow water equivalent thickness here to refer to studies where alpine snowpack was measured, but refer to thickness of water throughout the paper. We have revised the paper to clarify where this may be ambiguous (we only refer to snow thickness when discussing the stake height conversion).

P3 L2: Add 'the' between 'in' and 'center'

Corrected

P3: Given the work by Kodama and Paquet and Laval, what does this study add? From the Introduction, that is not obvious. Also note that the presented method, although perhaps more than a point measurement as presented by a stake or sonic ranger, is still limited in horizontal extent. In that sense, you need other methods such as remote sensing and/or modelling. This limitation is not mentioned in the introduction.

Better justification and reference to spatial limitations added to introduction as described above.

Instrument Deployment

P2 L12: (and throughout the manuscript) I prefer a consistent use of units m, kg, m or mm w.e.. Now it varies throughout the manuscript.

Corrected

Count Rate Correction and Conversion
In general, this part is not well written, parameter abbreviations and descriptions are mixed up, also in the figures.

This section has been thoroughly revised.

P3 L22: Remove the first 'background'. Note that you use different terms for this correction: background cosmic ray and solar correction. This is confusing.

Revised

P3 L25-27: What is the difference between reference sensor count rate and reference count rate?

Clarified.
P4 L12: Not clear what is plotted in figure 2a, caption and axis do not correspond with description here. N*/N0, where does this value come from, and is this plotted? Same for next line.

Corrected as described above

P4 L13: 'increase more steeply' I understand what you mean, but the line in the figure is less steep, which is confusing with this text. Rephrase.

Revised

P4 L15-16: sentence about 'resolution' reaching 1 cm for 1 count, is not clear. Please rephrase.

This section has been completely revised for clarity.

P4 L17: remove 'that' before 'resolution'

Revised

P5 L10: How is this assessed? in absolute sense, of in standard deviations?

Revised and clarified.

Results
P5 L12: provide parameter abbreviation (N?)

Added

P5 L19: the 30-day oscillation is not really visible, especially since the cycle in the annual pressure correction is so obvious.

Removed reference to this variation and now refer generally to short-term variability.

P5 L19: Value 0.5 given, shouldn't this be 0.05?

Revised

P5 L28: The stable periods and declines are hard to distinguish.

Time periods and descriptions are clarified in the text.

P6 L1: Note that the variations are large, figures are mostly plotted with logarithmic scales.

The non-linearity of the curves are described in multiple places in the text, beginning at the end of section 3.

P6 L10: here 'mean daily water equivalent accumulation rate' is used, the caption of figure 5 is slightly different, as is the axis description. Please make them consistent.

Revised

P6 L10: The value of 0.078 is not visible in the figure, and which figure do you refer to? 5a?

Revised

P6 L21: add unit for the value of 16 May.

Added

P7 L1: is this the difference between the cosmic method and the cores?

Revised

P7 L2 and L8: explain what r2 is.

No longer included.

Summary and conclusions

P7 L13: emphasise that the values given by the cosmic method are direct mass values, whereas most other methods need a conversion from height change to mass providing a density.

Added emphasis.

P7 L16: On P6L1, the value given is 0.007, here a value of 0.0071 is given, be consistent.

Revised.

P7 L21: Note that 2 years of observations is not enough to make any statement about seasonality.

Revised.

P8 L13: correct 'wouled' to 'would'.

Corrected

Figures and tables
Table 1: What is cph stand for? Better use 'hPa' instead of 'mbar'.

Revised

All figures: provide (correct) parameter abbreviations in ALL figure captions and axes.

Revised

Figure 2a: not clear what is plotted here. Description in caption, axis and in text do not correspond. Also mention correct parameter abbreviations in ALL figure captions and axes.

Revised

Autonomous Ice Sheet Surface Mass Balance Measurements from Cosmic Rays

Ian M. Howat[1,2], Santiago de la Peña[1], Darin Desilets[3], Gary Womack[3]

[1]Byrd Polar & Climate Research Center, Ohio State University, Columbus, OH. USA
[2]School of Earth Sciences, Ohio State University, Columbus, OH. USA
[3]Hydroinnova LLC, Albuquerque, NM. USA

Correspondence to: Ian M. Howat (ihowat@gmail.com)

Abstract. Observations of mass accumulation and net balance on glaciers and ice sheets are sparse due to the difficulty of acquiring manual measurements and the lack of a reliable remote sensing method. The methodology for recording the water equivalent accumulation of snowfall using the attenuation of fast neutrons generated by cosmic ray impacts was developed in the 1970's and has been employed in large-network snowpack monitoring but has yet to be applied to glaciers and ice sheets. In order to assess this potential method, we installed a cosmic ray neutron-sensing device at Summit Camp, Greenland in April, 2016. Hourly neutron count was recorded for ~24 months and converted to water equivalent thickness after correcting for variability in atmospheric pressure and incoming cosmic radiation. The daily accumulation estimates are analysed for noise level and compared to manual surface core and snow stake network measurements. Based on measurements of up to 56 cm of water equivalent, we estimate the sensor's precision to be better than 1 mm for water equivalent thicknesses less than 14 cm, and better than 1 cm in up to 140 cm, or approximately 0.7%. Our observations agree with the surface core measurements to within their respective errors, with temporary biases that are explained by snow drifting, as supported by comparison to the snow stake network. Our observations reveal large temporal variability in accumulation on daily to monthly scales, but with similar annual totals. Based on these results, cosmic ray sensing represents a potentially transformative method for acquiring continuous, in situ measurements of mass accumulation that may add constraint to glacier and ice sheet mass balance estimates from meteorological models and remote sensing.

1 Introduction

Ice sheets and glaciers gain mass from the accumulation of snow and lose mass primarily from meltwater runoff and iceberg calving, with smaller amounts from sublimation and basal melting. Accurate measurements of these terms are necessary for assessing the contribution of land ice to rising sea levels. All methods for estimating glacier and ice sheet mass balance, with the exception of satellite gravimetry, require observations or model estimates of the mass accumulated per time. Multiple remote sensing methods exist for measuring the volume of accumulation, including repeat satellite or airborne altimetry and snow penetrating radar, but the density of the accumulation and, therefore its mass, are unknown. Mass accumulation rates are most commonly obtained through in situ sampling from snow pits and firn cores, typically with an annual resolution corresponding to identifiable seasonal layering. Such methods are laborious, logistically expensive and provide only a point

measurement affected by local variability due to drifting and scouring. Active radar imaging of the upper snow surface has been employed successfully to measure mass accumulation, but the power and maintenance requirements of these systems, and their sensitivity to meltwater, make them currently impractical for long term (> 1 year) autonomous deployment in remote locations. Other methods typically used to monitor seasonal mountain snow packs, including snow pillows and

5 mechanical scales, are ill suited to glaciers and ice sheets. Snow pillows require the transport or on site generation of 100's of kg of water and antifreeze to fill the pressure bladder, and will still freeze at polar temperatures. Both methods require a large, level surface for deployment and they may underestimate mass due to stress bridging by strong layers in the snowpack.

