An assessment of two automated snow water equivalent instruments during the WMO Solid Precipitation Intercomparison Experiment

Craig D. Smith, Anna Kontu, Richard Laffin, and John W. Pomeroy

We would like to thank the referees for taking the time to review this paper submission. Your comments and suggestions are most appreciated and we have tried to address all of them in turn, providing clarification and revisions where necessary. Our responses are highlighted in the text below. Incorporating your suggestions and addressing your comments and concerns have certainly helped to improve this manuscript.

Response to referee comments

Anonymous Referee #1

Although snow water equivalent (SWE) is very important information for not only disaster forecasting but also for earth science, there is still room for discussion of the methods of automatic SWE measurement. They compared different automatic SWE measurement methods (CS725, SSG1000), which are based on the different principles, with the manual SWE measurement at three sites. Then they discussed the characteristics of the comparison results at each site. Although I am very interested in their works and do not doubt that their works give the basic important information for the improvement of automatic SWE measurement methods, their work only shows few scientific new findings in the present version. From this view, I think their topic should be suitable for rather GI (or some publications which mainly treat the topics of instrument, method) publication than TC publication. In order for this manuscript to be accepted for TC, the authors should completely reconstruct the manuscript to clear the new scientific findings and scientific contribution of their works.

Author's response

We will reconstruct the Results section and arrange the text by instrument, not by site, as suggested by Referee #2. We hope that this change will highlight our results including the new findings that came out of the SPICE intercomparison.

Author's changes in manuscript

- Text in **3 Results and Discussion** arranged by instrument and divided into two sections, **3 Results** and **4 Discussion**.

Anonymous Referee #2

The paper compares SWE manual measurements with automated sensors of attenuation of passive gamma radiation attenuation device in two SPICE sites and another site with rather different climatic and snow conditions; and also one scale snow in one of the SPICE sites. The topic is of high interest for a broad community as the sensors that are compared are starting to be very popular in many sites of the world, and it is necessary to discuss about their accuracy, possible sources of uncertainty, etc. Due to the limited length of snow observations, few locations analysed and some problems related with the experimental design it is not possible to give strong evidences on their accuracy and the reason of biases found between manual and automated measurements. However, I think that the paper still has enough interest for many readers, and it launches some hypothesis of interest that may serve as basis for further research.

In my opinion, the structure of the manuscript is not the best to present the results. I recommend to reorganize the presentation of them showing the equivalent figures for each site together, instead of doing subsections of the results for each site (indeed the discussion is presented in the way I suggest). Meanwhile table 1 and 2 can be combined (and added results from Fortress). In this way it would be reduced the final number of figures (the 14 current figures is excessive in my opinion). More important, it would be possible to identify common processes amongst sites and their differences, and the paper would gain consistency (in the current version some figures are made for one site but not for the others..e.g.). Thus it could be presented the validation of CS725 and SSG1000 (where available) in the three sites with a couple of multi-panel lots (one panel per site), and afterwards to show figures that allow explaining the patterns of accuracy/error shown (the figures relating air temperature and difference SWE, the soil moisture...

One concern is how to ensure that snow depth in automated sensors is the same than that were SWE is measured manually in the three presented sites. This could be another source of error not mentioned in the manuscript. In page 11 is mentioned the spatial variability in snow melting that could affect to different snow depth. Are there snow depth sensors installed above the measured areas with automated sensors. If this is the case, we could see how well the depths are similar and if there are differences, some plot could focus on estimated snow density.

Is there any evidence of a relationship between snow depth or amount of SWE with bias with the manual measurements. There are some references that CS725 may be inaccurate under very deep snowpack. - I think that Figure 11 should show the bias between manual and automated measurements to properly observe the coincidence between liquid water content and SWE differences.

It seems that the existence of water around SSG1000 may cause serious disruptions in the functioning of the device. Is it apparently due to problems in their installation or is a problem of design of the device?

Hoping my comments will result useful.

Author's response

The referee points out the limited length of observations, few locations, and problems with experimental design. This was largely related to the choice of limiting analysis to the SPICE instrumentation over the SPICE intercomparison period and using available reference measurements. We did add CS725 data from Fortress Mountain (which is not a SPICE site) to help support our theory that CS725 offsets were due to changes in soil moisture related to sandy soil substrates. To have more data and locations, we will add one more site, Weissfluhjoch, to the SSG1000 intercomparison which adds 2 seasons of instrument intercomparisons. However, as the referee points out in regards to the existing intercomparison, the methods for the manual reference measurements differ from the other sites. This contributes to the difficulty in assessing the level of uncertainty in the reference.

The referee suggests looking at automated snow depths if they are measured in the same field of view as the SWE measurements. This, unfortunately, was not the case as the snow depth sensors were located at different locations participating in their own assessment. This deficiency is noted and we will comment that this could assist future SWE intercomparisons. We did use manual snow depth measurements at Caribou Creek and Sodankyla to estimate site variability and will add a commentary on this.

As suggested, we will rearrange the Results and Discussion section by instrument, not by site. We will also combine relevant figures and tables to reduce their number and to make comparisons between sites easier.

The problems with SSG1000 and water are due to a design feature: the cables are so short that the electronics box must be installed below the instrument on the ground. We know that Fortress Mountain site has also experienced similar problems with the instrument. After SPICE, the Sodankylä team asked the manufacturer to replace the cables with longer ones. A third measurement winter was without instrument failures, as the electronics box was now mounted about 50 cm above ground. This will be briefly mentioned in the text.

We revised Figure 11 (now Figure 6) to include the difference between the CS725 and manual SWE measurements. This now shows the relative timing between the changes in the soil moisture/soil temperature and the changes in the instrument offset. However, the difference in the frequency of measurements means that interpretation isn't necessarily clear. As the manuscript states, it offers an explanation for some of the initial offset which is shown by the first intercomparison in mid-December but the large discrepancies shown in mid-February are not attributed to changes in liquid water in the soil anyways. It does mark the coincidence between spring soil thaw and infiltration of liquid water and a corresponding increase in the sensor bias.

Author's changes in manuscript

- Text in **3 Results and Discussion** arranged by instrument and relevant figures and tables combined. Text divided into two sections, **3 Results** and **4 Discussion**.
- Weissfluhjoch is added to the SSG1000 intercomparison.
- The SSG1000 problems will be addressed in the text.
- Revised Figure 11 (now Figure 6) to include a plot of the difference in SWE between the CS725 and the manual measurements in relation to the soil moisture/soil temperature changes
- We added further commentary in the Discussion section on how the experiment design could be improved for future SWE intercomparisons, including measuring snow depth in the same footprint as the SWE sensors.
- Added a commentary in the Discussion section on the snow depth variability at Caribou Creek and Sodankyla as an indicator of how the SWE could also vary with space and time.

Referee #3 Charles Fierz

General comments

The goal of this paper, as stated by the authors, is, "to assess the use and accuracy of two instruments that were tested during the WMO - Solid Precipitation Intercomparison Experiment the Campbell Scientific CS725 and the Sommer Messtechnik SSG1000 snow scale" as well as, "to inform users of the best way to use these instruments and of any potential measurement issues that may influence their data interpretation." Unfortuately, however, I don't feel I get the promised info by reading that paper. The deficiencies of the CS725 (first 5 lines on page 4) are simply confirmed and no convincing, in depth analysis of possible source of errors are addressed for any of the two instruments. Instead, spatial variability is invoked to explain the mismatch between the continuous measurements and the manual, punctual (in time) reference measurements, the error of which are not quantified either. In view of the above and my comments below, I can hardly recommend to accept that paper for publication.

Indeed, I really doubt that the authors have enough convincing arguments and data to bring the paper in line with their goals, even after major revisions.

Author's response

It is true that we confirm some deficiencies in the CS725, as the importance for soil moisture calibration of the CS725 is previously established (Martin et al. 2008; Wright et al. 2011). However, in previous comparisons the instrument has agreed much better with other measurements (automated or manual) than we have observed during SPICE. Our analysis shows that the CS725 assessed during SPICE did not compare as well with manual measurements as did previous intercomparisons and we propose that the larger offset is due to changes in soil moisture content after calibration and throughout the snow season. This was in turn caused by the soil type at both intercomparison sites which was predominately sand. The possible ways for soil moisture to change during winter are identified by Gray et al. (1985, 2001) and Lilbaek and Pomeroy (2008). We argue that these soil conditions did not exist for previous intercomparisons and instrument users need to be made aware of this. We also imply that this is not necessarily an instrument deficiency and some hydrological users may be able to use this measurement principle to their advantage.

We agree with the need of more in-depth analysis to support the hypotheses, and will provide this in a revised manuscript or make recommendations to guide future SWE sensor intercomparisons to fill these deficiencies.

The specific and minor comments are addressed after each comment. The resulting changes to the manuscript are also addressed after each comment.

Specific comments

•p. 1, lines 25-26: "These manual measurements are considered to be the reference for the intercomparison." This is one of the crux of that paper. The devices are hardly looked at while a whole lot of blame goes on these manual measurements, the error of which are hardly addressed though.

We now address some of the error in the manual measurements in the Discussion section. The literature addresses the mean bias associated with some manual samplers but the variability of the potential error in field sampling is quite large and depends highly on the skill of the user and the condition of the snow pack.

•p. 1, line 30: "throughout the intercomparison periods" is absolutely misleading and false. One full ablation period is missing and the problems of the instruments were not looked at.

One ablation period is missing for one sensor only…the SSG1000 at Sodankyla for the 2014/2015 period. However, we recognize that this limits the sample size of the intercomparison during ablation for this sensor (and for intercomparison with the other sensor). We will change the text "throughout the intercomparison periods" to "when data were available" where applicable. In addition, we will add SSG1000 intercomparison from an additional site

•p. 1, line 33: "seasonal melt" suggest replacing by "ablation period" [throughout the paper]. Furthermore, is pre-melt = accumulation period? I strongly suggest that you define these terms properly once and use them consistently throughout the paper. See for example on p. 6, line 31 for "point of maximum seasonal SWE"

We agree that consistent terminology is important. We will replace 'pre-melt" with "accumulation period" and "seasonal melt" with "ablation period".

•p. 3, line 13: "2 Instrumentation and Methods" Should try to not give interpretation in that paragraph but include it in the results section, for example as "previous intercomparison"

The commentary about previous intercomparisons has been moved to the "Introduction" section and all interpretation has been moved to "Results" or "Discussion".

•p. 3-4; lines 26-5: Is this the correct place for such comments? Should be moved to discussion part as an introduction to it.

Agreed

•p. 4, line 15: "... impact of the move are considered to be negligible." Why? Later you speak of spatial variability influencing the results.

We resolve this by including commentary on the spatial variability at this site and demonstrate how small this is over the distances of the sensor move.

•p. 4, lines 29-30: "... to stabilize the overlying snowpack and prevent ice bridging." Why does the snowpack need to be stabilized? How is ice bridging prevented? What observations do corroborate this?

This is a statement made by the manufacturer (and is stated as such) to justify their design of a larger platform with a smaller platform in the centre that does the actual weight measurement. It is out of our scope to validate their design other than to comment on the quality of the measurements.

•p. 5, line 3: "..., and the only snow scale provided ..." is incorrect. There is another SPICE site (Weissfluhjoch) equipped with a snow scale ... and a snow pillow next to it from the same provider.

The text in the manuscript is correct. The snow scale in Weissfluhjoch was provided by the site, not by manufacturer for testing during SPICE. However, based on this comment as well as some other comments about lack of snow scale intercomparison data (due to instrument failure), we have chosen to add the snow scale intercomparison for Weissfluhjoch..

•p.5, line 8: "... reliable manner ..." but not always. The simple regression does not reveal the true problems!

This statement is in reference to the actual functionality of the instrument (see the context) rather than an assessment of the measurement accuracy. The instrument functioned with a low failure rate.

•p. 5, lines 13-15: "The sensitivity..." Such a sentence belongs to the summary and conclusion section.

We agree. The sentence will be moved.

•p. 5, lines 24-25: "... has a mean measurement error less than 0.5 %." 0.5 % of what? Does this refer to the repeatability of measurements? Overall, the number looks very optimistic and the reference Farnes 1983 is hardly available. From other publications by the same author (1980 and 1982, see Kinar & Pomeroy, 2015b), this figure can hardly be reproduced. I'd strongly suggest to be more precise here.

Farnes et al. (1982) state that ESC-30 overmeasures by -0.3 % (i.e. undermeasures) of the true SWE, and that the correction factor for ESC-30 is 1.00 (no correction required). The accuracy is quoted by Goodison et al. (1987) which was added as a reference. This paper may be more available than the original Farnes et al.

paper. Of course, these errors are in ideal situations, as stated by Kinar & Pomeroy who cite Powell (1987) in reference to errors in measuring more difficult snow packs. We add a discussion on this as it relates to measurements during SPICE (especially at Caribou Creek).

•p. 5, line 25: “were taken just inside the footprint of the CS725” How do these disturbances affect the measurements?

Since the sample is 30 cm^2 inside an 80 m^2 sensor footprint, the impact is negligible but the sample area was filled in with discarded snow when possible. This was clarified in the text.

•p. 7, lines 14-15: “the instrument trends are the same as for the manual measurements” In my view your simple regression analysis fails here and does not look at problematic features. For example, how do you explain the apparent loss of mass around mid February 2014? Similar unexpected wiggles are also seen at other times on both seasons. These spurious measurements are also known to occur at Weissfluhjoch and are not to be expected from a well designed, continuously recording device.

Mid feb 2014, March 2014, March-April 2015: very cold periods (-30 C) after positive air temperatures resulted in ice bridging. The snow supports itself and the weight is not on the load cell. After adding the Weissfluhjoch data, we see it here as well (as the reviewer points out). This issue will be discussed in the revised paper in the Discussion section.

•p. 7, lines 15-21: This comparison or ‘tracker’ does not appear to work very well. Indeed, in January 2014 there is a large increase in the ‘Difference in SWE’ while air temperature plummets! Similar behaviour can be found at other times. In summary, there is another reason behind these large increases, but which?

Agreed, the correlation between temperature increase and the immediate corresponding increase in sensor bias isn’t always clear. Sources of error are now presented in the Discussion section, including the potential issues with manually sampling a snow pack that persists after a melting period as a result of ice layers, etc,

•p. 11, line 11: “systematic sampling errors” Can these be avoided?

I see the reviewer’s point. This was poorly worded. These sampling errors could be systematic but not necessarily so. The text was changed.

•p. 11, lines 24-27: “Although ...” A somehow simplistic view. In the paper you never assess any of the errors you assign the outliers to. This is definitely the biggest weakness of that paper.

