

**Review of Senese et al (2017) "Snow data intercomparison on remote and glacierized high elevation areas (Forni Glacier, Italy), August 18, 2017**

The paper by Senese et al. presents a methodology for calculating **accumulating** snow water equivalent (SWE) at a remote site located on a high elevation glacier in northern Italy. The method uses an average new snowfall density estimated for the site and calculates accumulating SWE on an event by event basis from snow depths measured by the automated snow depth sensor. The accumulated SWE is then compared to estimates made from periodic snow pits (usually only once per season near peak accumulation) and snow pillow measurements. Since the focus of the paper (as written and as per my suggestions below) is on SWE, I would suggest revising the title to "Estimating snow water equivalent on remote and glacierized high elevation areas (Forni Glacier, Italy)".

The technique is interesting in that it provides an estimate, albeit a potentially crude estimate, for total SWE using a relatively inexpensive, low power, low maintenance, and reliable (in my opinion) instrument and is presented as an alternative for estimating SWE at remote locations such as glacier monitoring and research in high alpine regions. The technique itself is filled with potential issues (and the authors point some of these out) and should only be used to produce an approximation of total SWE.

I have a concern about how the average new snowfall density is estimated. The way I understand it from the methodology, density is backwards calculated from snow depth and pit derived SWE using the same data as shown in the intercomparison (Fig 4) so the intercomparison between snow pit SWE and depth derived SWE is not independent. An independent estimate of average new snow density would have been more appropriate. An alternative would be to calculate the average new snow density from the first couple of seasons and then apply this to the remaining seasons. The lack of independence should be stated clearly and the intercomparison should focus on the snow pillow data as independent (mostly) validation.

As I mentioned, and as the authors point out, the technique has some potential issues: 1) it is highly sensitive to the estimation of average new snow density, which can vary substantially depending on both atmospheric and surface conditions, so the error in individual events could be quite large and therefore the value is only in the estimation of peak SWE and 2) the technique depends on determining event based changes in snow depth and the subsequent calculation and accumulation of SWE from those events, so missed events due to gaps in snow depth data invalidates the calculation of peak SWE.

My biggest concern with this paper is that the issues are not adequately addressed in the discussion. For example, the discussion of issues related to the use of a mean snow density is only a few lines long and should be expanded substantially. Discuss the potential errors in individual events and how this impacts your peak estimation. A large event (big increase in snow depth) that has a large deviation from the mean density will result in larger errors (e.g. a heavy wet snowfall). What is the potential for this to occur at this site? Add a discussion about missing data as this is the greatest threat to failure of the

technique. Can you do gap filling with photographed snow stakes? Would you recommend redundant sensors? More specific comments are listed below.

Note that the units for SWE should be reported in mm water equivalent (w.e.) or  $\text{kg m}^{-2}$  and not m w.e. . Snow depths should also be reported in cm and not m. It would also be useful if you used an abbreviation for “sonic ranger-derived SWE” such as  $\text{SWE}_{\text{SR}}$  and use this throughout the paper.

## Abstract

Page 1, Line 15: “...on the Forni Glacier **in Italy.**”

Lines 19-20: This sentence misses the mark. From what I have read, you are not really assessing the mean value of new snow density...this was done elsewhere. I think you miss explicitly stating the aim of the analysis and the value of this paper. You should state here that you are using mean new snowfall density and automated depth measurements to estimate the SWE of new snowfall and accumulating to estimate peak SWE and evaluating against other methods.

Line 21: “rather good” is a vague and subjective description of the estimation. Avoid this and/or quantify the estimation.

## Introduction

Page 2, Line 37: “...**often** only snowfall measurements are available...”

Line 38: “assess” should be “calculate” and “depending” should be “depends”. Fresh snow density also depends on surface conditions, correct?

Line 50: CryoNet is more of a network attached to the GCW initiative, rather than a “project”

Line 71: “detail” not “details”

Page 3, Lines 89-93: For item ii, I’m not sure that you are “defining the reliability of...” because you don’t really have a solid reference (more on that later) to be able to do that. I would rather you said that you were “assessing the capability of...”. You also use the term “obtained SWE” here...now I assume that you are referring to the derived SWE from the depth measurements so you need to be more specific here.

## Data and Methods

Page 3, Line 100: You refer to “These sensors” but you should rather say “These measurements were made at the two stations...” since not all of the measurements were made with sensors.

Line 106: Was the T/RH sensor shielded? I assume yes.

Page 4, Line 110: You should cite Beaumont (1965) here when introducing the snow pillow methods. Also some grammar issues with the first sentence on this page.

Line 116: I think “constrictions” should be “challenges”.

Line 121: “represented” should be “provided” or “supplied”. “deep” should be “thorough”.

Line 122: “working” should perhaps be “operation”.

Line 123: “...due to ice flow, etc.) **is required** before installation...”

Lines 125-128: Not sure if this paragraph adds anything to the paper

Line 130: I think that I know what you mean by “adjust” but you should clarify this.