Cosmic ray neutrons are generated through collision of cosmic rays, high-energy particles generated from supernova, with

10 the Earth's atmosphere. The hydrogen in water attenuates such neutrons, with attenuation increasing predictably with the mass of water surrounding a measurement sensor. In a series of experiments, Kodama et al. (1975; 1979) and Kodama (1980) designed and deployed passive sensors that used the attenuation of cosmic ray neutrons by accumulating snowfall to estimate time series of snow water equivalent thickness of mountain snowpack. These sensors were able to measure daily water equivalent thickness with an accuracy of 3-4% (Kodama, 1980). Additionally, the sensor is sensitive to snowfall over

15 relatively large area (10's m^2), providing an aerially averaged estimate. Further, since the maximum limit of observable water thickness is determined by the minimum number of counts required to provide a statistically significant mean, the method could ostensibly work in water equivalent thicknesses exceeding 10 m. While this method was further refined and adapted successfully to monitoring soil moisture (Kodama, 1985), it was not widely applied to measuring snowpack until 1998 when the French electric utility installed a network of 40 cosmic ray snow gauges for hydroelectric monitoring (Paquet

20 and Laval, 2005; 2008). An extensive comparison between snow cores and cosmic ray neutron sensor (CRNS) estimates revealed accuracies in water equivalent thickness between 12 and 20%, with much of the discrepancy due to spatial variability in the snowpack between the cores and the sensors, as well as a significant uncertainty due to variations in the moisture of the underlying soil. Accounting for these differences resulted in hourly water equivalent thickness estimated with accuracies better than 5% (Paquet and Laval, 2008), consistent with the results of Kodama (1980).

25 Cosmic ray sensing therefore provides a potentially effective method for measuring mass balance in the accumulation zones of ice sheets and glaciers. Since the cosmic ray neutron count rate is only sensitive to the mass, and not the density, of the firn, it integrates the processes of snowfall, sublimation, deposition, and vertical vapour and meltwater fluxes into a single measurement of local mass balance. Glaciers and ice sheets are also particularly suitable to cosmic ray sensing because, firstly, neutron counts increase with altitude and latitude, due to decreasing atmospheric attenuation, which increases the

30 accuracy and resolvable maximum thickness for a given temporal resolution. Secondly, unlike seasonal snowpack, there is no uncertainty associated with variable soil moisture or, in the case of cold polar ice sheets, horizontal water transport. Thirdly, the sensor's effective cone of measurement provides an aerial average that should be less sensitive to spatial variability caused by drifting. Fourthly, since the sensor is passive and primarily consists of only polyethylene in a steel case, it is lightweight, compact, durable and has a low power requirement, providing ease of deployment in extreme polar

environments. Finally, while cosmic ray neutron sensing only provides a local measurement of accumulation, continuous measurements from networks of sensors may be used to correct and validate ice-sheet wide mass balance estimates from atmospheric models and satellite altimetry.

Here we assess the potential for cosmic ray neutron sensing of glacier and ice sheet surface mass balance through deployment of a sensor at Summit Camp (72.57°N, 38.46°W), located 3216 m above sea level in the centre of the Greenland Ice Sheet. We describe the deployment setup, the characteristics of the raw neutron count data, the correction and validation datasets and compare the raw to the corrected count data and water equivalent accumulation estimate. We present our daily and monthly water equivalent accumulation rates at Summit Camp and then compare those estimates to manual observations of surface accumulation for validation.

2 Instrument Deployment

Summit Camp was chosen for its continuous power supply and climate-controlled instrumentation housing. This, and the year-round presence of support staff to troubleshoot if needed, simplified this initial deployment and reduced the risk of power or communications failure, allowing the focus to be only on assessment of the cosmic ray neutron counting methodology. Summit Camp personnel were also able to perform the validation surveys. Summit Camp has an annual water equivalent accumulation of 24 cm and an average surface density of 0.28 g cm^{-3} (Alley et al. 1993; Montgomery et al. 2018). Snowfall occurs throughout the year, with some uncertainty about the seasonality of accumulation (Dibb and Fahnestock, 2004).

We installed a Hydroinnova Snowfox™ cosmic ray neutron-sensing instrument at Summit Camp on April 30th, 2016. The Snowfox™ is a 81-cm long and 20 cm diameter tube that was placed horizontally in a shallow trench in the firn so that the top of the Snowfox™ was ~20 cm below the surface (Fig. 1). The trench was then allowed to fill with wind-blown snow, burying the sensor. A 100-m long power and communications cable connects the sensor to a data logger, telemetry modem and continuous power supply housed in the climate-controlled Mobile Science Facility (MSF) at Summit Camp. The sensor recorded hourly counts of neutron impacts, as well as hourly average barometric pressure and temperature.

3 Count Rate Correction and Conversion

To obtain an estimate of the water equivalent thickness of snow accumulation, h_w, from hourly counts of neutrons recorded by the sensor, N, corrections must be applied to account for variability in barometric pressure and incoming cosmic radiation (Andreasen et al., 2017). We do not apply a correction for atmospheric water variability because nearly all fast neutrons impacting the sensor are produced in the snowpack, rather than the atmosphere.

A correction factor accounting for variations in incoming cosmic ray intensity is obtained from the atmospheric pressure-corrected counts observed simultaneously at a second, reference neutron sensor located above the snow surface. If the corrected count rate at the reference sensor is I_m, the correction factor, f_i, is:

$$f_i = 1 + \beta \left(\frac{I_0}{I_m} - 1 \right), \tag{1}$$

where I_0 is an arbitrary, time-invariant reference rate, such as the long-term mean, and β is a scaling parameter. For application to Summit Camp, we use the neutron monitor located at Thule (THUL) Greenland operated by the Bartol Institute at the University of Delaware and distributed via the Neutron Monitor Database (www.nmdb.eu/nest/). The unitless scaling parameter β accounts for differences in the magnitudes of solar-induced variations between the reference and local sensors due to differences in latitude and elevation, only the latter of which will be significant in this case (Hawdon et al. 2014). We use a value of $\beta = 1.19$ based on regressions to the global neutron monitor dataset. While a single value produces unbiased corrections on weekly or longer timescales, we expect some errors associated with short term variation, such as individual Forbush decrease events.

Following Zreda et al. (2012), the barometric pressure correction factor, f_p, is obtained from:

$$f_p = exp \left(\frac{P_0 - P}{L} \right), \tag{2}$$

where P is the observed barometric pressure at the sensor at the time of the neutron count measurement and P_0 is an arbitrary, time-invariant reference pressure. The mass attenuation length for high-energy neutrons, L, is 130 g cm^{-2} at the latitude of Summit Camp (Desilets et al. 2006). Applying the incoming cosmic ray intensity and barometric pressure correction factors to the raw counts at the sensor, N, gives the corrected count rate:

$$N_{cor} = f_i f_p N, \tag{3}$$

which is then smoothed with a daily moving average and divided by the solar and pressure-corrected count rate at the surface obtained prior to burial of the sensor to obtain the relative corrected count rate, N^*. The water equivalent accumulation thickness, h_w, is then obtained from the relative corrected count rate N^* as:

$$h_w = -\Lambda^{-1} \log N^*, \tag{4}$$

with:

$$\Lambda = \frac{1}{\Lambda_{max}} + \left(\frac{1}{\Lambda_{max}} - \frac{1}{\Lambda_{min}} \right) \left(1 + exp \left[\frac{a_1 - N^*}{a_2} \right] \right)^{-a_3}. \tag{5}$$