The text referenced here has been changed and we now address the errors in more detail, including ice bridging, in the Discussions section.

•p. 11, lines 30-36: Poor conclusions! What does this linear relationship show? Would you calibrate the CS725 with a SSG1000? Were the deficiencies of the CS725 not already known (see your introduction)?

The linear relationship is meant to show that it’s not just the manual measurements causing the bias but rather the measurement principle of the sensor. We have tried to make this clearer in both the Discussion and the Conclusion sections. The previous literature does point out some known deficiencies in the CS725 but the behavior seen at Sodankyla, especially during melt, has never been documented. The comparison with the SSG1000 supports our conclusions that the increased bias is related to infiltration of meltwater into the sandy soils, and serves as a warning to users who may be using the instrument in similar situations.

•p. 12, line 4: "... not all increases in the bias... " Interesting, you don't even mention those in the discussion!

This has been corrected, see note above about sources of error

•p. 12, line 8: "... errors in the manual SWE measurements...." I agree that measurement errors can amount to a certain percentage of SWE. But you don't even quantify these errors, even though you use them as reference. Blaming not quantified errors for the observed mismatch seems simplistic indeed.

It is difficult to estimate the measurement error in the field because it highly depends on the measurement conditions and the observer, however we do attach some loose estimates on this error in the Discussions section on Manual SWE measurements.

•p. 12, line 32: "... have a good agreement ..." Here I have really hard times to follow the logic of your conclusions. First you blame the manual measurements for mismatch and then you claim a good agreement!

The point was that there was a good agreement considering the potential issues with the manual measurements and the measurement principle of the sensor. Hopefully the revised text makes this clearer.

Minor comments

•Please ALWAYS put a space between numbers and units (often wrong)

We will correct these.

•p. 1; line 28: Replace "(w.e.)" by "(mm w.e.)".

We will replace this.

•p. 1; line 29: Replace "Creek respectively" by "Creek, respectively".

We will correct this.

•p. 2; line 7: Are these two units equivalent?

Yes they are, assuming that density of water is 1 kg/m3.

•p. 3; line 1: Replace "(SPICE; Nitu" by "SPICE (Nitu".

Replaced with "(SPICE) (Nitu"

•p. 3; line 28: Replace "snow cores" by "snow courses".

No, they actually compared to 8 snow samples (cores), not 8 snow courses.

•p. 3; lines 28-29: What are "snow pit densities"? Please describe. Sampler?

It means that a density profile of 5 cm high samples is measured, SWE is determined for each sample and the total SWE is calculated as a sum of the layers. This is thought to be more accurate than SWE from bulk snow sample (core). Details will be added to the revised manuscript.

•p. 4; line 30: "0.3 % of full scale" that is 3 mm w.e.! Under what conditions? Moreover, the high resolution seems useless.

This is what the manufacturer tells about the technical details of their instrument. They do not specify the conditions, or give reasons for such high resolution.

•p. 4, line 35: What is drifting? Snow? The electronics?

This was in reference to drifting snow. This has been clarified in the text.

•p. 5; line 23: “snow tube” Is this the correct term? I’d suggest using “snow sampler” -as found elsewhere in the literature –throughout your paper.

Snow sampler can be any kind of sampler, from 5 cm high wedge sampler to a long tube. “Snow tube” defines the type of sampler used, and is a widely used term (e.g. by Kinar and Pomeroy 2015).

•p. 7; line 5: Replace “almost” by “by almost”.

We will replace this.

•p. 7; line 29: Replace “offset” by “intercept”, throughout, as used in the tables.

We will replace this.

•p. 7; line 32: What does “differential” mean?

Different melt rates at different locations at the site.

•p. 8; line 8: Try to read “2013/2014 only due to data unavailability for 2014/2015” and replace that sentence.

Clarified this sentence

p. 8; line 30: “melt and re-freeze occurred” That really questions the term “pre-melt” used elsewhere in the paper.

‘Onset of snow-melt’ and ‘snow-melt season’ are commonly used terms for the final snow melt (ablation) in the spring. They do not rule out the possibility of short melt-refreeze cycles during the accumulation period. However, we will replace ‘pre-melt’ with ‘accumulation period’ in the text for clarity.

•p. 9; lines 32-33: “gravimetric water fraction” vs “volumetric water content ”What is the relation?

Volumetric water content (theta) is the volume of water divided by the total volume. Gravimetric water content or fraction (u) is the mass of water divided by dry soil mass. They are related by theta = u SG, where SG is the soil specific gravity and depends on its density.

Author’s changes in manuscript

Changes to the manuscript are integrated into the specific responses to the referee’s comments above.

An assessment of two automated snow water equivalent instruments during the WMO Solid Precipitation Intercomparison Experiment

Craig D. Smith[1], Anna Kontu[2], Richard Laffin[3], John W. Pomeroy[4]

[1]Environment Canada, Saskatoon, S7N 3H5, Canada

[2]Finnish Meteorological Institute, Sodankylä, FI-99600, Finland

[3]Campbell Scientific, Edmonton, T5L 4X4, Canada

[4]Centre ~~of~~ for Hydrology, University of Saskatchewan, Saskatoon, S7N 5C8, Canada

Correspondence to: Craig D. Smith (craig.smith2@canada.ca)

Abstract

During the World Meteorological Organization (WMO) Solid Precipitation Intercomparison Experiment (SPICE), automated measurements of snow water equivalent (SWE) were made at the Sodankylä (Finland), Weissfluhjoch (Switzerland) and Caribou Creek (Canada) SPICE sites during the ~~northern~~ Northern ~~hemisphere~~ Hemisphere winters of 2013/2014 and 2014/2015. Supplementary intercomparison measurements were made at Fortress Mountain (Kananaskis, Canada) during the 2013/2014 winter.~~ and xx site~~. The objectives of this analysis are to assess automated SWE measurements against a reference, comment on their performance, and make recommendations on how to best use the instruments and interpret ~~its~~ their measurements. Sodankylä, Caribou Creek and Fortress Mountain hosted a Campbell Scientific CS725 passive gamma radiation SWE sensor. Sodankylä and Weissfluhjoch ~~also~~ hosted a Sommer Messtechnik SSG1000 snow scale. The CS725 ~~measurement~~ operating principle is based on measuring the attenuation of soil emitted gamma radiation by

the snowpack and relating the attenuation to SWE. The SSG1000 measures the mass of the overlying snowpack directly by using a weighing platform and load cell. Manual SWE measurements were obtained at the SPICE sites on a bi-weekly basis over the accumulation/~~melt~~ablation periods using bulk density samplers. These manual measurements are considered to be the reference for the intercomparison. Results from Sodankylä and Caribou Creek showed that the CS725 generally overestimates SWE as compared to manual measurements by roughly 30 to 35 % with correlations (r^2) as high as 0.99 for Sodankylä and 0.90 for Caribou Creek. The RMSE varies from 30 to 43 mm water equivalent (mm w.e.) and 18 to 25 mm w.e. at Sodankylä and Caribou Creek, respectively. The correlation at Fortress Mountain was 0.94 (RMSE of 48 mm w.e.) with no systematic overestimation. The SSG1000 snow scale, having a different measurement principle, agreed quite closely with the manual measurements at Sodankylä and Weissfluhjoch throughout the ~~intercomparison~~ periods when data were available (r^2 as high as 0.99 and RMSE from 8 to 24 mm w.e. at Sodankylä and 56 to 59 mm w.e. at Weissfluhjoch). When the SSG1000 ~~is~~ was compared to the CS725 at Sodankylä, the agreement ~~is~~ was linear until the start of ~~seasonal melt~~ablation period when the positive bias in the CS725 increases substantially relative to the SSG1000. Since both Caribou Creek and Sodankylä have sandy soil, water from the snowpack readily infiltrates into the soil during melt but the CS725 does not differentiate this water from the un-melted snow. This issue can be identified, at least during the spring ~~melt~~ablation period, with soil moisture and temperature observations like those measured at Caribou Creek. With a less permeable soil and surface runoff, the increase in the instrument bias during ~~melt~~ ablation period is not as significant, as shown by the Fortress Mountain intercomparison.

1 Introduction

The measurement of snow water equivalent (SWE) is vital for flood and water resource forecasting, drought monitoring, climate trend analysis, and hydrological and climate model initialization (Barnett et al., 2005; Gray et al., 2001; Bartlett et al., 2006; Laukkanen, 2004). Many of these applications require accurate and timely information about how much water is being held within the snowpack (Pomeroy and Gray, 1995). SWE measurements can be made in-situ, either manually or via automated instrumentation, or derived from remote sensing platforms, and are usually expressed as units of mass per area (kg m^{-2}) or in equivalent units of millimetres of water equivalent (mm w.e.).

Manual measurements of SWE are typically made using a multi-point bulk density sampling technique along an established transect or snow course (WMO, 2008). Snow course measurements are often time consuming and expensive, especially if required in remote locations (Pomeroy and Gray, 1995). This means that manual SWE measurements may be infrequent or only undertaken when the snowpack is estimated to be at its seasonal maximum. Prohibitive costs of manual snow course observations have led to the reduction of these measurements by many agencies, including Environment and Climate Change Canada, where operational snow course numbers have decreased from over 100 in the 1980s to less than 30 (Barry, 1995; Brown et al, 2000). Since the early 1990s, manual SWE measurements have been augmented or replaced by remote sensing techniques such as passive microwave retrievals (Goodison and Walker, 1995) but these techniques still require accurate and reliable in-situ measurements for ground-truthing and retrieval development (Derksen et al., 2005; Takala et al., 2011).

With the reduced availability of manual SWE measurements, automated instruments for the measurement of SWE are becoming more necessary and more commonplace. Snow pillows have been used for the automated measurement of SWE in remote locations since the 1960s (Beaumont, 1965) by measuring the overlying pressure of the snowpack on a fluid filled bladder. The SNOTEL network in the United States is based on snow pillow measurements (Serreze et al., 1999). More recently, similar measurements are obtained using snow scales that use a weighing surface and load cell to measure the weight of the overlying snow (Beaumont, 1966; Johnson et al., 2007). Several indirect methods exist to measure SWE that include the use of neutron probes (Harding, 1989) in which a radiation source is placed under the snowpack and the scattering of neutrons through the snow is measured by a detector. Cosmic ray proton probes (Kodama et al., 1979;

Rasmussen et al., 2012) work in a similar manner but do not require an active source. The probes described by Kodama ~~and Nakai~~ et al. are installed under the snow while the system described by Rasmussen et al. (called COSMOS) is installed above the snow. Kinar and Pomeroy (2007; 2015a) outline a method of non-invasive sonic reflectometry through the snowpack to determine snow density, liquid water content, and temperature. Other passive radiation sensors are mounted above the surface and measure the attenuation of naturally emitted radiation from the soil as it passes through the snowpack and then relate~~s~~ this attenuation to SWE content (Choquette et al., 2008; Martin et al., 2008). Each of these instruments and techniques have advantages and disadvantages, which are not discussed here (see Kinar and Pomeroy (2015b) for a more comprehensive description of snow measurement methods and related issues). Rather, this analysis assesses the use and accuracy of two instruments that were tested during the World Meteorological Organization (WMO-Solid Precipitation Intercomparison Experiment (SPICE)~~;~~ (Nitu et al., 2012; Rasmussen et al., 2012), namely the Campbell Scientific CS725 and the Sommer Messtechnik SSG1000 snow scale.

The CS725 (previously known as GMON or GMON3) has been previously field tested by Hydro Québec ~~(as referenced above)~~ (Choquette et al., 2008; Martin et al., 2008) as well as by Wright et al. (2011). Previous results by Choquette et al. (2008) showed an average error of +18 % when comparing to 8 manual snow cores over 3 seasons in Quebec. They got a somewhat better agreement with total SWE calculated from density profiles (with an average error of +5 %) but only had 4 samples over 2 seasons. Wright et al. (2011) showed intercomparison results between GMON3 sensors and snow pillows, precipitation gauges, and snow courses at Sunshine Village (Alberta, Canada) and Tony Grove Ranger Station (Utah, USA). Results showed high correlations between the sensor and (unadjusted) accumulated precipitation (0.99) and between the sensor and snow pillow observations (0.99) but lower correlations (0.83) with snow course observations (during one season at Sunshine Village). The authors question the quality and inherent biases in the snow course samples but do not comment on the sources of error or the proximity of the snow course to the instrument.

Instrument intercomparisons that included the SSG1000 have been limited but some results are reported by Stranden and Grønsten (2014), who showed parallel SWE measurements between snow pillows, snow scales, and manual snow courses. With mitigating circumstances (e.g. snow drifting and scale issues), they concluded that the measurement surface area had an impact on the measurement quality and that the Sommer scale gave “promising results” but that further intercomparison was required.

One of the objectives of the WMO-SPICE project is to assess the performance of automated instrumentation for the measurement of snow, including snow on the ground (SoG). This is accomplished by comparing the tested instruments to an established reference measurement. In total, fifteen countries are participating in the WMOSPICE project with about 20 intercomparison sites. Of these, 7 countries and 9 intercomparison sites are hosting SoG instrumentation. The instrumentation for SPICE has either been provided by the instrument manufacturers or by the site hosts. For SoG, 13 different instruments are under test with 9 measuring snow depth and 4 measuring SWE. The CS725 and the SSG1000 SWE instruments examined here were installed at the Sodankylä (Finland), ~~and~~ Caribou Creek (Canada) and Weissfluhjoch (Switzerland) intercomparison sites (Fig. 1). To supplement the CS725 data collected for WMO-SPICE, data was added from an additional CS725 instrument installed at the Fortress Mountain ski area in the Kananaskis region of the Canadian Rocky Mountains. ~~and site xx.~~

2 Instrumentation and Methods

2.1 Campbell Scientific CS725

The CS725 (Fig. 2 left) is a passive gamma sensor developed by Hydro Québec in collaboration with Campbell Scientific (Canada) Corp. (Choquette et al., 2008; Martin et al., 2008). The instrument is installed above the snow surface and determines SWE by measuring naturally emitted gamma radiation from Potassium and Thallium sources in the soil that ~~are~~ is attenuated by the snowpack. Each gamma ray detected by the sensor element is counted over a user defined period, the resulting distribution is compared to the distribution when there was no snow cover, and the difference is used to calculate SWE. The sensor field of view (FOV) is approximately 60° from centre resulting in a ~~field of view (~~FOV~~)~~ of approximately 80 m^2 when installed 3 m above the snowpack and with the collimator attached. The collimator serves to shield the instrument from gamma rays emitted from sources that are not in the target area. The effective range of the instrument is 0 to 600 mm w.e. with a measurement accuracy of +/- 15 mm w.e. from 0 to 300 mm w.e. and 15 % from 300 to 600 mm w.e. (Campbell Scientific CS725 Manual, https://s.campbellsci.com/documents/ca/manuals/cs725_man.pdf).

