The Cryosphere Discuss.,
https://doi.org/10.5194/tc-2018-286-RC2, 2019

The Cryosphere
Discussions
Open Access
EGU

Interactive comment on "Converting Snow Depth to Snow Water Equivalent Using Climatological Variables" *by* David F. Hill et al.

Adam Winstral (Referee)

adam.winstral@ars.usda.gov

Received and published: 5 March 2019

The aim of this study is to develop a simple means of estimating snow densities to convert observed snow depths to snow-water-equivalent. The authors seek to use long-term climatological variables rather than station or modeled data so that snow depths garnered in remote locations without direct meteorological observations (e.g. crowd-sourced or Lidar data) can be easily and accurately converted to SWE. There is a growing need for improved means of characterizing snow density as greater amounts of snow depth data are becoming available (e.g. Lidar). Therefore this type of research is certainly warranted. Given that snow depths have always been more readily available than SWE or density data, other researchers have similarly produced methods of estimating densities. While not all of the previously developed approaches tackle the specific case presented here (i.e. meteorological data immediately preceding snow depth observation not required), the Sturm et al. (2010) and Jonas and Magnusson (2009) approaches do. The authors clearly acknowledge this and make a positive comparison of their method measured against the Sturm method. However, I find the presented Sturm comparison to be biased against the Sturm method (see further comments). They also hint at (lines 413, 493), but never provide evidence nor specifically claim to be, better than the Jonas approach. I would like to see the authors present in a more convincing manner how and why their method represents a substantial advancement over the previously published methods before I am ready to consider this manuscript worthy of publication.

This is my major concern:

The authors randomly split the aggregated CONUS, AK, and BC data into training and validation datasets (Section 2.2). They then use the "held-out" validation dataset to make the Sturm comparison (Section 3.1). So, essentially they have trained their model on data from the same locations with the same statistical metrics present in the comparison dataset. On the other hand, the comparison dataset is 100% independent of the Sturm training data. In order to present a fair comparison this needs to be done with a dataset that is totally independent from the derivation of both. The northeast dataset would be one ideal dataset for conducting this test and I'm not sure why this wasn't done. That said, it would certainly be more convincing if the inter-model comparisons were conducted over a wider range of conditions. I would also like to see direct comparisons to the Jonas method. As I stated in the above paragraph, the authors must present a convincing case that the new methodology represents an improvement over existing procedures. I just don't find that in the current manuscript.

Moderate concerns that need addressing:

I don't understand why rmse was normalized with respect to mean annual precipitation (Section 3 and Figure 8). This obviously biases the normalizations low where summer

Printer-friendly version

Discussion paper

precipitation is more common. Artifacts of this can be seen in Figure 8 (e.g. low ratios in Arizona, New Mexico, Alaska where summer precipitation can be considerable compared to winter; high ratios in eastern Sierras where synoptic summer storms are rare). This type of normalization might be appropriate for annual or longer hydrologic studies, but for this snow-based, winter-focused research the normalization should be based on either mean wintertime precipitation or better yet, mean annual snowfall. Both mean wintertime precipitation and mean annual snowfall should be easily derivable from the PRISM data already used in this study.

Graphs. There are way too many data points in the scatter plots to understand what is really going on in Figures 6 and 9, and some of the plots in Figure 11. These should be presented as either heat plots or randomly select and plot a subset of these data. Additionally and partly due the aforementioned reason, the overlapping plots in Figure 9 are impossible to fully discern.

I had difficulty accepting the reasoning for the residuals and mean biases apparent in the Figures 6b and d. I think these residuals, which are present in the validation dataset are also related to the choice of fitting a power law relationship rather than a linear least squares one. Given that the training and validation data should maintain the same statistical metrics then these residuals should be present in the training data as well. If, in fact, this is the case then the combination of a power law fit and the predominance of accumulation season samples would be the reason. My suspicion is that if a linear least squares fit was chosen then there should be near zero mean biases in both the training and validation sets given that the two sets maintain the same characteristics. I would expect that in the linear scenario, there should be a wider spread in residuals (i.e. higher rmse) but very little change in mean bias. Of course, this would be entirely different if the validation set was truly independent.

How the different datasets were used needs better clarification. I didn't understand the purpose of the manually sampled Chugach data. As far as I can tell, these data were not included in the calibration nor the validation analyses. What do these data show?

Why were they included? How do these data add anything new to the analysis? This should be clearly stated and incorporated into the story or leave the Chugach data out.

Section 2.1.2. Do these PRISM climatological variables, based on sparse station data and resolved at 800m, really pick up the heterogeneity you're aiming to capture as expressed on lines 132-37. It would be nice if you could show a spatially explicit example showing these capabilities.

Tidbits:

The residuals (e.g. Figure 6) should be presented as modeled minus observed. In this manner the underestimations of SWE appear as negative residuals rather than the positive residuals currently presented. I find this much easier to understand.

Lines 44-47 and 72-74. Each of these sentences contain two distinct thoughts that would perhaps be better if split into two sentences.

Lines 120-22. I didn't think this sentence was necessary . . . unless you turn it into reasoning that this just adds a layer of computational costs / complexity that aren't necessary for your desired application.

Lines 141-2. Might want to add something about why you would also prefer to not use NWP data that could possibly substitute for the lack of observations (i.e. computational costs, errors in NWP data).

Line 169. You also used snow pillow data from the northeast US. You might want to make that clear here . . . as in "Snow data for this project, aside from the aforementioned SNOTEL data, . . ."

Section 2.1.1.5. Might want to mention that these issues are most common in summer when vegetation grows beneath the sensor.

Line 440. Roughness of underlying terrain is certainly one factor, but couldn't there be others as well (e.g. wind redistribution).

Interactive comment on The Cryosphere Discuss., https://doi.org/10.5194/tc-2018-286, 2019.