The empirical, location-independent parameters Λ and a are determined through calibration and field validation experiments by the sensor manufacturer. Their values are listed in Table 1. The resulting relationship between h_w and N^* is plotted in Fig. 2A. There is a gradual increase in h_w with decreasing N^* to $h_w = 34$ cm at $N^* = 0.3$. Below this value for N^*, h_w increases more steeply, rising to 200 cm at $N^* = 0.1$, 300 cm at $N^* = 0.04$ and 490 cm at $N^* = 0.01$. Thus, the sensitivity of the neutron counter to changes in water equivalent thickness decreases non-linearly with increasing thickness. Figure 2B shows the change in h_w versus the change in corrected count rate, N_{cor}, for zero water equivalent thickness ($h_w = 0$) corrected count rates, N_{cor}^0, of 2000, 4000 and 8000 counts per hour (cph). The count rate is ~25 times more sensitive to variations in h_w at 10 cm than at 100 cm, and 50 more times sensitive than at 200 cm. The inverse of this sensitivity is the measurement resolution; the larger the change in h_w per change in corrected count rate, N_{cor}, the coarser the resolution of the neutron counter. There is,

therefore, a corresponding, nonlinear decrease in the resolution of h_W with increasing h_W, as well as a linear decrease with a decreasing N_{cor}^0. At h_W=100 cm, the sensor resolutions are 0.50, 0.23 and 0.11 cm of h_w per cph, for N_{cor}^0 of 2000, 4000 and 8000 cph, respectively. The resolution declines to a 1 cm change in h_W per 1 cph change in N_{cor} near 400 cm for a N_{cor}^0 of 4000 cph. However, the fractional change is such that the maximum resolution is better than 0.25% of h_w for h_w between 5 and 350 cm. These are the maximum resolutions based only on the relationship between count rate and water equivalent thickness in Eqns. 4 and 5. The actual uncertainty in h_w will therefore be this resolution in addition to errors in the corrections for variability in incoming cosmic rays and barometric pressure, as well as sensor noise.

4 Validation Datasets

In order to validate the cosmic ray neutron sensor (CRNS) observations, water equivalent accumulation was measured manually every ~8 days beginning March 17, 2017 from a location approximately 10 m from the CRNS. The manual observations utilized the "snow board" method in which a shallow, rectangular pit is excavated and a piece of plywood is placed at the floor of the pit. The pit is then allowed to fill with snow and settle over a period of ~2 weeks. A PVC tube is used to remove a core sample of the snow from the surface to the plywood, which serves as a depth reference for each subsequent sample. The sample is taken from a different location each time, as measured from flagged poles at the corner of the plywood, to provide an undisturbed sample. The surface snow core is then weighed to the nearest 0.1 g, the weight of the core tube is removed and the snow weight is divided by the cross sectional area of the core to give a measurement of h_w. For redundancy, the snow core sample is allowed to melt and the water volume is recorded to the nearest mL. This volume is divided by the cross sectional area of the core to give another measurement of h_w. The difference between these two measurements provides a check on the accuracy of the sample. For the 54 observations, the mean difference was 0.20 cm water equivalent thickness with a standard deviation of 1.48 cm, and with the water volume measurment giving a larger estimate on average than the weight measurement. This standard deviation is larger than the uncertainty predicted by the measurement precisions and, therefore, may be due to unconstrained errors in the measurment procedure. We therefore assume ±1.48 cm water as the error in h_w obtained from the surface cores. We note that the mean density of the snow core sample is also recorded. However, the time-varying depth of the sample and the possible impact of the base board on the compaction, via effects on vertical heat and vapour fluxes, makes the applicability of these density values uncertain.

We additionally compare CRNS h_w estimates to ~weekly accumulation measured from a network of 120 stakes (i.e. the "bamboo forest") at Summit Camp (Dibb and Fahnestock, 2003). The change in surface height measured through repeat snow stake measurements (here Δh_s, SHC in Dibbs and Fahnestock, 2003) gives the integrated change due to surface mass balance processes and densification of the firn into which the stake is anchored (initially ~1.2 m and increasing with burial). A major benefit of the snow stake measurements is that they provide an estimate of average variability over a relatively large area, as opposed to the point measurement provided by the surface core samples. This enables an assessment of the possible influence of spatial variability on surface core measurements relative to the CRNS.

The relationship between changes in snow water equivalent thickness, Δh_w, as measured by the CRNS, and mean surface height measured by the snow stake network, Δh_s, can be approximated by:

$$\Delta h_w = \rho^*(\Delta h_s - \Delta h_f), \qquad\qquad\qquad\qquad\qquad (6)$$

where Δh_f is the change snow or firn thickness from the stake bottom to previous measurement surface and ρ^* is the ratio to water of the mean density of the mass either added to or subtracted from the surface (i.e. the surface density). We use mean monthly ρ^* for the top 10 cm from the Surface Mass Balance and Snow on Sea Ice Working Group (SUMup) dataset (Montgomery et al. 2018). These range from a minimum density of 0.23 g cm^{-1} in August to a maximum of 0.33 g cm^{-1} in April, with an average standard deviation of 0.06 g cm^{-1} or 20%. The compaction term Δh_f will also vary seasonally with snow temperature but has a much higher relative uncertainty due to inter-annual variability, variations in accumulation history and the changing depth of the stakes. Model estimates by Zwally and Li (2002) predict seasonal variations in compaction ranging from approximately 10 to 40 cm a^{-1}, with a value of 18 cm a^{-1} based on mean temperatures and steady state accumulation. Thus, for a ρ^* of 0.28 ±0.06 g cm^{-1} and a Δh_f of -0.39 ±0.20 cm (18 cm a^{-1} over the average 8 days between measurements), and assuming ρ^* and an Δh_f are uncorrelated, an average Δh_s of 1.70 cm yields an Δh_w of 0.37±0.10 cm water, or an uncertainty of 26%. While this error is small relative to the surface core measurements, these errors may be systematic, resulting in a larger cumulative error. Here we apply Eqn. 6 using the average monthly SUMup densities and by adjusting Δh_f to give the best fit between Δh_w from the CRNS and Δh_s, comparing the result to expected values of Δh_f based on previous observations.

5 Results

Fig. 3A plots the time series of uncorrected hourly neutron counts, N, recorded by the CRNS. Starting from $N=7000$ cph, when the sensor was exposed at the bottom of the firn trench, the count rate dropped to ~4200 cph over 16 days as the trench filled with snow. The count rate then held above 3000 cph until March, when it dropped to 2200 cph by May. The rate then stayed between 2000 and 3000 cph through the end of the record.