The two CS725 instruments for WMO-SPICE were both installed in October 2013 to Sodankylä, Finland, and Caribou Creek, Canada, and operated over the Northern Hemisphere winters of 2013/2014 and 2014/2015. Both instruments were mounted so that the bottom of the instrument was approximately 2 m above the ground. Both instruments were installed with the manufacturer provided collimator. Data was output every 6 hours. ~~Each~~ The instruments performed in a reliable manner exhibiting a measurement rate higher than 95 % at both sites over the course of the two winter seasons. No malfunctions were noted and no maintenance was required. The instrument at Sodankylä was moved approximately 10 m during the summer of 2014 to avoid some buried cables in the footprint, but any potential impact of the move are considered to be negligible and addressed in Section 4.

The third CS725 used in this analysis was not a WMO-SPICE instrument, but was loaned to the University of Saskatchewan for testing and intercomparison by the instrument manufacturer. This instrument was installed in a clearing near the Fortress Mountain ski resort in the Kananaskis Valley, Alberta, Canada. The CS725 was mounted at a height of approximately 3.5 m above the ground. The distance to the trees around the instrument was approximately 10 m from the centre of the instrument FOV, putting them outside of the response area. Data collected by this instrument from October 2013 through June 2014 is used in this analysis. Like the other CS725 instruments, SWE data was output every 6 hours.

2.2 Sommer SSG1000

The SSG1000 snow scale (Fig. 2 right) manufactured by Sommer Messtechnik, Austria, measures SWE through the use of a weighing platform and load cells. Unlike the CS725, it makes a direct measurement of the weight of the snowpack on top of the weighing platform and converts this weight to SWE. The entire platform consists of 7 perforated panels, each 0.8 m x 1.2 m, that are attached to a frame and installed level with the surface of the ground. The entire instrument surface is 2.8 m x 2.4 m (6.72 m^2) but only the centre panel is weighed by the load cell. According to the manufacturer, the purpose of the larger surface surrounding the centre measurement panel is to stabilize the overlying snowpack and prevent ice bridging (http://www.sommer.at/en/products/snow-ice/snow-scales-ssg). The SSG1000, as tested for WMO-SPICE, has a measurement range of 0 to 1000 mm w.e.~~ SWE~~, and a manufacturer stated resolution and accuracy of 0.1 mm w.e. and 0.3 % of full scale (3 mm), respectively.

~~Instrument intercomparisons that included the SSG1000 have been limited but some results are reported by Stranden and Grønsten (2014) who showed parallel SWE measurements between snow pillows, snow scales, and manual snow courses. With mitigating circumstances (e.g. drifting and scale issues), they concluded that the measurement surface area had an impact on the measurement quality and that the Sommer scale gave “promising results” but that further intercomparison was required.~~

The SSG1000 snow scales in this analysis~~, and the only snow scale provided by the manufacturer for WMO-SPICE, was~~were installed in the Sodankylä ~~Intercomparison~~ and Weissfluhjoch SPICE sites~~Field~~. The Weissfluhjoch instrument was provided by the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) Institute for Snow and Avalanche Research (SLF). Data collection from the instrument started in October 2013 and continued for the 2013/2014 and 2014/2015 Northern Hemisphere winters. The SSG1000 was located in the North East quadrant of the Sodankylä SPICE ~~Intercomparison~~ Field, approximately 22 m southeast of the original location of the CS725. At Weissfluhjoch, it is located in the southwest corner of the instrument field. SWE observations from the instruments were recorded once per minute during the two intercomparison seasons. The instruments at both sites worked in a reliable manner during the accumulation periods but the instrument at Sodankylä malfunctioned due to water damage to the electronics late in the spring of 2014 and again early spring of 2015. At Weissfluhjoch, 99% of the 1-minute data for both years are usable for intercomparison while ~~, resulting in~~ 83 % and 67 % of the 1-minute ~~measurements~~data is ~~being~~ useable for the respective intercomparison periods at Sodankylä. Other than this, no malfunctions were reported or maintenance required during the intercomparison. ~~An SSG1000 was also installed at the Fortress Mountain Ski Resort site near the CS725 by the University of Saskatchewan, but failed to operate during the comparison winter due to damaged electronics from a possible lightning strike.~~ ~~The sensitivity of this instrument to atmospheric electrical phenomenon, even when located in a sheltered forest clearing is a concern.~~

The electronics box of the SSG1000 is designed to be installed below the instrument on the ground, which is flooded during snowmelt. After SPICE, the manufacturer was asked to provide longer cables allowing the installment of the electronics box about 0.5 m above the ground. After this modification, there were no problems with water and electronics during the following snowmelt season~~. The sensitivity of this instrument to atmospheric electrical phenomenon, even when located in a sheltered forest clearing is a concern~~.

2.3 Reference SWE measurements

The reference SWE manual measurements for this intercomparison differ by site. ~~are~~ All except Weissfluhjoch were bulk density snow samples made with a snow sampling tube of a known diameter that has one end capable of penetrating and cutting into the snowpack. The tube ~~is~~ was inserted into the snowpack down to the surface of the ground and the sample ~~is~~ was extracted. Along with the sample, the depth of the snowpack ~~is~~ was also obtained. The sampled snow ~~is~~ was then either bagged and weighed or ~~is~~ was weighed inside the tube using a cradle and balance. The snow sampler used in Canada is different than the tube used in Finland and these differences, as well as any other differences in sampling technique, are described below. The SWE measurements at Weissfluhjoch were done via snow pit density samples and depth measurements.

At Caribou Creek, the reference SWE measurements were obtained using an ESC-30 snow tube with a 30 cm^2 cutting area. Farnes et al. (1983) and Goodison et al. (1987) show~~demonstrated~~ that the ESC-30, when used correctly, has a mean measurement error of less than 0.5 % of the true SWE. Bulk density samples at Caribou Creek were taken just inside the ~~footprint~~FOV of the CS725, bagged, and weighed. A 30 cm^2 area is assumed to have a negligible impact on future sensor measurements considering the sensor footprint is 80 m^2, but it was filled in with discarded snow when possible. These manual SWE measurements were made about every two weeks in conjunction with a full 5 point snow course across the Intercomparison Field and into the forest canopy on each side.

At Sodankylä, the reference SWE measurement was made using a Finnish bulk density sampling tube, with a sampling area of 78.54 cm^2, and balance (Kuusisto, 1984) at roughly the same location in the Intercomparison Field every two weeks. Only one sample was measured at a time. During the winter of 2013/2014, the bulk density SWE sample was obtained approximately 12 m from the centre of the CS725 FOV and approximately 16 m from the centre of the SSG1000. In 2014/2015, after the CS725 was moved, the manual sampling was done approximately 6 m from the CS725 FOV and approximately 25 m from the SSG1000.

An ESC-30 snow tube was used at the Fortress Mountain site. A full snow survey was conducted at the site once per month, transitioning to bi-weekly during the ~~melt~~ablation period. Although the actual ~~The actual~~ snow survey course was ~~conducted~~ through the forested area, supplemental measurements were taken in the clearing ~~and not the clearing~~ where the instrumentation is located. The distance between the sensor and the manual measurements was approximately 10 m~~, but three concurrent samples of SWE were obtained in the clearing closer to the CS725~~.

The manual SWE measurements at Weissfluhjoch were made by SLF and derived via bi-weekly snow pit density ~~samples~~profiles obtained in the centre of the instrument field. The distance between the sensor and manual snow measurement varied from observation to observation as the location of the snow pit was relocated for each bi-weekly measurement. The average distance was approximately 20 m.

2.4 Intercomparisons

The intercomparisons are not completely consistent amongst the four~~three~~ sites because of the different instrumentation and manual methods for measuring reference SWE. At Sodankylä and Weissfluhjoch, the ~~CS725 and the SSG1000~~sensors can both be compared with the manual SWE measurements made nearby, although the manual measurements are ~~is~~ not within the FOV of either instrument, as the destructive nature of the manual measurements would have prevented further automated measurements. The timestamps of both instruments were matched as closely as possible to the manual observation time. Since the CS725 only reports every 6 hours, the ~~output~~ measurement output closest to the manual observation time was used for the intercomparison. Since the SSG1000 reports every minute, no time adjustment was necessary. The same procedure was used to compare the CS725 to the SSG1000. No SSG1000 was present at Caribou Creek or ~~operating at~~ Fortress Mountain and no CS725 sensors were installed at Weissfluhjoch.~~, so and the comparisons at these sites were only between the CS725 and the manual SWE measurements.~~

For the CS725, which outputs a SWE value derived from both the Potassium (K) and Thallium (Tl) counts, the manufacturer suggests that the output with the higher count~~s~~ is generally the most reliable. For Sodankylä, the K/Tl ratio is always greater than 1 (varying from 3.5 to 8.0) indicating that the Potassium counts are greater than the Thallium counts. For Caribou Creek, the ratio varies from 2.8 to 4.0. For Fortress Mountain, the ratio varies from 0.3 to 8.5 but is above 1 approximately 70 % of the time. Therefore, the CS725 analysis is based on the Potassium output although the statistics for Thallium are shown in brackets in Tables 1 and 4. This will allow us to determine if there ~~are~~ were any obvious differences in the statistics related to the output derived from one source or the other.

3 Results~~ and Discussion~~

3.1 CS725 vs. m~~M~~anual~~Sodankylä~~

The comparison between the CS725 measurements and the manual SWE observations are shown in Fig. 3 with the Potassium output in red circles and the Thallium output in blue triangles. The black line in the figure represents the 1:1 line. Figure 4 shows the time series of automated and manual SWE measurements. Figure 5 shows the difference between the CS725 (red, difference divided by 2 for visualization) and the measured air temperature (blue) through the two seasons. The regression analysis coefficients and summary statistics are listed in Table 1. The statistics are provided for each individual season and for the two seasons combined. The statistics for the individual seasons are also refined further to show results for the accumulation period (delineated from the ablation period by ~~determined as before~~ the timing~~point~~ of maximum seasonal SWE). This will help to eliminate the effects of snow melt on both the manual measurement and the various potential impacts on the CS725 measurement. These figures and tables are further analyzed for each site in the following subsections.

3.1.1 ~~CS725 vs Manual~~Sodankylä

~~The comparison between the CS725 measurements and the manual SWE observations are shown in Fig. 3 with the Potassium output in red circles and the Thallium output in blue triangles. The black line in the figure represents the 1:1 line.~~ Throughout the intercomparison periods at Sodankylä, the CS725 ~~overestimates~~ overestimated SWE on average by 30 % (mean relative bias or MRB) as compared to the manual measurements. ~~The regression analysis coefficients and summary statistics are listed in Table 1. The statistics are provided for each individual season and for the two seasons combined. The statistics for the individual seasons are also refined further to show results for the periods prior to snow melt (determined as the point of maximum seasonal SWE). This will help to eliminate the effects of snow melt on both the manual measurement and the various potential impacts on the CS725 measurement.~~

From Table 1, the regression analysis for the CS725 as compared to manual SWE over the entire season results in a slope (β) of 1.24 for 2013/2014 and 1.06 for 2014/2015. The difference in β between the K and Tl outputs is small. The ~~offset~~intercepts (ε) for the entire seasons are 8.77 mm w.e. for 2013/2014 increasing to 26.9 mm w.e. for 2014/2015. This difference might be in part a result of moving the instrument to a new location. The correlation coefficient, r^2, is 0.92 for 2013/2014 and 0.96 for 2014/2015. With the period of ~~seasonal melting~~ablation eliminated from the analysis, the impact on β and ε are relatively small although the ~~offset~~intercept ε decreases almost 9 mm w.e. and 4 mm w.e. for the respective seasons. The ~~pre-melt~~accumulation period r^2 increases to 0.97 and 0.99 for the 2013/2014 and 2014/2015 seasons respectively suggesting that more scatter is introduced into the relationship during the ~~melting~~ ablation period. This is expected, as the snow does not melt evenly at the site and a melting snowpack is more difficult to sample with a snow tube. ~~Some of the difference in the relationships from one season to the next is possibly due to the relocation of the sensor but this is difficult to ascertain.~~

Figure 4 (top) shows the time series for the 2013/2014 (left) and 2014/2015 (right) seasons at Sodankylä. ~~Superimposed on this graph are the measurements made by the SSG1000 that will be discussed further in the next section~~. In this figure, the overestimation of the CS725 (red and blue lines) can be seen when compared to manual SWE (black circles). In general, the instrument trends are the same as for the manual measurements with differences between the measurements increasing after the start of the ~~melt~~ablation periods and in January 2014 and December 2014. ~~. Figure 5 shows the difference between the CS725 (red, difference divided by 2 for visualization) and the measured air temperature (blue) through the two seasons~~. Although it appears from ~~Fig~~Fig. ~~.~~ 5 that the difference between the measurements ~~are~~ is simply increasing with time (or SWE amount), we believe that at least part of this increase is a result of melting in the snowpack which occurs during some relatively warm days. In 2013/2014 (Fig. 5, left), a large increase in the difference occurs after the > 0 °C temperatures in mid to late April. In 2014/2015 (Fig. 5, right), there is a moderate increase after some > 0 °C temperatures in March but a much larger jump after the beginning of the ~~melt~~ablation period in April.