Line 132: What are ring faults?

Line 146-147: “...from snow depth acquired by sonic rangiers **and estimated new snowfall density.**” Clarify what you mean by the last sentence “In particular...”. This is a crucial piece of the methodology and you should describe this better. What time did you use for start/end? Did you do any noise filtering? How do you account for redistribution, settling, and sublimation during the day? (should also be discussed in the Discussion section).

Page 5, Lines 152-153: What do you mean by “unique date”?

Lines 150-153: This is the methodology that isn’t clearly described (see note above).

## Results

Beside expressing the errors or biases in absolute units, it would be helpful if the relative bias was stated. Same can be said for subsequent sections.

Lines 159-161: Can you comment here or in the discussion on any potential impacts on the analysis due to the site move.

Lines 166-169: The intercomparison between the SR50 and the USH8 is interesting but largely irrelevant for this paper. I would omit this. What you could mention is that the two instruments had a correlation of ?? and that from 2015 onwards, redundancy in the snow depth measurement could mean better data for the SWE estimate.

Line 174: What is “a very good agreement”?

Line 180: “elaborate” should be “accumulate”

Figure 4: it would be very useful if the missing SR50 data could be indicated (e.g. different colour line, etc). This would certainly help with the interpretation of the graph. One might argue that all of the estimated SWE should be set to missing after large gaps in the snow depth data since it is an accumulated result.

Page 6, Line 188: “...thus suggesting a correct working of the sensor.” This is more than a sensor, but rather a technique or process. This sentence is awkward and should be reworded.

Line 192: typo “derided”.

Line 194: Change “raises” to “increases”.

Lines 196-197: You mention the snow course data using the snow tube and suggest a “large spatial variability”. Perhaps you could report the average and standard deviation of this data and in the discussion, relate this variability to the differences you see between the SR50 SWE estimation and the snow pillow and snow pit.

Lines 200-201: Oversampling of the tube is a potential error and should be included in the discussion and not in the results. You need to estimate this error and put it into context with the differences between the methods.

## **Discussion**

The start of the Discussion paragraph needs a couple of sentences of introduction to “frame” the problem and set up the following paragraphs.

Lines 205-208: The whole success of technique hinges on the approximation of the average new snowfall density but yet this only gets a few sentences in the discussion. This needs to be expanded. When would you expect errors to be greatest? At this site, can you potentially get large snowfalls that have much greater (or lower) densities than the average, which could then bias your accumulated SWE? Any suggestions on how to better estimate new snowfall SWE, perhaps at sites where you don’t have snow pits to back-calculate the density.

Line 205: “Once our procedure was verified, we performed...”

Line 209: What is a “general good agreement”. Quantify this.

Line 214: “not constant” should be “inconsistent”

Page 7, Lines 223-231: You reference some issues with the snow pillow and then state that your snow pillow only seems to be working with depths greater than 50cm. Can you relate any of the referenced potential issues to the errors that you are seeing at the site? I’m not convinced that your issues are related to those referenced so this discussion needs to be stronger.

Lines 232-246: This paragraph is more of a justification for using this technique and is more appropriate for the introduction than the discussion. The sentences (Lines 246-250) about the snow pillow is relevant, however. The subsequent discussion about the snow pit should have its own paragraph.

Line 248-249: You should refer to Fig 5 and discuss this error in context with what you are seeing at the site. This should be combined with the snow pillow discussion. Also, is there a reference for this statement about minimum snow pillow measurements?

Line 252: What do you mean by “whole glacier accumulation amount”?

Line 253: "...the **peak** cumulative SWE..." Same applies for the next sentence.

Line 257: Typo "closed" should be "close"

Lines 259-261: The two sentences beginning with "Finally..." seem out of place or awkward. Please re-organize these.

Page 8, Lines 261-263: The discussion about the SR50/USH8 is largely unsubstantiated and should be omitted.

I would like to see me recommendations in the discussion about how to make this technique better and how to adapt it for other remote sites. Perhaps make some recommendations on further testing, possibly to take advantage of other data collected during SPICE.

## **Conclusions**

Line 267: Was the acronym SPICE not defined earlier in the paper? If it was, it doesn't need to be defined here.

Line 269: The sentence beginning with "This has allowed..." needs some commas to read correctly.

Line 270: "...SWE values using a fresh snow density..."

Lines 270-273: I have some problems with the statement "The results achieved...". You only tested two sensors for measuring snow depth so I wouldn't go as far as saying that the SR50 is the "most suitable device" ...perhaps it is but you don't have enough information to substantiate this. You present a technique for estimating peak SWE on a glacier and you compare this technique against another technique (e.g. the snow pillow) but you don't really have a true reference so I'm not sure you can say that it is "most suitable". Rather, I would reference your relative error and despite the issues, suggest that the technique is "suitable" for estimating SWE.

Line 274: What are the relative variations as a result of the density change?

Instead of ending the conclusions with a comment about the error, perhaps end with a comment on the applicability to other remote sites.