Corrections for solar activity, f_i, and atmospheric pressure, f_p, are shown in Fig. 3B. Short term (weekly) variations with an amplitude of ±0.02 are visible in f_i with an overall decline of 0.11 over the period punctuated by a 0.05 increase in July 2017. Relative variability, and thus the impact on count corrections, is larger for f_p, with short term (days) variablity of up to ±0.1 and an annual cycle with an amplitude of ±0.12 and a maximum in July and minimum in January. Applying these corrections to the raw count data gives the corrected series shown in gray in Fig. 3A. The correction substantially smooths the variabilty in count rate yields and a more linear decline of ~1000 cph between July 2017 and May 2018. The corrected count rate is then divided by the initial corrected count rate when the sensor was uncovered (8181 cph) to give the corrected relative count rate and snow water equivalent thickness, h_w, shown in Fig. 3C. The initial, rapid 8 cm water equivalent (w.e.) increase is consistent with infilling of the approximately 20-cm deep trench assuming a density of wind packed snow of 0.4 g cm^{-3}. From June to September, 2016, h_w remains near constant before increasing to 30 cm w.e. by the following spring. Another

stable period of h_w occurs in the summer 2017, followed by another 15 cm w.e. increase between October 2017 and May 2018, interupted by a loss of 3 cm w.e. between 25 March and 4 April that was regained by 7 April. This abrupt change is visible as a spike in the raw count rate that is not corrected by f_i or f_p.

We expect the noise level in h_w to increase as N^* decreases (and h_w increases) because the noise in N^* becomes larger relative to the zero snow count rate and the signal to noise ratio, and thereofore the measurement precision, decreases. This noise increase is visible in the curves in Fig. 3C. We assess the increase in noise, or loss of precision, by plotting the daily mean of hourly h_w estimates against their standard deviation, omitting the period of rapid trench infilling prior to 14 May, 2016 (Fig. 4). The standard deviation in water equivalent thickness increases from under 0.1 cm at h_w=15 cm to 0.5 cm at h_w=50 cm. The best fit line predicts a standard deviation of 0.005 cm at h_w=0, increasing by 0.007 cm per cm of h_w, so that the standard deviation would reach 1 cm by h_w=140 cm. The lower range of daily standard deviations corresponds closely to the water equivalent thickness-dependent resolution of hourly h_w measurements derived in Section 3. However, the best fit line of daily mean h_w to its standard deviation is ~10 times larger than the thickness-dependent resolution (Fig. 4), indicating that noise in the count rate is the dominant source of error. In relative terms, the standard deviation drops below 1% for h_w larger than 5 cm, declining to 0.7% for h_w greater than 30 cm. The increasing noise however, also decreases the precision of change measurements, given by the root of two times the squared deviations, or 0.01 cm per cm of h_w, assuming daily errors are uncorrelated. Thus the standard deviation in daily change measurements is approximately ±1% of h_w.

Differences between daily mean water equivalent thickness, h_w, are shown in Fig. 5A. The mean change in daily mean h_w is 0.07 cm w.e. with a standard deviation of 0.37 cm w.e. . The maximum single day of accumulation was 2.2±0.3 cm w.e. on 15 September, 2017, while the maximum negative accumulation (ablation) was -1.3±0.6 cm w.e. on 25 February, 2018. After the initial period of infilling of the trench, including two days with near 2 cm w.e. day^{-1} on 14 and 15 May, 2016, there was a sustained period of low deviation in daily rates from June to October, 2016, followed by increasingly large scatter. We expect increasing noise in these data due to a declining relative neutron count rate with burial. Plotting the daily change in h_w as a percentage of h_w (Fig. 5B), we find a consistent daily variation throughout the record with a standard deviation of ±1.2%, consistent with ±1% precision error determined from the daily standard deviation of hourly measurements shown above. We would also expect variability to correspond with wind strength due to the effect of drifting and scouring. Some anomalously rapid changes in water equivalemt thickness, such as the 2.2 cm w.e. gain on 15 September, 2017, occurred on days with high winds (Fig. 5C). Additionally, the loss of 3 cm w.e. between 25 March and 4 April occurred during a period of sustained high winds, including three days of > 10 m/s mean wind speed on March 23, 25 and 26. Overall, however, the correlation between mean or maximum (not shown) daily wind speed and accumulation rate is not significant, indicating that much of the variability is due to noise and other factors.

We plot the monthly accumulation rate, calculated as the change in monthly mean h_w, normalized for days in each month, in Fig. 6. The monthly accumulation rate shows a high degree of variability, with a standard deviation of 57% of the 1.9 cm

w.e. per 30-day mean. In both 2016 and 2017 the lowest accumulation rate was in July, with the July 2016 accumulation totalling only 0.2 cm w.e., or 11% of the mean annual rate. November and October were the highest accumulation months in 2016 and 2017, respectively. For the 7 months (June through December), measured in both 2016 and 2017, those in 2017 had higher accumulation rates in all but August, with more than double the 2016 rates in October and November. Despite this large difference, total annual accumulations where similar for the two years. In the first year of the record, from 15 May, 2016 (after infilling of the trench) to 14 May, 2017, the total accumulation was 23.2±0.2 cm w.e. yr^{-1}, while the total accumulation in the 2nd year, between May 16, 2017 and May 15, 2018 , was 24.1±0.5 cm w.e. yr^{-1}.

6 Validation

While the analysis of short-term variablity in accumulation presented above provides a constraint on the precision of the CRNS, the accuracy may be assessessed only through validation. We compare the change in water equivalent accumulation thickness, Δh_w, estimated from the CRNS to surface core and snow stake observations (Fig. 7). Since both the surface core and snow stake observations have a one day timestamp, we compare these data to the daily means of the hourly cosmic ray estimates. On average, Δh_w measurements obtained by differencing the ~8-day surface core measurments are 0.04 cm w.e. greater than the cosmic ray estimates, with a standard devation of 1.72 cm w.e., which is less than the combined errors of the snow core difference measurments (±2.09 cm w.e.) and the CRNS (±0.30 cm w.e.). The cumulative Δh_w over the observation period averaged 1.68 cm w.e. greater for the surface cores, with a standard deviation of 2.03 cm w.e. (Fig. 7A). More separation between the surface core and cosmic ray estimates is apparent after 30 January, 2018, when the surface core recored an increase of 2.5 cm w.e. by 7 February, followed by a 3 cm w.e. increase between 20 and 27 March. This latter increase corresponds with a period of high winds, as mentioned above, and occurs immediately before the abrupt 3-cm w.e. decrease recorded by the CRNS. Prior to 30 January, 2018, the three largest outliers occur on 19 September 2017, 9 November 2017 and 17 January 2018 and all after large (> 2 cm w.e.) accumulations recorded by the surface core. The difference then goes to near or greater than zero in the next one or two core measurements. The CRNS estimated approximately one standard deviation lower Δh_w than the surface cores in June 2017, moving back to zero difference by the end of August. These large deviations, followed by gradual convergence, suggest spatial variablity due to drifting, the effect of which may have increased after 30 January, 2018. Prior to this, the positive bias in cumulative Δh_w recorded by the surface core was only 0.76 cm with a standard deviation of 1.19 cm w.e.. The surface core Δh_w had returned to within 2.24 cm w.e. of the CRNS estimate as of the most recent measurement (8 May, 2018).

We compare the measurements of changes in surface height at the network of snow stakes to estimates of water equivalent thickness change from the CRNS using the conversion approach described in Section 4. Using the monthly SUMup density dataset (Montgomery et al. 2018), we obtain a consistently close fit between the snow stake and cosmic ray estimates by applying a compaction terms (Δh_f in Eqn. 6) of 0 cm day^{-1} from the start of the record until 1 January, 2017 and 0.05 cm day^{-1} (18.25 cm a^{-1}) thereafter. The latter is close to the seasonal compaction rate (18 cm) modeled by Zwally and Li (2002)

and is 40-50% larger than that estimated from two years of snow pits by Dibb and Fahnestock (2004). While zero compaction is not realistic, it could be explained by a lower than average surface density the in 2016. Applying a 0.05 cm day^{-1} compaction rate throughout the record and uniformly decreasing the mean surface density by 25% prior to 1 January, 2017 gives nearly identical results.