3.1.2 Caribou Creek

The comparison of the CS725 instrument and the manual SWE measurements made at Caribou Creek are shown in ~~Fig~~Fig. ~~.~~ 3 and summarized in Table 1. As with Sodankylä, the difference between the two sensor outputs (Potassium vs. Thallium) is negligible. Also like Sodankylä, the CS725 at Caribou Creek consistently overestimates total SWE such that the MRB is~~on average by~~ 35 %. However, the relationships between the instrument and the manual SWE measurements are different than at Sodankylä. At Caribou Creek, the slopes of the regression line, β, are less than 1 for all scenarios in Table 1 with the exception of the accumulation period in 2014/2015. The intercepts, ε, are all larger than seen at Sodankylä, with the accumulation period in 2014/2015 being the exception once again. The r^2 values range from 0.90 for the combined (2013/2014 and 2014/2015) data to 0.55 for the accumulation period in 2014/2015. ~~Correlations are lower at Caribou Creek for several reasons. The spatial and seasonal variability are greater at this site and the sample size is lower. This is especially the case for 2014/2015 where sample size is lower due to a shorter and more variable winter where melt and re-freeze occurred several times over the course of the season. Melting and re-freezing not only results in higher spatial variability, but also makes the manual SWE measurement more difficult and prone to error. Eliminating the accumulation period improved the comparison statistics for 2013/2014 but made the statistics for 2014/2015 much worse due to the reduced sample size.~~

For both the 2013/2014 and 2014/2015 seasons, the time series for Caribou Creek (Fig. 4 middle) shows a rapid increase in SWE in early winter related to heavier, wet snowfall events that most likely began as rain and transitioned to snow. For 2013/2014, the CS725 time series generally follows the trend of the manual SWE measurements with a large deviation developing mid- to late-March with the onset of seasonal melting. Figure 5 (middle) shows the time series of the difference between the CS725 and manual SWE (in red, divided by 2 for visualization purposes) and the temperature time series (blue) for both seasons. In 2013/2014 (Fig. 5 middle left), there is an increase in the difference that occurs in late January. This could be due to a melt period where temperatures at the site exceeded 4 °C preceding the increase in the instrument bias. A much larger jump in the difference occurs mid-March possibly due to significantly higher temperatures (exceeding 10 °C) earlier that month. In 2014/2015 (~~Figure~~Fig. 5 right), the deviation between the measurements occurs earlier in the season (mid- to late-January) coinciding with a January snow melt period characterized by above 0 °C air temperatures and high wind speeds (not shown). Differences decrease after snowfall events in February only to increase again after the start of ablation in March.

3.1.2.1 Soil m~~M~~oisture and soil ~~T~~temperature

In reaction to an observed offset after the 2013/2014 intercomparison season, soil moisture and temperature probes were installed at the Caribou Creek site with the objective of correlating post-calibration, overwinter, and ablation soil moisture changes with sensor offsets.~~After the 2013/2014 winter, a profile of Stevens Hydra Probe soil moisture sensors was installed within the FOV of the CS725 at Caribou Creek. The objective was to monitor changes in soil moisture that might influence the sensor's soil moisture calibration at the autumn/winter transition or changes in soil moisture throughout the winter that might be related to mid-winter snowpack melt, both of which could cause the instrument to overestimate as compared to the manual measurement.~~ The instruments were installed at three depths: 0 to 5 cm (vertically), 5 cm (horizontally) and 20 cm (horizontally). Unfortunately, the probes only measure liquid water fraction by volume (WFV) so the analysis is mostly limited to when the soil temperatures (also measured by the probe) are above 0 °C when we assume that most of the water in the soil is unfrozen.

Formatted: Indent: First line: 1 cm

Figure 6 shows the time series of soil moisture near the surface (0 to 5cm) along with the difference between the CS725 and manual measurements (scaled by a factor of 100 for visualization) for the 2014/2015 season. The red markers indicate when the soil temperature at this level is above 0 °C. It is easy to see from the time series when the liquid soil moisture (near the surface) freezes in late fall resulting in a rapid drop in measured ~~liquid water content~~WFV. Following the freezing of the near surface layer, which occurs 8 November 2014, the measured soil moisture in this layer remains fairly static until mid-March 2015 when a period of positive air temperatures (~~Figure~~Fig. 5 middle right) raises the near surface soil temperatures above freezing, transitioning frozen soil moisture to liquid and allowing for infiltration of snow melt water into the sandy soil at this site.

The freezing of the 0 to 5 cm depths in early November is preceded by rain/snow events in late October that are represented by the large jump in CS725 SWE shown in ~~Figure~~Fig. 4 (middle right) and confirmed with snow depth measurements (not shown). During the transition from rain to snow and prior to the surface freezing, ~~Figure~~Fig. 6 shows fluctuations in near surface soil moisture related to the precipitation events in late October and early November. Considering that the soil moisture calibration of the sensor was entered as 0.10 (gravimetric water content of GWC~~fraction~~), the increase in measured ~~volume~~WFV ~~tric water content~~ to 0.18 prior to freezing has the potential to create a perpetuating offset in the CS725 SWE estimates and may explain at least some of the bias shown by the instrument beginning in mid-December~~. Initial estimates provided by Campbell Scientific (personal communication, 2015) suggest that a 10 % increase in gravimetric soil moisture would approximate a 10 mm w.e. response in the SWE estimate of the CS725. Using these approximation would explain roughly 30 % of the CS725 offset over the 2014/2015 period. More work is needed to establish some confidence in these early season SWE measurements that are made before the ground freezes and the soil moisture content becomes less mobile.~~

In addition to the offset in the CS725 SWE measurements that occurs at the beginning of the season, it was anticipated that the rapid increase in the difference between the CS725 and Manual SWE at the end of January 2015 could also be attributed to a change in near surface soil moisture, as this was a time of mid-season snow melt. However, a change in the liquid soil moisture during the melt period could not be detected so it is unlikely that the increase in the instrument offset can be attributed to infiltration of melt water into the sandy soil. A more plausible explanation are manual measurement errors that could result from attempting to sample a complex snow pack containing ice layers in the pack or at the snow/soil interface. Ice layers would have formed due to mid-season melt and re-freezing. The increase in the difference between the manual measurement and CS725 in mid- to late-March could be a result of snow melt infiltrating into the top layers of the sandy soil as the soil thaws or forming a basal ice layer (Woo et al., 1982) on top of the soil. A corresponding spike in measured soil moisture during early spring snow melt is shown in ~~Figure~~Fig. 6.

3.1.3 Fortress Mountain

The intercomparison of the CS725 instrument and the manual SWE measurements made at Fortress Mountain are shown in ~~Figure~~Fig. 3 (bottom) and summarized in Table 1. Unlike the other two sites, the CS725 and manual SWE measurements generally fall on the 1:1 line with no systematic overestimation (MRB < -5 %). This can also be seen in the time series shown in Fig. 4 (bottom). The slope of the regression line is 0.88~~1~~ with a small decrease to 0.764 when excluding the ablation period. The intercept is 32.4 mm w.e. increasing to 84.4 mm w.e. when excluding the ablation period. The r^2 is comparable to Sodankylä at 0.92 (increasing to 0.94 by excluding the ablation period). It is unfortunate that the sample size is relatively small (n=8) but regardless, the instrument compares quite well to the manual measurements at this site. ~~Unlike Caribou Creek and Sodankylä that both have very sandy soils, there doesn't appear to be an increase in the instrument bias as compared to the manual measurements following a mid-winter or end of season melt event. In fact, from Fig. 5 bottom, we see the bias drop into high negative numbers after a melt period at the end of March 2014 and again at the end of April 2015. The absence of a trend of increasing bias following melt is most likely due to low infiltration of meltwater into these saturated frozen soils. The water here may be more inclined to drain away from the mildly sloping target area rather than infiltrating.~~

3.2 SSG1000 vs. Manual

~~The time series of the SSG1000 measured SWE, also shown in Fig. 4, suggests a close relationship with the manually measured SWE.~~ The regression analysis for the SSG1000 intercomparisons are shown in Fig~~.~~ure 7 with the time series for both seasons shown in Fig~~.~~ure 8. The ~~The~~ comparison statistics are in Table 2. This analysis, as for the CS725 above, is organized by site.

3.2.1 Sodankylä

The SSG1000 regression analysis with the manual SWE measurements ~~is~~ ~~for the regression analysis sh~~shown in ~~Figure~~Fig. ~~.~~ 7 (top) and summarized in Table 2 has ~~6 indicate that the~~an r^2 for the entire 2014/2015 period of ~~is quite good at~~ 0.99 but is only 0.84 for 2013/2014 period. However, the SWE data from the SSG1000 is not available for the ~~seasonal snow melt~~ ablation period in 2014/2015 due to an instrument malfunction. To have a consistent intercomparison for the two seasons, the ~~melt~~ablation period (post maximum SWE) was removed from the 2013/2014 period and the r^2 becomes 0.97, very similar to 2014/2015. Combining the two seasons, the slope of the regression, β, becomes 0.99 with an offset ε of -7.27 mm w.e. with an r^2 of 0.88. The ~~average bias~~MRB for the two seasons combined is -11 %.

The time series of these data are shown in Fig~~.~~ure 8 (top) for both the 2013/2014 (left) and 2014/2015 (right) seasons. For both seasons, the sensor measurements track quite well with the manual measurements. The outliers that appear in Fig~~.~~ure 7 (top) can also be seen in the 2013/2014 time series (Fig~~.~~ure 8 top left) beginning mid-way through the ablation period. It is unknown if this occurs during the 2014/2015 ablation period due to missing data.

3.2.2 Weissfluhjoch

The regression analysis for the SSG1000 and the manual SWE measurements is shown in Fig~~.~~ure 7 (bottom) with the time series in Fig~~.~~ure 8 (bottom). This alpine site has a much deeper snow pack than ~~both~~either Caribou Creek ~~and~~or Sodankylä but comparable to Fortress Mountain, which unfortunately did not have concurrent SSG1000 measurements. The r^2 for both seasons is quite high at 0.97, similar to the accumulation period intercomparison at Sodankylä, but β is less (0.72 and 0.82) and ε is much higher (91.7 and 79.0 mm w.e.) for both seasons (2013/2014 and 2014/2015). The outliers are obvious in ~~ure~~Fig. ~~.~~ 8 (bottom) when the manual SWE measurements are substantially higher than the sensor measurements. Unlike Sodankylä, these outliers mostly occur before maximum seasonal SWE, which is why we don't break the season down as we do with Sodankylä. They are, however, likely a result of sensor bridging which is discussed more in Section 4. There are outliers that occur late in the ablation periods where the sensor substantially overestimates SWE. These are perhaps due to issues with the manual sampling of a complex (melting or melting/refreezing) snow pack. When combining the two seasons, the resulting low MRB of 8% (for combined seasons) is somewhat surprising given the obvious outliers.

~~The outliers in Fig. 6 can be attributed to the melt~~ablation ~~period in late April to May 2014 but it is difficult to ascertain if the errors are related to the instrument or to the manual measurement. The most likely explanation is the differential melt of the snowpack at the site combined with errors associated with manually sampling a melting snowpack.~~

3.3 CS725 vs. SSG1000

The intercomparison with manual measurements for both the CS725 and the SSG1000 are suggesting that the agreements are the most favourable ~~prior to snow melt~~in the accumulation period. Figure 7 shows the relationship between the CS725 and the SSG1000 for both seasons at Sodankylä with the 2014/2015 season shown in red circles and the 2013/2014 season shown in blue circles (changing to blue triangles at the approximate onset of ~~seasonal melt~~ablation). The relationship for both years appears to be linear up to the time where maximum SWE is reached. At the onset of ~~melt~~ablation, the relationship between the instruments ~~(2013/2014 only due to data unavailability for 2014/2015)~~, shown by the magenta circles, deviates substantially from linear. This is confirmed by Table 3 which shows a higher r^2 when the 2013/2014 ~~melt~~ablation period is not included in the analysis. This analysis could only be completed for the 2013/2014 season since the sensor data is missing for the 2014/2015 ablation period due to malfunction.

~~Some of the difference between the two instruments during melt~~ablation ~~can certainly be attributed to differential melt in the intercomparison field. We know that both sensors agree less well with manual SWE measurements after the onset of melt~~ablation~~. Using hourly web camera photos from Sodankylä (not shown), we can qualitatively confirm that some areas of the intercomparison field melt faster than others due to exposure. However, it is still difficult to ascertain how much of the difference between the instrument~~

~~measurements during melt~~ablation ~~is due to the unequal melting rates and how much can be attributed to the difference in measurement principle of the instruments. While meltwater drains away from the SSG1000 and is no longer measurable as SWE, meltwater that infiltrates into the top layer of the soil could still be interpreted as SWE by the CS725. This issue is discussed further in subsequent sections.~~

4 Discussion

The regression analysis between the CS725 and the manual SWE measurements resulted in r^2 values ranging from 0.55 to 0.99, depending on site and season. Although generally lower than the correlations of 0.99 reported for intercomparisons with other instruments by Wright et al. (2011), our correlations are (on average) similar to the r^2 of 0.83 that they reported for snow tube measurements. The average bias shown here, which ~~averages~~was between 30% and 35%, is substantially higher than the 18% reported by Choquette et al. (2008). The exception to this is the CS725 at Fortress Mountain which has a mean negative bias less than 5% when compared to the manual measurements. Besides the maximum SWE observed at Fortress, the two major differences that Fortress Mountain has from Caribou Creek and Sodankylä are the soil and the topography. Soils at the Fortress Mountain site have higher clay and loam content, overlain with a layer of organics, and generally remain frozen and saturated for the duration of the winter. This, combined with the sloping terrain and faster meltwater runoff via drainage channels, likely minimizes the change in soil moisture during the transition seasons and thereby minimizes potential offsets in the CS725 measurements. Furthermore, the correlations for the CS725 for Caribou Creek are substantially lower than for Sodankylä and Fortress Mountain. ~~Correlations are lower at Caribou Creek~~This could be for several reasons. The spatial and seasonal variability are quite high~~greater~~ at Caribou Creek ~~this site~~ and the sample size is low~~er~~. This is especially the case for 2014/2015 where sample size is small~~lower~~ due to a shorter and more variable winter where melt and re-freeze occurred several times over the course of the season (Fig. 5 middle right). Melting and re-freezing generally ~~not only results in higher spatial variability, but also~~ makes the manual SWE measurements more difficult and prone to error, creates basal ice, and results in higher spatial variability. Eliminating the ablation~~ccumulation~~ period improved the comparison statistics for 2013/2014 but made the statistics for 2014/2015 much worse due to the reduced sample size. Potential sources of error in the CS725 intercomparison are discussed further in the following sections.

The SSG1000 was quite highly correlated with the manual SWE measurements at both Sodankylä and Weissfluhjoch with r^2 values as high as 0.99 at Sodankylä (when excluding the ablation period) and 0.97 at Weissfluhjoch. However, when the ablation period is included in the intercomparison for 2013/2014 at Sodankylä (it is not present in 2014/2015 at Sodankylä due to sensor malfunction), the r^2 drops to 0.84. The more significant result at Sodankylä is the smaller MRB, which is -2% to -15% (depending on the exclusion of ablation), much lower than the bias reported for the CS725. The magnitude of the MRB is similar at Weissfluhjoch except that the bias here is a positive 8%. This is surprising considering the many occurrences of negative sensor bias (as seen in Fig~~.~~ure 8 bottom) but these negative outliers are balanced by some large (albeit inconspicuous) positive outliers at the end of the ablation periods. The outliers for Sodankylä in Fig~~.~~ure 7 (top) can be attributed to the ablation period in late April to May 2014 but it is difficult to ascertain if the errors are related to the instrument or to the manual measurement. The most likely explanation is that these are related to the occurrence of bridging. Bridging is also suspected as the cause of the pre-ablation outliers at Weissfluhjoch since the sensor seems to agree quite well with the manual measurements up to mid-March and early-April for both seasons. An intercomparison with a collocated snow pillow (not shown here) suggests a similar albeit smaller negative bias during the same period. Errors associated with bridging are discussed further ~~Section 4~~in this section.