5 On average, the Δh_w estimated from the snow stake measurements is 0.01 cm w.e. less than the cosmic ray measurments for periods between stake measurements, with a standard deviation of 0.98 cm w.e.. Cumulatively, the snow stakes estimated an average Δh_w 0.35 cm w.e. greater than the CRNS with a standard deviation of 0.99 cm w.e. (Fig. 7B). The snow stake measurements, as with the cosmic ray measurements, tend to show smaller variability than the surface core measurements and do not show the large increases and decreases during the outler events mentioned above, nor the large loss and rebound

10 recorded by the CRNS between 25 March and 7 April, 2018. This provides further evidence of the effect of drifting on the surface core and, possibly, CRNS relative to the large-area average provided by the snow stake network.

7 Summary and Conclusions

Cosmic ray neutron sensing offers a potential method for obtaining practical, autonomous, in situ mass balance measurements on glaciers and ice sheets, providing continuous measurements of changes in surface mass, rather than

15 volume. To test this potential, we deployed a cosmic ray neutron sensor (CRNS) at Summit Camp, Greenland between April 2016 and May 2018. Based on the the daily scatter in hourly measurements, we obtain a 1σ error estimate of 0.005 cm water equivalent (w.e.) for 0 cm w.e., increasing with thickness at a rate of 0.007 cm w.e. per cm of water equivalent thickness, giving errors of 0.1, 0.5 and 0.7 cm w.e. in depths of 10, 50 and 100 cm w.e., respectively and daily-difference errors of approximately 1% the depth. We observe single-day accumulation and ablation events of up to approximately 2.2 and -1.3

20 cm w.e. per day, respectively, with daily variabilty due to both increasing sensor noise with depth and wind drifting and scouring. Monthly accumulation rates show large month-to-month variability, exceeding 50% of the mean, with a minimum accumulation in July of consecutive years and a maximum in autumn.

We validate CRNS measurements through comparison to repeated surface snow core and stake network measurements. We detect little or no bias in CRNS water equivalent thickness change estimates for the ~8-day periods between surface core and

25 snow stake measurements, with the standard deviation of the differences between estimates near or within the expected errors. Spatial variablity due to drifting results in periods of divergence in the cumulative thickness change recorded by the surface core from both the CRNS and snow stake measurements. Divergence of cumulative thickness change between the CRNS and snow stakes estimates, particularly during one large event in March, 2018, indicates that the CRNS is also suseptiable to short-term (days to weeks) variablity, due to either to count correction errors or drifting that is not

30 representative of the large surface area captured by the snow stake survey. Thus, a challenge to further validation will be obtaining validation observations that are of high enough accuracy and have similar spatial and temporal resolution as the CRNS.

Our test supports cosmic ray neutron sensing's potential as an effective and practical method for obtaining, for the first time, continuous and autonomous measurements of surface mass balance within accumulation zones (i.e. where annual snowfall exceeds ablation). The very high accuracy of the instrument, exceeding 1 mm per hour, opens up the possibility of acquiring mass balance measurements over the low accumulation, but vast interior of the Antarctic Ice Sheet, where few direct measurements of surface mass balance exist and which represents among the largest sources of uncertainty in ice sheet mass balance measurements. These low accumulation, polar regions are also the most difficult to measure manually due to the limited precision of core samples and the lack of chronology in the upper firn. In areas with higher accumulation, the increase in noise and loss of resolution with depth can be mitigated by increasing the count period, so that the duration of autonomous sensor deployment would be most likely limited by the instrument power and communications (e.g. the height of the tower supporting the telemetry antenna and solar panels or the battery lifespan). This still could be several years in the case of very high accumulation, or decades in the case of polar deserts. Finally, additional accuracy on short time scales (< 1 week) may be obtained by deploying a local reference sensor above the surface, reducing uncertainties count corrections.

The portability, ease of deployment and low power of this passive sensor are ideal for measuring accumulation in remote locations, where manual measurements (i.e. cores and snow pits) are currently cost or logistically prohibitive. Combining the CRNS with observations commonly made by automated weather station observations, including temperature, wind speed and repeat measurements of surface height by echo sounder, would provide new information about the processes of wind redistribution and firn compaction, for which mass and density are currently unknown variables. This information is critical for obtaining ice sheet mass balance from repeat altimetry measurements, such as from NASA's ICESat missions. Finally, these measurements would inform regional and ice-sheet scale surface mass balance models for which direct water equivalent accumulation and wind distribution observations are currently sparse or non-existent.

Since our implementation requires burial in the underlying firn, the CRNS is most applicable for measuring accumulation where meltwater infiltration is shallow enough that water does not infiltrate below the depth of the sensor. It is possible, however, that the sensor could be used to measure water transport at the surface or in the firn by observing the decrease in mass during periods of no precipitation, in a similar manner currently used for soil moisture measurements. Finally, borehole applications may exist for measurement of basal and englacial water transport.

Acknowledgments

This work was supported by U.S. National Aeronautics and Space Administration Grant NNX14AH90G and the Ohio State University Office of Research. The authors thank the Summit Camp staff for their diligent maintenance of the instrument and collection of validation observations.

References

Alley, R., Meese, D., Shuman, C., Gow, A., Taylor, K., Grootes, P., White, J., Ram, M., Waddington, E., Mayewski, P. and Zielinski, G.: Abrupt Increase in Greenland Snow Accumulation at the End of the Younger Dryas Event, Nature, 362, 6420, 527-529, doi: 10.1038/362527a0, 1993.

Andreasen, M., Jensen, K.H., Desilets, D., Franz, T.E., Zreda, M., Bogena, H.R. and Looms, M.C.: Status and Perspectives on the Cosmic-Ray Neutron Method for Soil Moisture Estimation and Other Environmental Science Applications, Vadose Zone Journal, 16, 8, doi: 10.2136/vzj2017.04.0086, 2017.

Desilets, D., Zreda, M. and Prabu, T.: Extended scaling factors for in situ cosmogenic nuclides: New measurements at low latitude, Earth Planet.Sci.Lett., 246, 3-4, 265-276, doi: 10.1016/j.epsl.2006.03.051, 2006.

Dibb, J., and Fahnestock, M.: Snow accumulation, surface height change, and firn densification at Summit, Greenland: Insights from 2 years of in situ observation, Journal of Geophysical Research-Atmospheres, 109, D24, D24113, doi: 10.1029/2003JD004300, 2004.

Hawdon, A., McJannet, D. and Wallace, J.: Calibration and correction procedures for cosmic-ray neutron soil moisture probes located across Australia, Water Resour.Res., 50, 6, 5029-5043, doi: 10.1002/2013WR015138, 2014.

Kodama, M.: Continuous Monitoring of Snow Water Equivalent using Cosmic-Ray Neutrons, Cold Reg.Sci.Technol., 3, 4, 295-303, doi: 10.1016/0165-232X(80)90036-1, 1980.

Kodama, M., Kudo, S. and Kosuge, T.: Application of Atmospheric Neutrons to Soil-Moisture Measurement, Soil Sci., 140, 4, 237-242, doi: 10.1097/00010694-198510000-00001, 1985.

Kodama, M., Nakai, K., Kawasaki, S. and Wada, M.: Application of Cosmic-Ray Neutron Measurements to the Determination of the Snow-Water Equivalent, Journal of Hydrology, 41, 1-2, 85-92, doi: 10.1016/0022-1694(79)90107-0, 1979.

Montgomery, L., Koenig, L. and Alexander, P.: The SUMup dataset: Compiled measurements of surface mass balance components over ice sheets and sea ice with preliminary analysis over Greenland, Earth Syst. Sci. Data Discuss., doi: https://doi.org/10.5194/essd-2018-21, in review.

Paquet, E., M. Laval, L.M. Basalaev, A. Belov, E. Eroshenko, V. Kartyshov, A. Struminsky and V. Yanke (2008), An Application of Cosmic-Ray Neutron Measurements to the Determination of the Snow Water Equivalent, paper presented at Proceedings of the 30th International Cosmic Ray Conference, Merida, Mexico.

Paquet, E., and Laval, M.: Operation feedback and prospects of EDF cosmic-ray snow sensors, Houille Blanche-Revue Internationale De L Eau, 2, 113-119, doi: 10.1051/lhb:200602015, 2006.

Zreda, M., Shuttleworth, W.J., Zeng, X., Zweck, C., Desilets, D., Franz, T.E. and Rosolem, R.: COSMOS: the COsmic-ray Soil Moisture Observing System, Hydrology and Earth System Sciences, 16, 11, 4079-4099, doi: 10.5194/hess-16-4079-2012, 2012.

Table 1. Parameter Values Used in Equation 5.