The CS725 and SSG1000 measurements at Sodankylä correlate very well with each other showing correlations as high a 0.99 when excluding the ablation periods. The key result here, as shown in Fig. 9, is the deviation from this linear correlation at the onset of melt in the 2013/2014 season. Although some of this deviation can be blamed on differential melting at the site, we attribute a large portion of the deviation to the different measurement principles of the sensors. At the onset of melt and the ripening of the snow pack, meltwater drains out of the snow pack towards the ground surface. Once reaching the surface, the meltwater can pool and re-freeze (potentially forming a basal layer of ice), runoff from the measurement area, or infiltrate into the soil. Due to the flat measurement area and the sandy soil at Sodankylä, runoff is unlikely; therefore the meltwater is either infiltrating into the sandy soil or re-freezing at the surface. Either way, the same meltwater is likely draining through and away from the measurement plate of the SSG1000 and therefore no longer being measured as SWE in the snow-pack. However, this meltwater, whether infiltrated into the top layer of the sandy soil or pooling at the surface, is still being registered by the CS725 as SWE. This contributes to the overestimation of SWE by the CS725 as compared to the SSG1000 and to the non-linearity of the intercomparison shown in Fig. 9 after ablation. Also, this meltwater is either difficult or impossible to include in a snow tube sample, increasing the bias between the CS725 and the manual measurements.

4.1 Sources of error

There are several possible sources of error that affect both the automated and manual SWE measurements. They are discussed and analyzed for each instrument/method in this section.

4.1.1 Soil moisture (CS725)

A large potential source of error for the CS725 can arise from a poor pre-snowpack soil moisture calibration or a large post-calibration change in soil moisture prior to the freezing of the ground surface. Overwinter soil moisture changes (Gray et al., 1985), infiltration of snowmelt water into soils (Gray et al., 2001) or formation of a basal ice layer between the snowpack and the soil (Lilbaek and Pomeroy, 2008) could also result in deviation between the manual and CS725 SWE measurements. Since the CS725 calculation of SWE is based on gamma ray counts during wet and dry periods with no snow cover, incorrect measurements or faulty assumptions with respect to the soil moisture calibrations could result in a sensor offset. Furthermore, if soil moisture levels in the soil change significantly prior to freeze-up or during winter or ablation period, then the SWE estimates derived from the sensor are less reliable. The approximate error associated with an inaccurate

Formatted: First line indent, Indent: First line: 0 cm gravimetric soil moisture calibration, as provided by the manufacturer, is roughly 1 mm w.e. of SWE for 0.1 GWC soil moisture error. Figure 6 shows an increase in soil moisture at Caribou Creek up to a water fraction of 0.18 prior to freeze up in the fall of 2014. Given that the gravimetric water content soil moisture calibration was approximately 0.10 (and assuming that the GWC and WFV are roughly the same), the resulting calibration offset could explain up to 40 % of the early season jump in SWE shown in Fig. 4 (middle right) and much of the offset between the sensor and the manual measurement shown by the first intercomparison point in Fig. 5 (middle right). This calibration issue would then perpetuate through the winter period. It is unfortunate that this same soil moisture and soil temperature data is not available for Sodankylä or for the first season at Caribou Creek as this would have provided some verification for the calibration offset.

From Fig. 6, there appears to be a coinciding jump in the CS725 bias and the jump in soil moisture (due to above freezing soil temperatures and infiltration) in the spring of 2015 at Caribou Creek. Although the bias is not as large as that seen in mid-winter, it is a significant increase of approximately 10 mm w.e. for each of the final two intercomparison points in mid-March and early-April. Much of this 20 mm w.e. increase could be explained by a corresponding increase in soil moisture from 0.18 WFV (estimated at freeze up) up to 0.45 WFV (spike at thaw), assuming that the CS725 is interpreting this near surface soil moisture as SWE.

Although there is some ambiguity in the results, we think that these soil measurements are useful for explaining at least some of the offsets seen between the sensor and the manual SWE measurements, especially during the transition periods. More work is needed on these linkages before a reliable sensor adjustment can be derived.

~~3.3.1~~4.1.2 Ice bridging (SSG1000)

Ice bridging is a known issue affecting SWE measurements that are made by weight, such as snow pillows or the snow scale (e.g. Engeset et al., 2000). Bridging typically occurs when air temperature reaches 0 °C and then cools creating a melt-refreeze crust layer on the snow surface. This layer is very hard and supports the weight of the snow, thus decreasing the weight on the pillow or scale. Probable bridging situations can be seen in

~~Figure~~Fig. 8 both at~~in~~ Sodankylä and at~~in~~ Weissfluhjoch. At~~In~~ Sodankylä, in December 2013, March 2014 and February-March 2015 the SWE values measured by the SSG1000 decrease and are lower than the manual measurements. At the same time, air temperature first goes above 0 °C and then cools to as low as~~even~~ -30~~ 0~~ °C creating perfect conditions for ice bridging. At Weissfluhjoch this is not so obvious, but it is difficult to explain the differences between manual and SSG1000 measurements otherwise. The snowpack was homogeneous (verified with terrestrial laser scans) and even though a co-located snow pillow (not shown here) showed some underestimation compared to the manual measurements, the underestimation was much smaller than at the SSG1000. However, snow pillows have been found to be less prone to ice bridging issues d~~D~~ue to their larger surface area~~, snow pillows have been found to be less prone to ice bridging issues~~ (Beaumont, 1966; Tollan, 1970).

4.1.3 Snow s~~S~~patial ~~V~~variability

Another potential source of error in this analysis is due to the spatial variability at the intercomparison sites impacting the relative SWE between the sensor and manual measurement locations. At Sodankylä, the maximum distance between the sensors and the manual SWE measurements was 12 m for the CS725 (6 m after the move prior to the 2014/2015 season) and 25 m (16 m in 2013/2014) for the SSG1000. Unfortunately, only one SWE measurement is made at the intercomparison site, but generally the spatial variability is low with snow depth exhibiting a coefficient of variation (COV) under 6 % (with a maximum snow depth of just over 80 cm). Therefore, the impact of spatial variability in SWE, even with a 25 m separation, is likely quite small for most of the season. However, both webcam photos and snow depth measurements provide evidence that snow melt rates during ablation vary across the site, largely dependent on exposure. Manual snow depth measurements suggest that spatial differences in the area around the SWE measurements are small and are perhaps as high as 4 cm in mid-April of 2014 and less in mid-April of 2015. These differences obviously account for very little of the late season SWE deviation shown in Fig. 5 (top). This also suggests that the CS725 move prior to the 2014/2015 season had a low impact on sensor bias from one season to the next.

At Caribou Creek, with maximum snow depths of 56 cm and 41 cm for the two consecutive seasons, exhibits a much higher spatial variability. Here, COV is about 15 % (19 %) at peak snow depth but increases to 30 % (90 %) during ablation for 2013/2014 (2014/2015). With a full 5 point snow course ~~is~~ performed here, mean SWE maximum is approximately 125 mm w.e. in 2013/2014 and 75 mm w.e. in 2014/2015 with COV very similar to those shown for snow depth. The manual measurement used in the intercomparison is made just inside the footprint of the sensor, approximately 5 m from the centre. Although relatively close, the higher spatial variability could result in a spatial bias, especially during ablation. For example, in 2013/2014, we estimate SWE to increase across the sensor FOV by approximately 10 mm w.e. in late ~~a~~April due to differential melting as a result of exposure. With the manual measurement closer to the lower SWE in the sensor FOV, up to 25 % of the difference in SWE between the sensor and the manual measurement (as shown in Fig. 5 middle left) could be explained.

The spatial variability is not assessed for Fortress Mountain or Weissfluhjoch.

4.1.4 Experiment design

Some aspects of the design of the SWE intercomparison are less than ideal and often were a result of compromise amongst the overall SPICE objectives, site host resources, and nationally accepted practices. These compromises potentially contribute to some ambiguity of the study results and this commentary could form the basis for recommendations on the design of future SWE intercomparisons.

Ideally, the manual reference at each site should have been identical using the same sampling equipment at a prescribed offset from each SWE sensor. Rather, each site host used their nationally accepted method of sampling SWE (as described in Section 2.3). Distances between the manual SWE measurement and the sensor varied from 5 m to 25 m, depending on site, but perhaps more significantly, the variation within the sensor FOV (especially for the CS725) was not properly assessed. This could certainly have been a factor at Caribou Creek but the intense sampling within the FOV of the sensor would have caused too much disturbance. Also, increased frequency (i.e. weekly) of manual measurements is desirable, albeit at the risk of disturbance, especially after significant changes in the snow pack. Manual measurements should pay special attention to the existence of basal ice layers which may have an impact on the overall accuracy of the SWE estimate.

Another ideal situation would have been the co-location of both SWE sensor types at each site. This, in combination with soil moisture and temperature sensors within the FOV of the CS725 sensors at all sites and for both seasons, would have provided additional information for the assessment of sensor bias. Another good addition would be the automated and high frequency measurement of snow depth within the sensor footprints to provide an indicator of snow density and melt rates and perhaps and indicator of snow bridging on the SWE weighing type sensors.

4.1.5 Manual SWE ~~M~~measurements

As noted above, the manual SWE measurements differed by site, the exception being Caribou Creek and Fortress Mountain that both used the ESC-30 snow tube and bagged and weighed the sample. However, we won't comment further on possible bias associated with different samplers (Farnes et al.,1983; Goodison et al. 1987), as these are generally small as compared to the differences in the measurements shown in these results. We do, however, want to address possible errors associated with the manual measurement of a complex snow pack (i.e. a snow pack with ice layers or during melt), especially with a snow tube.

During the intercomparison, both Caribou Creek and Sodankylä experienced several freeze and thaw cycles over the course of the winter (as seen in Fig. 5 top and middle) but one was especially pronounced at Caribou Creek during mid- to late-January 2015 (Fig. 5 middle right). The result of freeze/thaw is usually a "crusty" snow pack with several ice layers. In general, these characteristics make a snow pack difficult to sample with a snow tube as the tube cutters need to cut through multiple ice layers without snow escaping from the bottom of the tube (Powell, 1987). It is anticipated that even an expert user will have difficulties obtaining an accurate sample in these conditions, exacerbated even more by the shallow pack found at Caribou Creek in 2014/2015. It is difficult, even during the course of the sample, to estimate measurement error, but it could easily result in a 5 to 10 % underestimate of SWE. Although this may explain some of the bias in the

CS725 measurements, especially at Caribou Creek, it is countered by the relatively good agreement between the manual and SSG1000 measurements for Sodankylä. However, mid-winter melting could also result in basal ice as the meltwater percolates through the snow and refreezes at the surface (providing that the surface is below 0 °C) or in the top layer of the sandy substrate. Not only would this ice layer be difficult to measure with a snow tube (which is difficult to cut through and often results in an underestimate), the meltwater may drain off of the SSG1000 measurement surface and be underestimated by that measurement as well. This may partially explain the often (but sometimes inconsistent) increase in sensor bias shown by manual SWE measurements following mid-winter freeze/thaw cycles in Fig. 5 (top and middle) Unfortunately, the field notes did not indicate when a basal ice layer was observed so much of this is speculation.

During ablation, measures were taken to sample the snow pack before it ripened but this could not always be accomplished due to travel time to the site (Caribou Creek). Because the sample was bagged and weighed rather than weighed in the tube, a wet sample would experience some errors because of bagging and result in a small underestimate of SWE (perhaps 5 % as a rough estimate).

~~The sensitivity of this instrument to atmospheric electrical phenomenon, even when located in a sheltered forest clearing is a concern.~~

~~4~~5 Summary and Conclusions

Two automated SWE instruments were tested at three~~two~~ WMO-SPICE sites (Sodankylä, Weissfluhjoch and Caribou Creek) and at one additional Canadian site (Fortress Mountain) during the WMO-SPICE intercomparison (Northern Hemisphere) winters of 2013/2014 and 2014/2015. Instrument measurements were compared to periodic manual measurements of SWE at the sites and cross referenced with ancillary measurements of air temperature and soil moisture and soil temperature (at Caribou Creek) to try to determine causality for some of the bias seen in the intercomparison. The objective is not necessarily to determine which instrument makes the most accurate measurement, but to inform users ~~of the best way to use these instruments and~~ of any potential measurement issues that may influence their data interpretation.

Intercomparison results for the CS725 show that it overestimates SWE on average by 30 % and 35 % at Sodankylä and Caribou Creek respectively with higher correlations at Sodankylä (r^2 ranging from 0.92 to 0.99) than Caribou Creek (r^2 ranging from 0.55 to 0.90). The difference in correlations between the sites can be attributed to smaller sample size, higher spatial variability of SWE, and ice layers in the snowpack at Caribou Creek. Offsets were generally higher at Caribou Creek which could be indicative of an inaccurate soil moisture calibration of the instrument, a change in soil moisture relative to the calibration prior to or after the soil freezing, or ~~systematic~~ sampling errors in the manual SWE measurement due to a more complex snowpack. Correlations at Fortress Mountain are also quite high (r^2=0.94) with a mean negative bias of less than 5 %, which is more comparable to the results of Wright et al. (2011) in similar conditions. ~~At Sodankylä, the bias tends to increase with increasing SWE. This does not occur at either Caribou Creek or Fortress Mountain and the reason for this is not entirely clear.~~ At the two sandy SPICE sites, ~~T~~the agreement between the CS725 and the manual SWE measurements are generally better ~~at the two sandy SPICE sites~~ prior to the start of seasonal ~~melt~~ablation. We believe this occurs largely because of early spring melt percolating through the snow pack and either forming a basal ice layer or infiltrating into the sandy substrate. Either way, this water is difficult to measure with a snow tube. However, because this water continues to attenuate the gamma radiation signal detected by the CS725, the sensor still interprets this water as SWE and therefor appears to overestimate as compared to the manual measurements. Seasonal ~~melt~~ablation has no significant impact on the agreement at Fortress Mountain due to saturated frozen soils that restrict infiltration and a mild slope that promoted runoff of meltwater from the site.