Λ_{max}	1.144×10^2
Λ_{min}	14.11
a_1	3.133×10^{-1}
a_2	8.268×10^{-2}
a_3	1.117

Figure 1: Photograph of Hydroinnova Snowfox™ cosmic ray neutron sensor (CRNS) exposed in firn trench on the first day of deployment. Inset shows the location of Summit Camp on the Greenland Ice Sheet, as denoted by the red circle. The red triangle shows the location of the THUL reference neuton monitor maintained by the University of Delaware Bartol Institute.
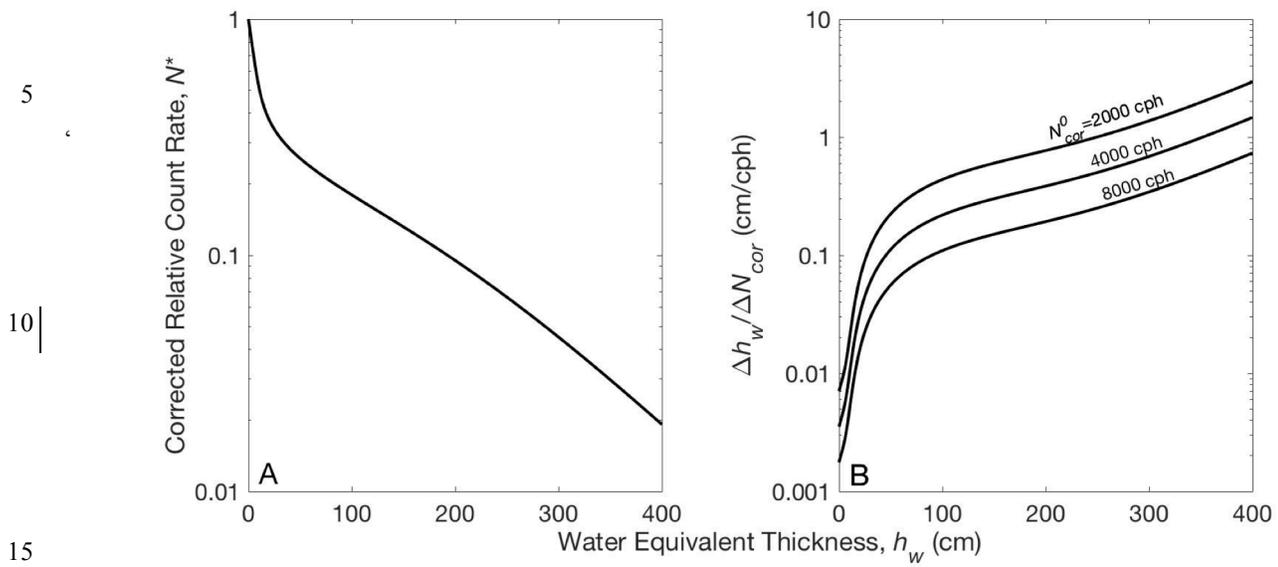
Figure 2: (A) Relationship between water equivalent thickness, h_w and the relative hourly corrected count rate, N^*, measured by the cosmic ray neutron sensor (CRNS), as obtained from Equations 4 and 5 and the parameters provided in Table 1. (B) The change in h_w per change in hourly corrected count rate, N_{cor}, for count rates at $h_w = 0$ (i.e. the corrected count rate when no snow is covering the sensor, or N_{cor}^0) of 2000, 4000 and 8000 counts per hour (cph).
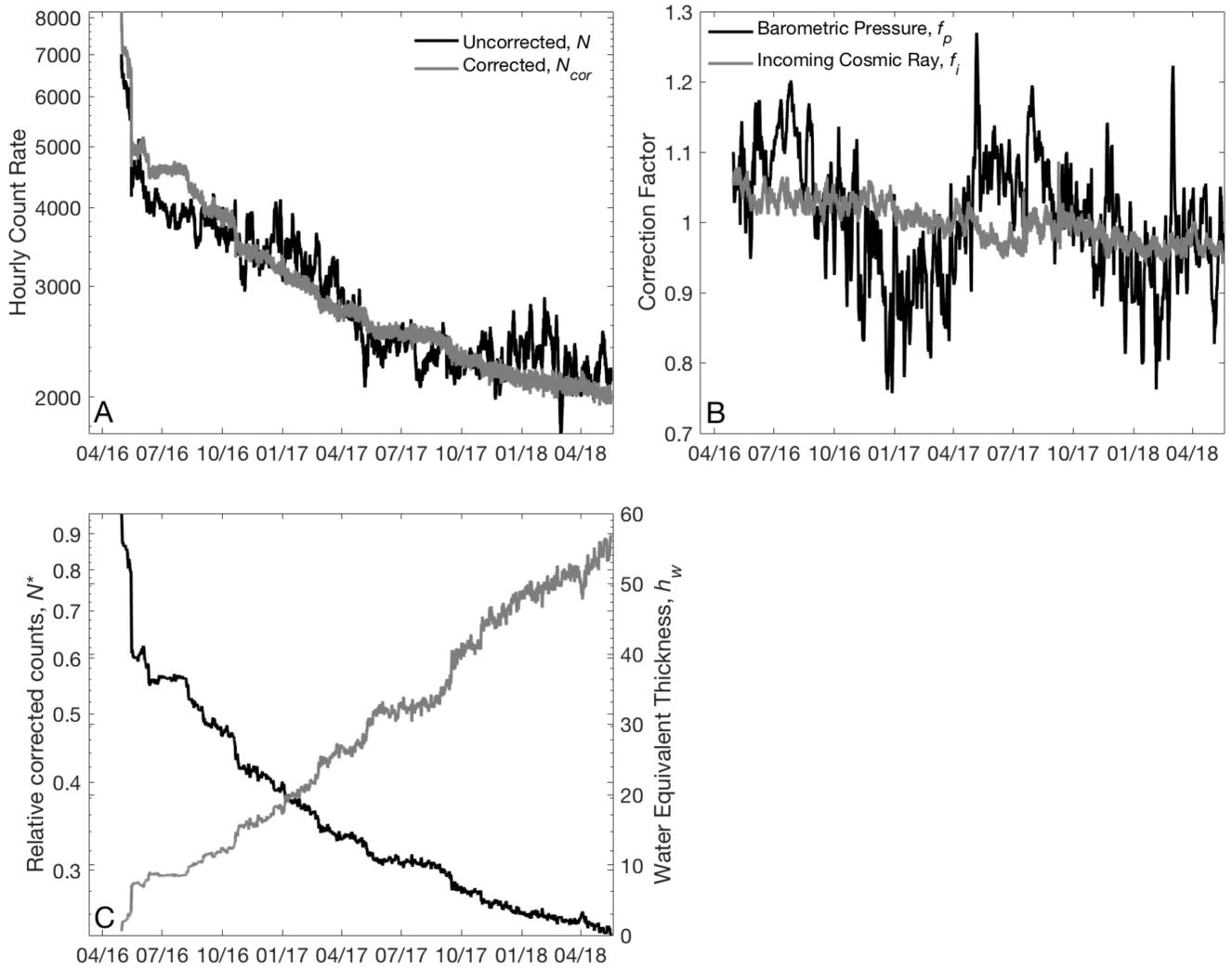
Figure 3: (A) (black) Uncorrected, *N,* and (grey) corrected, *N_{cor}* hourly neutron count rate measured at **Summit Camp**, **Greenland by a cosmic ray neutron sensor (CRNS) emplaced in the firn. (B)** Count rate correction factors for (black) atmospheric pressure, *f_p* and (grey) solar activity, *f_i*, as derived from equations 1 and 2 in Section 3 of the text and the Thule (THUL) neutron monitor for reference. **(C)** (black) Relative corrected counts, *N*,* and resulting (grey) water equivalent thickness, *h_w* from **Equations 4 and 5** and the parameters in Table 1.
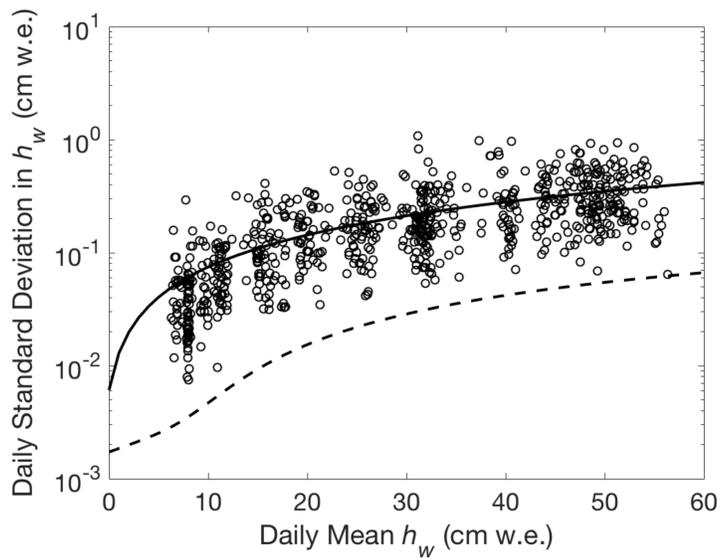
Figure 4: Daily means of hourly water equivalent (w.e.) thickness, h_w, **measured by the cosmic ray neutron sensor (CRNS), versus the daily standard deviation in hourly measurements. Solid curve is the line of best fit, with a slope of 0.007 and a y-intercept of 0.005 cm. Dashes are the change in** h_w **per change in hourly relative corrected count rate** $N*$ **given by equation 4, which is the maximum resolution of the CRNS.**
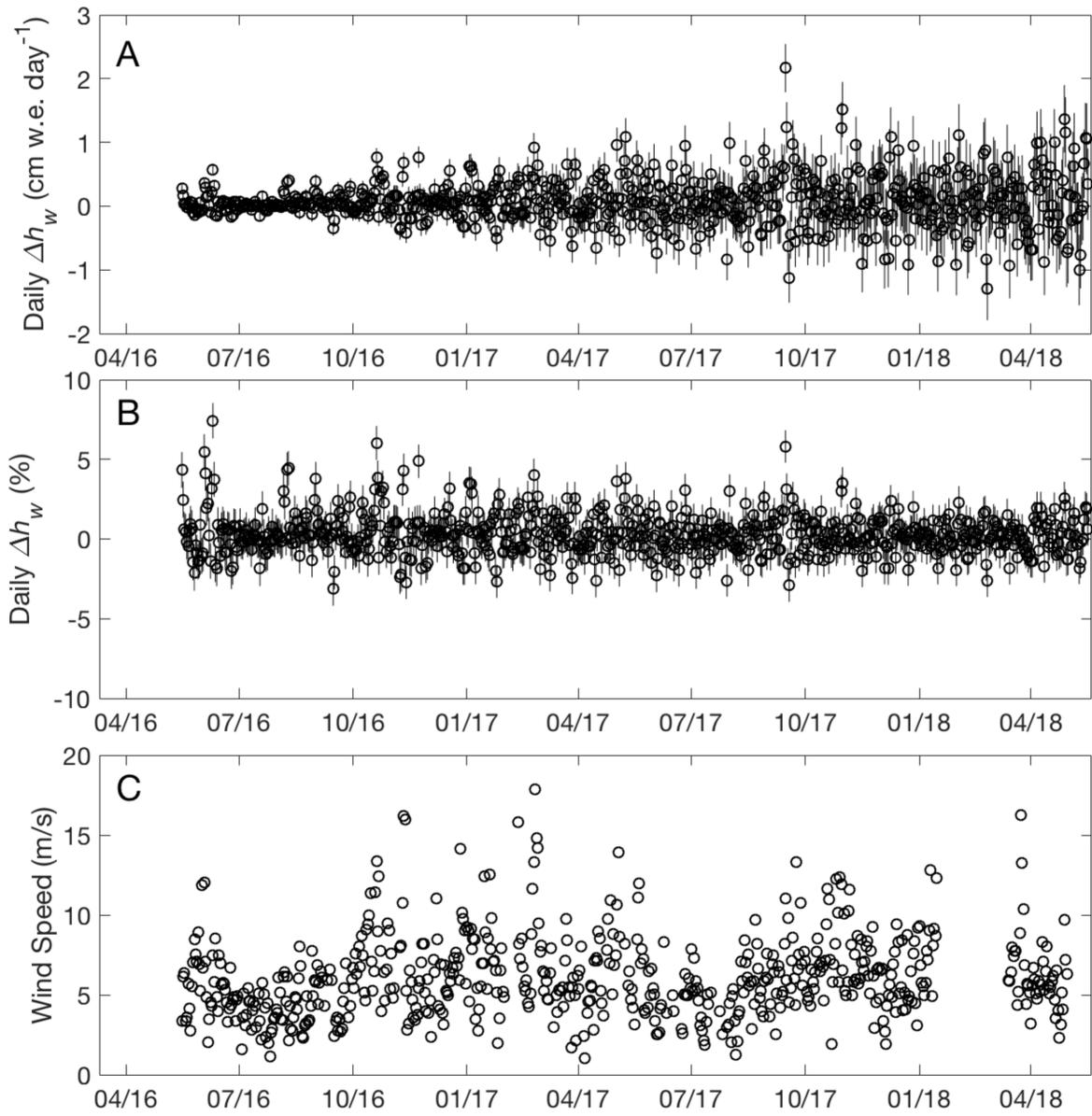
Figure 5: (A) Daily change in water equivalent (w.e.) thickness, h_w, and **(B)** change as a percentage of h_w at Summit Station, Greenland, measured by the cosmic ray neutron sensor (CRNS) with 1σ error bars estimated from the daily standard deviations in hourly measurements versus water equivalent thickness (curve in Fig. 4.) **(C)** Summit Camp daily mean wind speed measured at the automatic weather station.
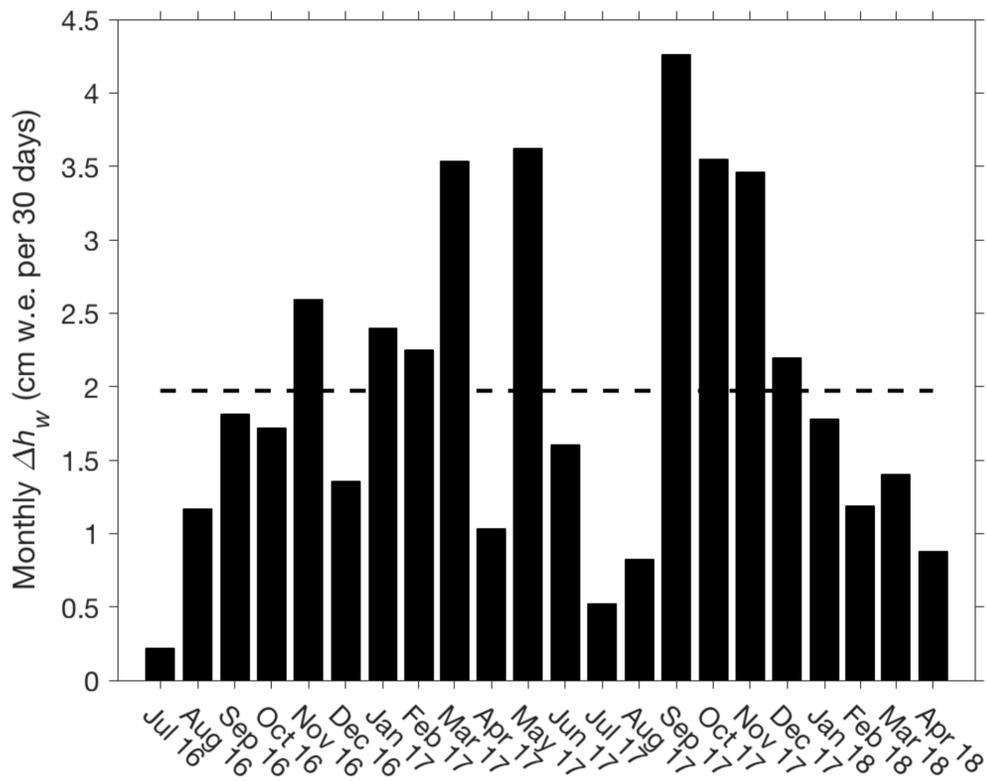
Figure 6: Monthly (30-day) change in water equivalent (w.e.) thickness (h_w) at Summit Camp, Greenland as measured by the cosmic ray neutron sensor (CRNS). Dashes are the average (1.9 cm w.e. per 30 days). The record begins with July 2016 because May is omitted due to infilling of the trench containing the sensor, so that the first difference is from June to July.
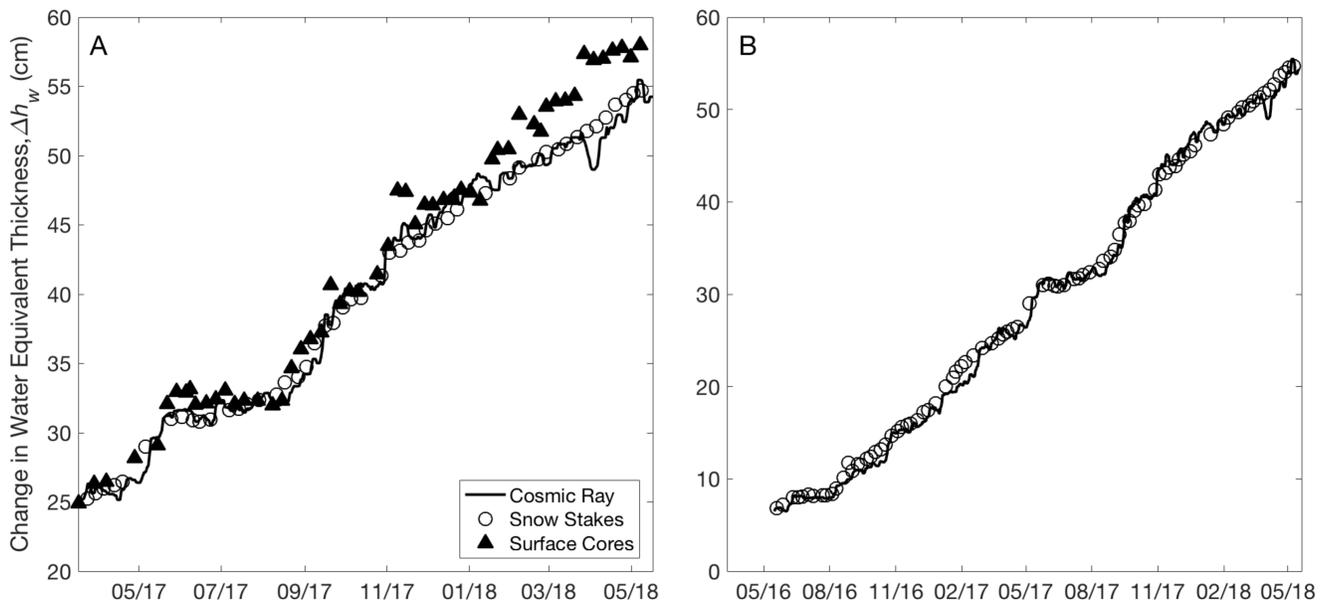
Figure 7: (A) Time series of cumulative change in water equivalent thickness, h_w, at Summit Camp, Greenland, over the period of surface core sampling (17 March 2017 to 8 May 2018) from (solid curve) the cosmic ray neutron sensor (CRNS), (triangles) surface core samples and (circles) the "bamboo forest" snow stake network. Conversion of mean surface height change of the snow stakes to water equivalent thickness is described in Section 4. Values for h_w are the depth of the CRNS, with initial h_w values for the surface core and snow stake set to the concurrent CRNS value.