The SSG1000~~, although only installed~~ at both Sodankylä and Weissfluhjoch, compared quite well to the manual SWE measurements showing~~ a~~ mean biases of ~~negative bias~~ less than -11 % and 8% at the respective sites. It did, however, experience some technical issues at Sodankylä early in the 2014/2015 snowmelt period which limited the intercomparison for that season. The correlations were quite high with the r^2 ranging from 0.84 to 0.99 at Sodankylä and 0.97 at Weissfluhjoch. Many of ~~T~~the outliers in the SSG1000 intercomparisons are most likely due~~-~~to bridging of the snowpack on the weighing plate. At Weissfluhjoch, these events occurred prior to maximum seasonal SWE while at Sodankylä they occurred~~ing at Sodankylä~~ during ablation.~~occurred in 2013/2014 and were largely related to the melt period late in the season (which was not measured by the instrument in 2014/2015)~~. Removing the ablation~~melt~~ period in the 2013/2014 Sodankylä data resulted in a substantial increase in r^2 from 0.84 to 0.97. ~~Although ice bridging of the scale cannot be ruled out, it is difficult to attribute these outliers to the instrument. These outlier are more likely due to increased spatial variability in site SWE during melt~~ablation ~~or as a result of errors associated with manually sampling the melting snowpack.~~

The SSG1000 correlated~~agreed~~ very well with the CS725 at Sodankylä, especially during the ~~pre-melt~~accumulation period. Although the overestimation of SWE by the CS725 is quite apparent when compared against the SSG1000, the ~~pre-melt~~accumulation period r^2 was 0.98 and 0.99 for the two respective seasons. Intercomparison ~~The scatter plot (Figure. 7) clearly shows a very linear relationship between the two instruments (and the increasing bias in the CS725 with increasing SWE) but also very~~ of the two sensors clearly shows how the ~~the behavior of the~~ CS725 overestimates SWE at the onset of ~~melt~~ablation in March/April of the 2013/2014 season. Independent of the manual measurements, this indicates that the deviation of the CS725 from manual SWE during late season ~~melt~~ablation is most likely instrument related and a result of infiltration of melt water into the sandy soils and the misinterpretation of this water as SWE.~~ at this site~~. ~~While this meltwater drains away from the SSG1000 platform, the water is still available in the soil to attenuate the gamma radiation signal and therefore the CS725 still interprets this water as snow.~~

~~To examine some potential causality for the bias between the CS725 and manual SWE measurements, the bias was cross referenced with air temperature to qualitatively examine the impact of mid- and late-season snowmelt. At both Sodankylä and Caribou Creek, the bias seems to increase substantially after the occurrence of above freezing air temperatures over the course of the winter, although not all increases in the bias can be~~

~~attributed to this. As suggested above, the bias seems to increase 5 substantially again at the onset of seasonal melt~~ablation ~~in March and April. The mechanism for this could be two-fold. The first could simply be the creation of ice layers in the snowpack as a result of freeze/thaw cycles. This would not have an impact on the CS725 measurements but increases the likelihood of errors in the manual SWE measurements. The second could be formation of basal ice layers or infiltration of meltwater into the frozen sandy soil. Installed soil moisture/temperature instruments at Caribou Creek do not show any change related to meltwater infiltration during mid-winter melt, perhaps because either the sandy soil is frozen (in which case the sensors may not register a change) or the meltwater is pooling right at the surface creating a basal ice layer. This basal ice layer would be difficult to accurately sample with a snow tube and therefore result in an underestimation in the manual SWE observation.~~

~~The soil moisture/temperature data was used to qualitatively assess the impact of soil moisture change on CS725 measurements at the beginning of a winter season when precipitation generally transitions from rain to rain/snow to snow. In theory, a change in soil moisture leading up to the soil freezing (freezing locks the moisture in place for the season) could impact the sensor's ability to assess that first snowfall event, and potentially perpetuate an offset through the season. Even though the CS725 at Caribou Creek seems to correctly time the SWE accumulations and melt during the fall transition in 2014, measured soil moisture did fluctuate during these events as the soil was not frozen. However, it is difficult to ascertain if this caused the SWE overestimation for these events or more significantly, if the change in soil moisture at transition resulted in the offset for the remainder of the season. It is unfortunate that soil moisture data is not available following the calibration and transition in 2013. During seasonal melting~~ ablation ~~in March 2015 at Caribou Creek, we do see the large increase in soil moisture related to the thawing of the surface and infiltration of meltwater. This infiltration is most likely contributing to the increase in bias of the CS725 sensor relative to the manual SWE measurements made in mid- and late-March 2015.~~

When comparing SWE instruments to a manual reference, there are several considerations that must be made that ultimately impact the interpretation of the results. We know that the manual measurements of SWE are not free of error. Experience proves that making a snow tube bulk density sample in a snowpack containing ice layers or during melt is difficult and inherently prone to errors. We also have to consider the spatial variability of the snow that we are sampling as the CS725 (and the SSG1000 to a lesser degree) have a much larger measurement footprint tha~~t~~n the manual point sample. Taking this and the technical capabilities of the instruments into consideration, both sensors have a relatively good agreement with the manual reference measurements. We have identified that the SSG1000 has had some technical issues during snowmelt but are satisfied that these issues can be overcome with some installation modifications. The SSG1000 may also underestimate SWE on occasion due to bridging so users need to be aware of this potential. We have~~, however,~~ identified the potential for the CS725 measurements to be misinterpreted, especially when deployed over sandy soils and during melting conditions. ~~when basal ice layer formation or infiltration to soils can occur. Because of the CS725 measurement principle, it does not differentiate between changes in soil moisture (before freeze up in the fall and after thaw in the spring) or basal ice layers from changes in the SWE of the snowpack. Under these conditions, the CS725 will overestimate the actual SWE.~~ From a hydrological perspective, perhaps it is more useful to include this sub-surface moisture in the SWE estimate as it ultimately impacts the amount of water available for runoff. Nevertheless, it is certainly helpful to ~~colocate~~collocate this instrument with ancillary measurements of soil moisture, soil ~~and~~ temperature, and snow depth to guide the user in interpreting the data set. ~~The SSG1000 had some technical issues during snowmelt, but they can be overcome. The instrument may underestimate actual SWE in case of ice bridging.~~

Acknowledgements and Disclaimers

WSL Institute for Snow and Avalanche Research SLF kindly provided the SSG1000 and manual SWE measurements from Weissfluhjoch. Many of the results presented in this work were obtained as part of the

Solid Precipitation Intercomparison Experiment (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

Barry, R.G.: Observing systems and data sets related to the cryosphere in Canada. A contribution to planning for the Global Climate Observing System, Atmosphere-Ocean, 33, 771–807, 1995.

Barnett, T.P., Adam, J.C., Lettenmaier, D.P.: Potential impacts of a warming climate on water availability in snow dominated regions, Nature, 438(7066), 303–309, 2005.

Bartlett, P.A., MacKay, M.D., Verseghy, D.L.: Modified snow algorithms in the Canadian Land Surface Scheme: Model runs and sensitivity analysis at three boreal forest stands, Atmosphere-Ocean, 44, 207–222, 2006.

Beaumont, R.T.: Mt. Hood pressure pillow snow gage, J. Appl. Meteor., 4, 626–631, 1965.

Beaumont, R.: Snow accumulation, Proc. 34th Western Snow Conference, 3–6, 1966.

Beaumont, R. T.: Evaluation of the Mt. Hood pressure pillow snow gage and application to forecasting avalanche hazard, IASH Publications no. 69, pp. 341-349, 1966.

Brown, R.D., Walker, A.E., Goodison, B.: Seasonal snow cover monitoring in Canada: an assessment of Canadian contributions for global climate monitoring, 57th Eastern Snow Conference Proc. of the 2000 Annual Meeting, Hardy, J. (ed.), Syracuse, NY, USA, 131–141, 2000.

Choquette Y., Lavigne, P., Nadeau, M., Ducharme, P., Martin, J.P., Houdayer, A., Rogoza, J.: GMON, a new sensor for snow water equivalent via gamma monitoring, Proc. of the 2008 International Snow Science Workshop, Whistler, B.C., 2008.

Derksen, C., Walker, A., Goodison, B.: Evaluation of passive microwave snow water equivalent retrievals across the boreal forest/tundra transition of western Canada, Remote Sensing of Environment, 96(3/4), 315-327, 2005.

Engeset, R., Sorteberg, H., Udnaes, H.: Snow pillows: Use and verification, Snow engineering 2000: Recent advances and developments – Proceedings of the fourth international conference on snow engineering, Trondheim, Norway, 2000.

Farnes, P.F., Goodison, B.E., Peterson, N.R., Richards, R.P.: Metrication of manual snow sampling equipment, Final report Western Snow Conference, Spokane, Washington 106 pp., 1983.

Goodison, B.E ,Glynn, J.E. ,Harvey, K.D.,Slater, J.E.: Snow surveying in Canada: A perspective, Canadian Water Resources Journal, 12(2), 27-42, DOI: 10.4296/cwrj1202027, 1987.

Goodison, B.E., Walker, A.E.: Canadian development and use of snow cover information from passive microwave satellite data, Passive Microwave Remote Sensing of Land-Atmosphere Interactions, 245–262, Utrecht, The Netherlands, 1995.

Gray, D.M., Granger, R.J., Dyck, G.E.: Overwinter soil moisture changes, Transactions of the ASAE, 28(2), 443–447, 1985.

Gray D.M., Toth, B., Zhao, L., Pomeroy, J.W., Granger, R.J.: Estimating areal snowmelt infiltration into frozen soils, Hydrological Processes 15(16), 3095–3111, 2001.

Harding, R. J., Entrasser, N., Escher-Vetter, H., Jenkins, A., Kaser, G., Kuhn, M., Morris, E. M., Tanzer, G.: Energy and mass balance studies in the firn area of the Hintereisferner, Glacier fluctuations and climatic change, 325-341, Springer, The Netherlands, 1989.

Johnson, J. B., Gelvin, A., Schaefer, G.: An engineering design study of electronic snow water equivalent sensor performance, Proc. 75th Western Snow Conf., 23–30, 2007.

Kinar N.J., Pomeroy, J.W.: Determining snow water equivalent by acoustic sounding, Hydrological Processes, 21(19), 2623–2640, 2007.

Kinar, N.J., Pomeroy, J. W.: SAS2: the system for acoustic sensing of snow, Hydrological Processes, 29(18), 2015a.

Kinar, N.J., Pomeroy, J.W.: Measurements of the physical properties of the snowpack, Reviews of Geophysics, 53(2), 2015b.

Kodama, M., Nakai, K., Kawasaki, S., Wada, M.: An application of cosmic-ray neutron measurements to the determination of the snow-water equivalent, Journal of Hydrology, 41, 85-92, 1979.

Formatted: English (U.S.)

Kuusisto, E.: Snow accumulation and snowmelt in Finland, Publications of the Water Research Institute No. 55, National Board of Waters, Helsinki, Finland, 149 pp, 1984.

Laukkanen A.: Short term inflow forecasting in the Nordic power market, Master thesis, Physics and Mathematics, Helsinki University of Technology, Helsinki, 60 pp., 2004.

Lilbaek, G., Pomeroy, J.W.: Ion enrichment of snowmelt runoff water caused by basal ice formation, Hydrological Processes, 22, 2758-2766, 2008.

Martin, J.P., Houdayer, A., Lebel,C., Choquette, Y., Lavigne, P., Ducharme, P. : An unattended gamma monitor for the determination of snow water equivalent (SWE) using the natural ground gamma radiation, 2008 IEEE Nuclear Science Symposium Conference Record, 983-988, 2008.

Nitu, R., Rasmussen, R., Baker, B., Lanzinger, E., Joe, P., Yang, D., Smith, C., Roulet, Y., Goodison, B., Liang, H., Sabatini, F., Kochendorfer, J., Wolff, M., Hendrikx, J., Vuerich, E., Lanza, L., Aulamo, O., and V. Vuglinsky: WMO intercomparison of instruments and methods for the measurement of solid precipitation and snow on the ground: organization of the experiment, WMO Technical Conference on meteorological and environmental instruments and methods of observations, pages 16–18, Brussels, Belgium, 2012.

Pomeroy, J.W., Gray, D.M.: Snowcover accumulation, relocation, and management, National Hydrology Research Institute Science Report No. 7, National Water Research Institute, Saskatoon, Canada, 1995.

Powell, D.: Observations on consistency and reliability of field data in snow survey measurements, paper presented at 55th Annual Meeting Western Snow Conference, Vancouver, B. C., 1987, 69‑77, 1987.

Rasmussen, R., Baker, B., Kochendorfer, J., Meyers, T., Landolt, S., Fischer A.P., Black, J., Thériault, 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, Bull. Amer. Meteor. Soc., 93, 811–829, 2012.

Serreze, M.C., Clark, M.P., Armstrong, R.L., McGinnis, D.A. and Pulwarty, R.S.: Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data, Water Resources Research, 35(7), 2145-2160, 1999.

Stranden, H.B., Grønsten, H.A.: Filefjell Snow Research Station, Evaluation of recorded data for snow seasons 2009/2010 and 2010/2011 (English Summary), Norwegian water resources- and energy directorate, report no. 23/2011, 27 pp., 2011.

Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Kärnä, J-P., Koskinen, J.: Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements, Remote Sensing of Environment, 115, 3517–3529, 2011.

Tollan, A.: Experiences with snow pillows in Norway, Bulletin of International Association of Scientific Hydrology, 15:2, pp. 113-120, 1970.

Woo, M., Heron, R., Marsh, P.: Basal ice in high Arctic snowpacks, Arctic and Alpine Research, 14(3), 251-260,1982.

World Meteorological Organization: Guide to hydrological practices, Volume 1: Hydrology—From measurements to hydrological information, WMO-No. 168, 6th Edition, Geneva, 296 pp, 2008.

Wright, M., Kavanaugh, K., Labine, C.: Performance analysis of the GMON3 snow water equivalency sensor, Proc. of The Western Snow Conference, Stateline, NV, 2011.

Formatted: Font: Not Italic

Tables

Table 1: Regression coefficients and other statistical measures for the multi-season intercomparison of the CS725 with manual SWE at Sodankylä, Caribou Creek and Fortress Mountain (where β and ε are the slope and intercept of the regression line). ~~Pre-Melt~~Accumulation indicates that data occurring after maximum seasonal SWE is omitted from the analysis.

Site	**Season**	**β K(TI)**	**ε K(TI) mm w.e.**	**r^2 K(TI)**	**RMSE K(TI) mm w.e.**	**Mean Relative Bias (K)**	**n**
Sodankylä	2013/2014	1.24(1.27)	8.77(3.17)	0.92(0.92)	43.0(42.2)	30.1%	17
	2013/2014 (~~pre-melt~~accumulation)	1.24(1.28)	0.0123(-6.63)	0.97(0.97)	35.6(33.9)	24.6%	13
	2014/2015	1.06(1.13)	26.9(24.2)	0.96(0.96)	36.6(42.2)	30.9%	13
	2014/2015 (~~pre-melt~~accumulation)	1.05(1.12)	23.3(20.2)	0.99(0.99)	30.0(35.7)	28.1%	10
	Combined	1.16(1.21)	16.8(11.9)	0.92(0.92)	40.3(42.2)	30.4%	30
Caribou Creek	2013/2014	0.783(0.764)	40.6(46.9)	0.78(0.72)	22.8(27.5)	22.2%	12
	2013/2014 (accumulation)	0.982(0.997)	17.7(20.2)	0.79(0.75)	18.0(22.2)	15.4%	9
	2014/2015	0.849(0.849)	27.1(30.4)	0.77(0.71)	23.6(27.4)	63.0%	7
	2014/2015 (accumulation)	1.12(1.31)	-8.38(-14.5)	0.55(0.60)	25.4(29.5)	42.4%	4

	Combined	0.904(0.911)	27.5(31.0)	0.90(0.87)	23.1(27.4)	34.6%	19
Fortress	2013/2014	0.881	32.4	0.92	48.0	-4.5%	8
Mountain	2013/2014 (accumulation)	0.764	84.4	0.94	56.0	-3.6%	5

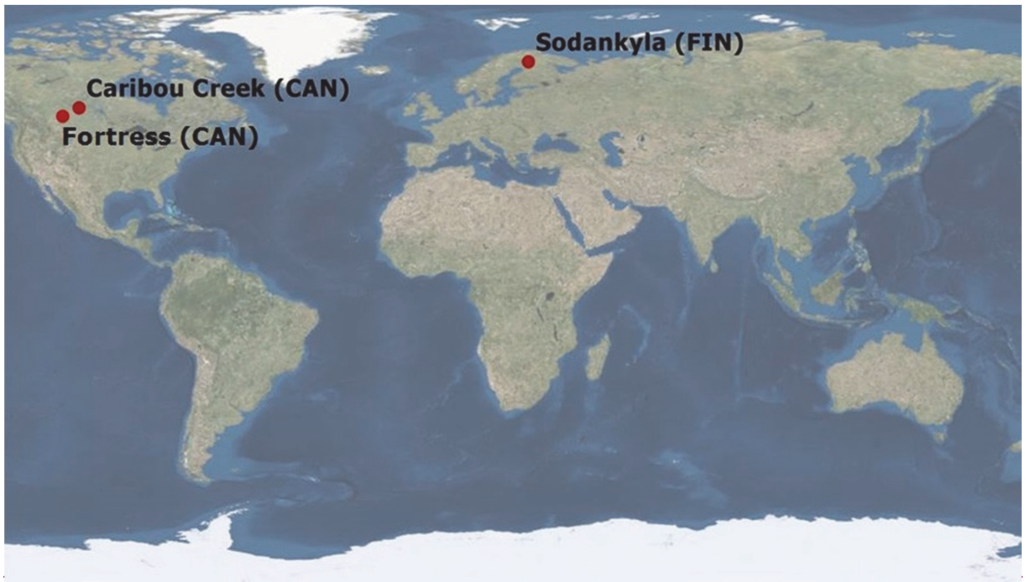
Table 2: Regression coefficients and other statistical measures for the multi-season intercomparison of the SSG1000 with manual SWE at Sodankylä and Weissfluhjoch (where β and ε are the slope and intercept of the regression line).

Site	Season	β ~~K(TI)~~	ε ~~K(TI)~~ mm w.e.	r^2 ~~K(TI)~~	RMSE ~~K(TI)~~ mm w.e.	Mean Relative Bias ~~(K)~~	n
Sodankylä	2013/2014	1.05	-15.5	0.84	24.2	-15.1%	17
	2014/2015	0.92	5.~~4~~5	0.99	7.9~~88~~	-2.3%	10
	Combined	0.99	-7.3~~27~~	0.88	19.8	-10.8%	27
Weissfluhjoch	2013/2014	0.72	91.7	0.97	55.5	4.2	14
	2014/2015	0.82	79.0	0.97	58.6	11.3	17
	Combined	0.79	77.2	0.96	57.2	8.1	31

Table 3: Regression coefficients and coefficient of determination for the multi-season intercomparison of the CS725 with the SSG1000 SWE measurements at Sodankylä (where β and ε are the slope and intercept of the regression line). ~~Pre-Melt~~Accumulation indicates that data occurring after maximum seasonal SWE is omitted from the analysis.

Season	β	ε mm w.e.	r^2
2013/2014	1.20	15.7	0.90
2013/2014 (accumulation~~pre-melt~~)	1.24	4.29	0.98
2014/2015	1.19	11.9	0.99

Figures



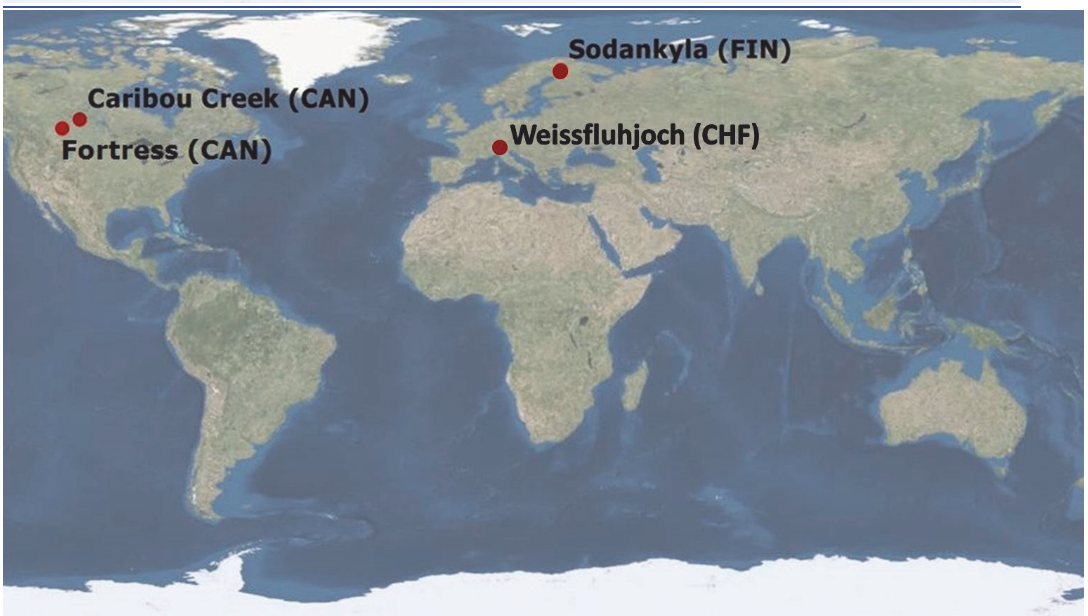


Figure 1: Location of the CS725 (Sodankylä, Caribou Creek, Fortress Mountain) and SSG1000 (Sodankylä and Weissfluhjoch) instrument intercomparisons~~stallations~~.

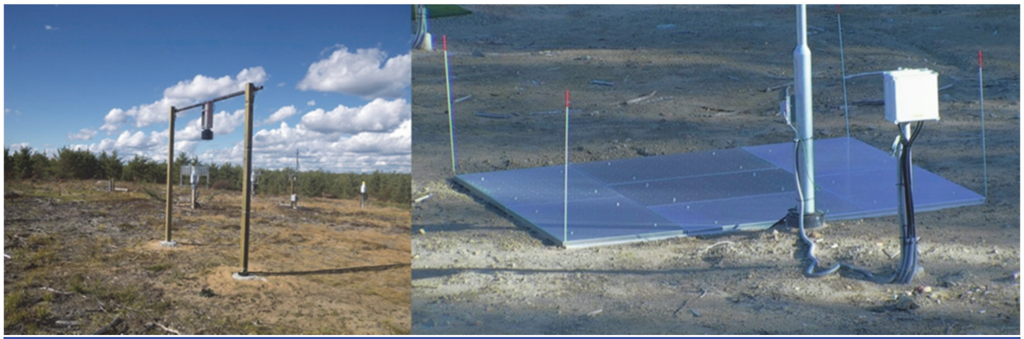


Figure 2: The Campbell Scientific CS725 (left) installed at Caribou Creek and the Sommer Messtechnik SSG1000 (right) installed at Sodankylä.

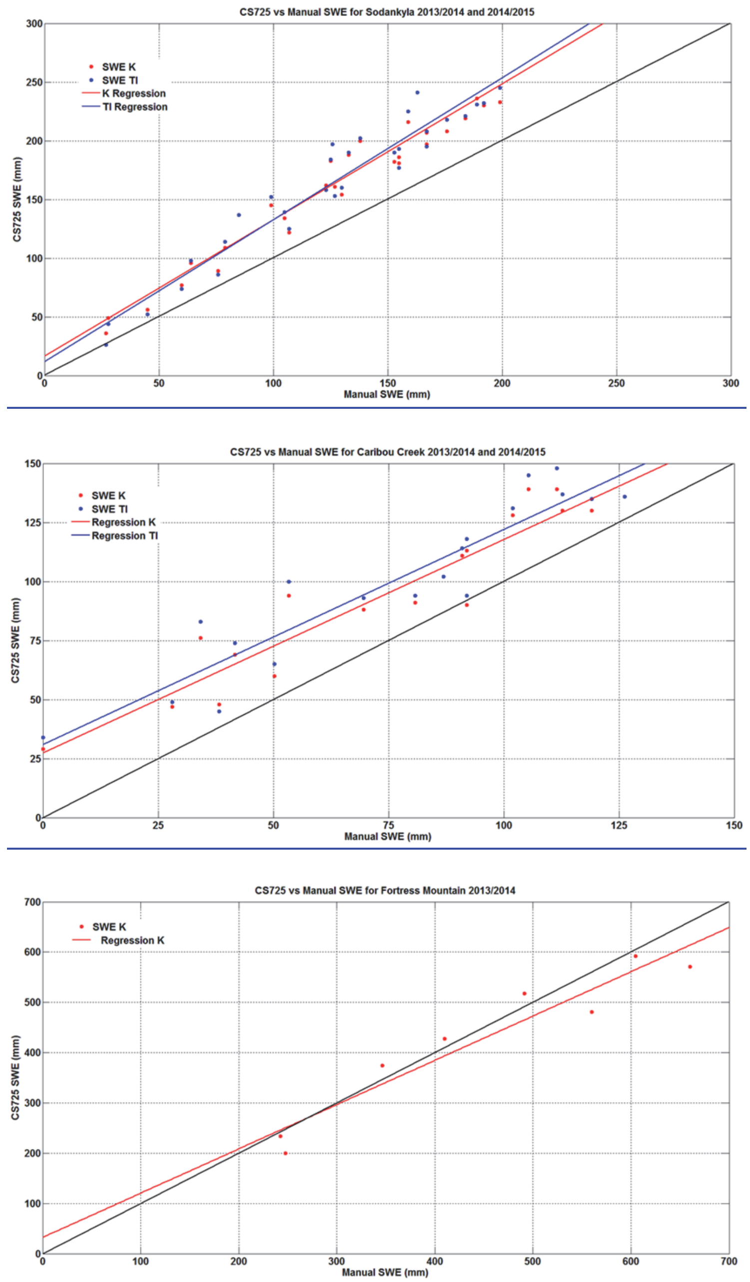


Figure 3: CS725 vs Manual SWE for Sodankylä (top) and Caribou Creek (middle) for the 2013/2014 and 2014/2015 seasons and Fortress Mountain (bottom) for the 2013/2014 season. Potassium output in red and Thallium output in blue. Black line is 1:1.

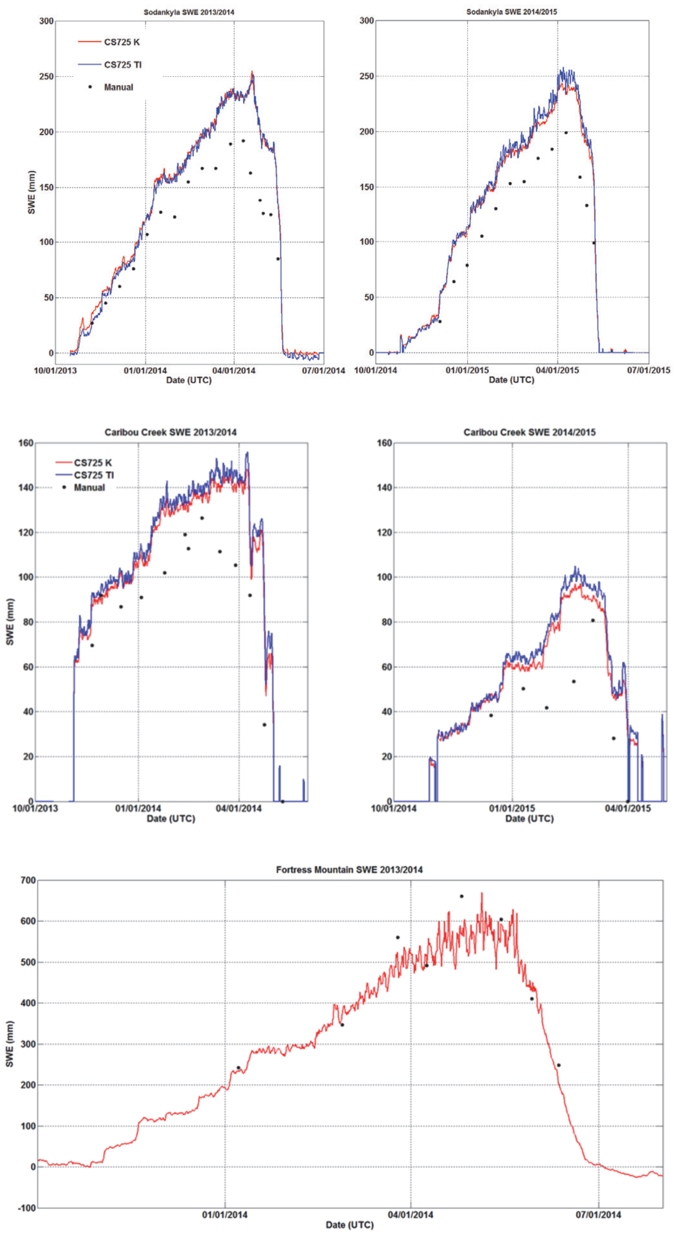


Figure 4: Time series of the CS725 SWE sensors and manual SWE measurements at Caribou Creek (top), Sodankylä (middle) for the 2013/2014 (left) and 2014/2015 (right) seasons, and Fortress Mountain (bottom) for the 2013/2014 ~~(left) and 2014/2015 (right)~~ season~~s~~.

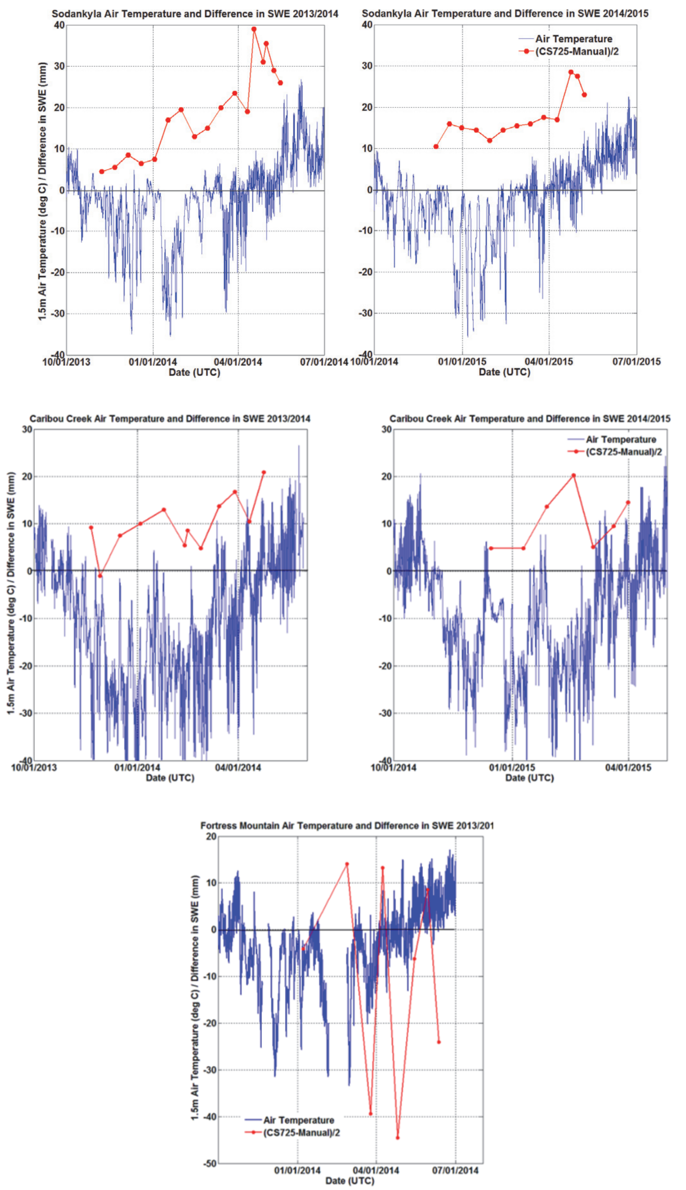
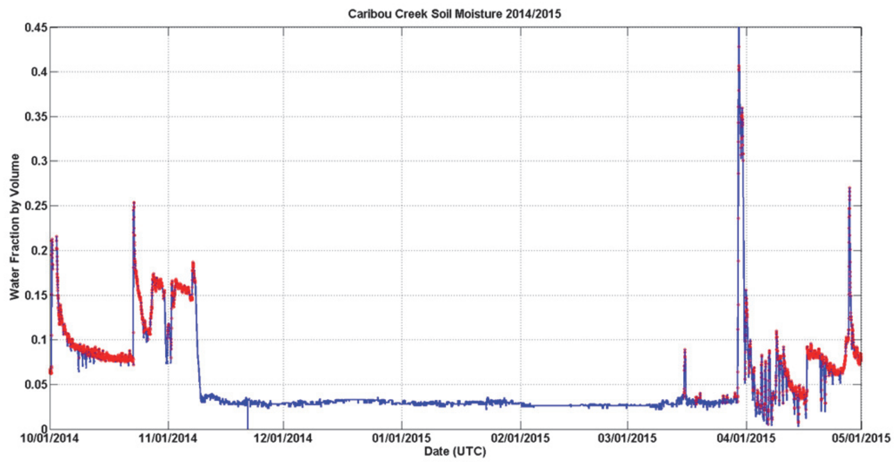


Figure 5: Time series of air temperature (blue) and difference between CS725 and manual measurements (divided by 2, red) at Sodankylä (top) and Caribou Creek (middle) for the 2013/2014 (left) and 2014/2015 (right) seasons, and at Fortress Mountain (bottom) for the 2013/2014 season.



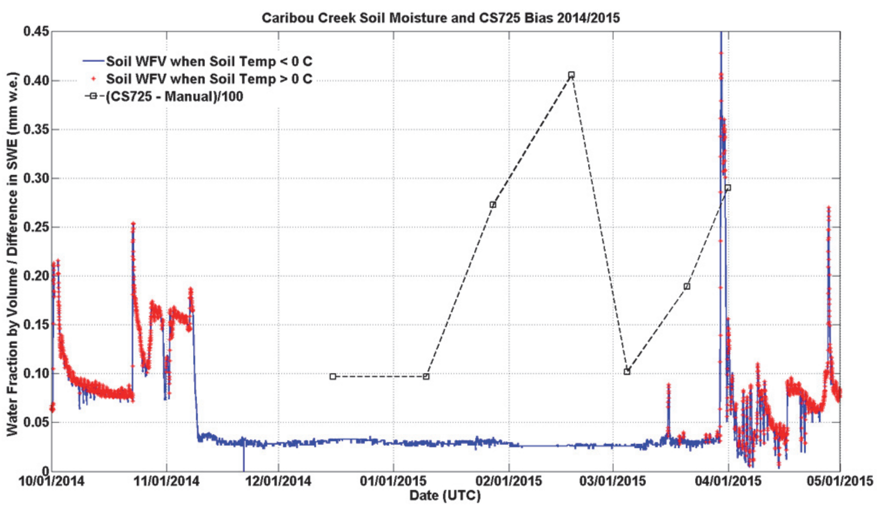


Figure 6: Time series of near surface (0 to 5 cm) soil moisture (water fraction by volume) and the difference between the CS725 and manual measurements (dashed line and black boxes) at Caribou Creek for the 2014/2015 season. Red markers show where near surface soil temperatures are~~is~~ above 0°C.

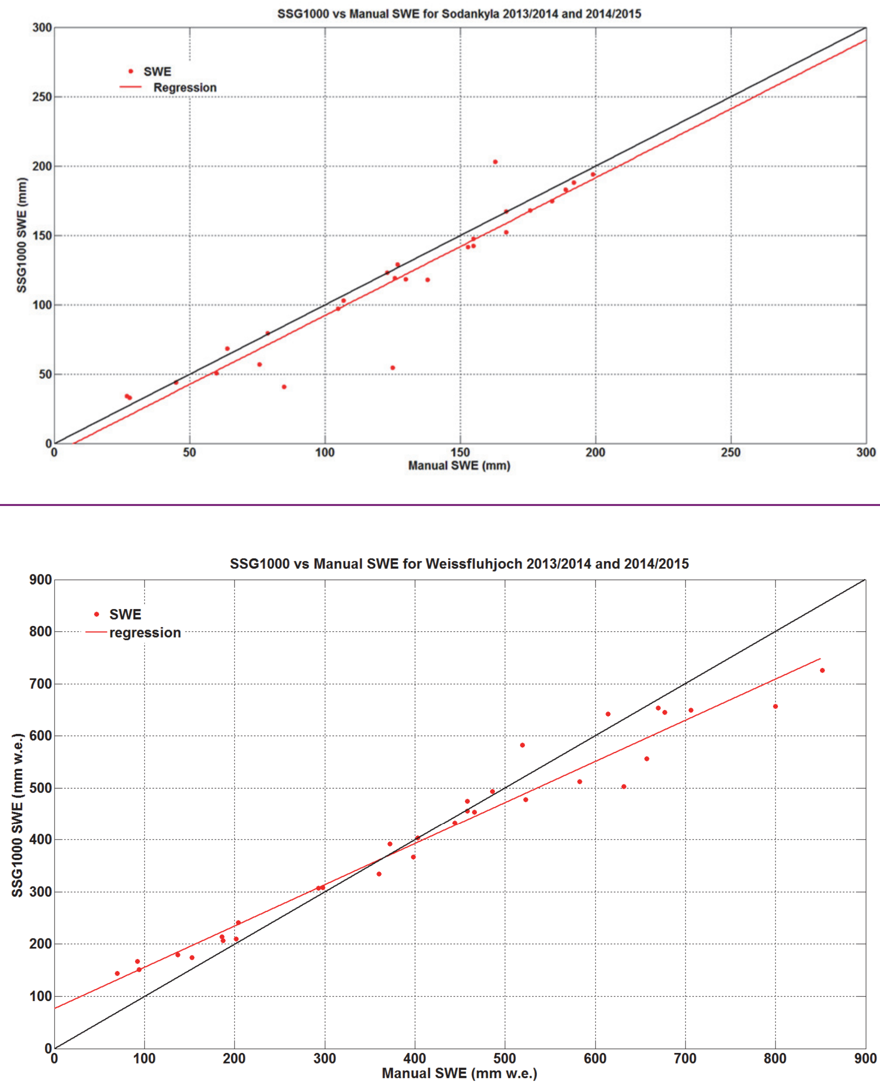


Figure 7: SSG1000 vs Manual SWE at Sodankylä (top) and Weissfluhjoch (bottom) for the 2013/2014 and 2014/2015 seasons. Black line is 1:1.

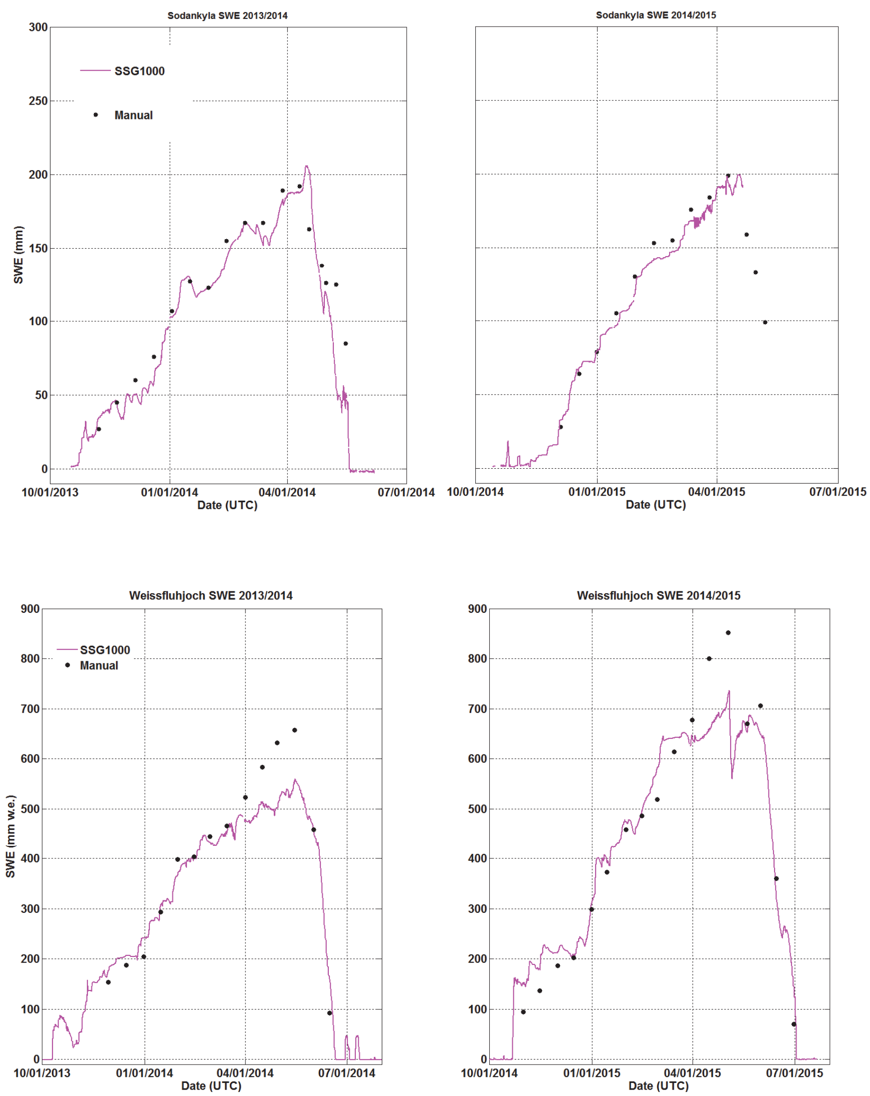


Figure 8: Time series of the SSG1000 SWE sensors and manual SWE measurements at Sodankylä (top) and Weissfluhjoch (bottom) for the 2013/2014 (left) and 2014/2015 (right) seasons.

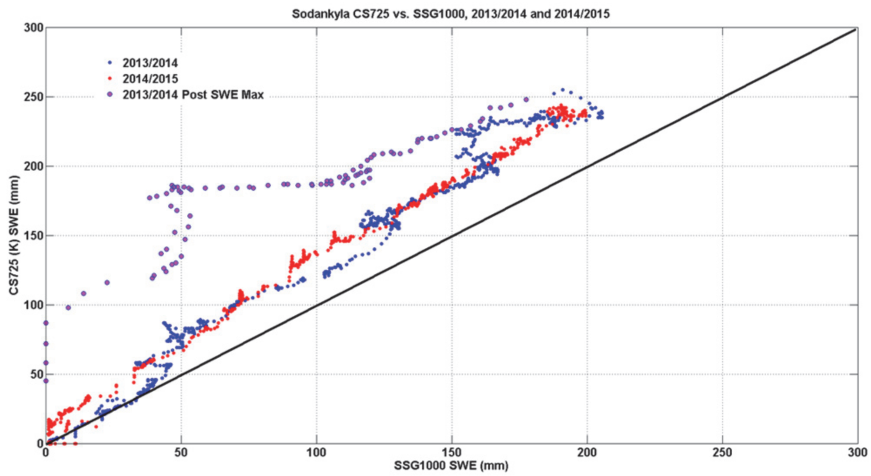


Figure ~~8~~9: CS725 vs SSG1000 for the 2013/2014 (blue/magenta) and 2014/2015 (red) seasons at Sodankylä. Black line is 1:1.