

## ***Interactive comment on “Climate, snow, and soil moisture data set for the Tuolumne and Merced River watersheds, California, USA” by James W. Roche et al.***

**J. Lundquist (Referee)**

jdlund@uw.edu

Received and published: 27 October 2018

Overall Review:

Data in remote mountain regions is hard to obtain and hard to organize and quality control. For this reason, this dataset is worth sharing, particularly as part of a special issue on mountain datasets. However, the paper as written is very general. It does not reference other datasets available for the same region or provide specific details about this particular dataset. Therefore, I recommend a number of revisions to improve the data archive and the paper before publication.

Printer-friendly version

Discussion paper



In particular, I cannot find photographs of the sites or sketches defining what various parameters mean relative to each other (such as what exactly does “edge”, “road” and “stream” in snow depth files refer to - - how far are these from each other, and what differences are they trying to capture). I also cannot find details on soil profiles, soil characteristics, or information on how calibration was conducted to convert the raw instrument readings to volumetric water content. For using the soil moisture data, it would also be very valuable to know how deep the soil is above bedrock. (For example, at Gin Flat, Flint et al. 2008 found that water pooling on the bedrock below their soil moisture sensors was incredibly important to the soil water content evolution: <http://tenaya.ucsd.edu/~dettinge/flint08.pdf> . This paper also gives a good example of laboratory measurements of soil samples from the site, which I could not find in the data files here.)

Please forgive me if I have missed anything that is in the data files, and in that case, please consider my comments an indication that the paper and readme files should be clearer on how to find such relevant data.

Thank you, Jessica Lundquist

More specific comments are below:

The paper should reference other data sets that this data set should be used with: These include:

Lundquist et al. 2016 streamflow data (you cite this later in the paper but don't mention it has data complimentary to this paper) <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2016WR019261>

Other iSnoB snow modeling in this region with associated datasets: Hedrick et al. 2018: Paper: <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2018WR023190>

From Hedrick's acknowledgements section: "The data set used to produce the results presented in this study is available at <https://doi.org/10.5281/zenodo.1343653>. The

Printer-friendly version

Discussion paper



interpolation from point to grid for the forcing data is available as a standalone Docker container in a software repository at <https://doi.org/10.5281/zenodo.1343647>."

Link to Hedrick's data archive: <https://zenodo.org/record/1228400#.W9OJzBNKjUI>

The ASO 50-m SWE data surfaces are also available in this repository and are relevant to the data you are sharing here.

page 5, line 10; The Tuolumne sensor is representative of its PRISM grid cell, but not its PRISM elevation band. Tuolumne is in a rain shadow, and the multiplication factor used in Lundquist et al. 2016 is based on the ratio of the 800-m PRISM climate normals for the Tuolumne snow site and the Dana Meadows snow site, so you can use PRISM climate normals to scale these and should explain that. (Feel free to contact me if you're confused by what I've written here.)

I'm presuming you got rid of the data from the RAWS radiometers that had shading by trees and surrounding terrain (these sensors have problems in this area), but it would be helpful to have more information on the QC criteria for doing so. Karl Lapo has a github repository on methods for Mountain station quality control (with particular focus on solar radiometers) that may be helpful: <https://github.com/klapo/moq>

page 6, line 12: data "were" (not was) collected.

Are your basin and grid data the same or different from those in Hedrick et al. 2018 in the zenodo repository? It would be good to comment on the differences/similarities.

Did you compare the canopy data (tree height, etc.) with that available from the lidar maps in this region? Given that this area has been extensively flown by lidar, I would imagine you checked it in at least a subset of the areas and could comment on how well it compares. It would inform subsequent users of which dataset they might prefer to use.

Table 1. I'm not sure where you're getting the numbers for instrument height, as my impression is that this is quite variable across the region. Please comment on how

[Printer-friendly version](#)[Discussion paper](#)

these values were obtained and how consistent you think they are. Even better, in the data files, include metadata on sensor heights for each particular sensor.

Table 4. The canopy parameters here don't make sense with the rest of the dataset. These look like they are model parameters and thus should be reported in your modeling paper. If these are directly linked to observations, you should better explain how these were observationally derived.

Some of your sensors have much longer records than 2010-2014. It would be very helpful to reference the total potential duration of each sensor, as well as link to an archive where someone who wants access to that data could acquire it (presuming they want to QC the data themselves). Also, explain why you focus your data reporting on this short period (is it the only time frame you had UC Merced snow depth and soil moisture?)

Figure 3. You're presenting a dew point lapse rate and goodness of fit. You need to explain which stations (all? or only some?) went into this fit (because you will get quite different answers depending on what you include) and then let the reader know that there are two papers that explicitly analyzed lapse rates (temperature, Lundquist and Cayan 2007, and dew point, Feld et al. 2013) for this same area, and these papers demonstrate variability beyond a linear fit in space and time, and those data are available in conjunction with the Lundquist et al. 2016 data paper. In this context, your data could be a nice supplement to anyone wanting to follow up by combining the two datasets.

Comments concerning the dataset files:

I was able to download the data from the link. (This is good, as I have had that `_not_` work in more papers than I care to mention.) Upon opening these files, I still had many questions, as detailed below:

How were the soil moisture measurements calibrated? How do you get volumetric

[Printer-friendly version](#)[Discussion paper](#)

water content? Do you have photos and sketches of the soil types and compositions at each depth where you have a sensor? I can't find these in the files.

What is a .JNB file? There are a lot of these under Merced Level\_2 measurements, and I can't open them with a text editor. There does not appear to be a README in that folder that explains what software is required here. Also, if it is specialized software, I strongly encourage you to include the data in a more general form.

Are there pictures of the sites? Maps that identify which sites are which? It's hard to navigate the giant zip folder of all the data. Some kind of supplementary road map to start with would help a lot. A pdf with maps and pictures and explanations would help quite a bit.

For example, I open a file called "SmokyJack," and it tells me there is snow and soil at the "edge", "met", "open", "road", "stream" and "under" - - what do those mean? Is there a photo of the site? Multiple photos would help. The paper doesn't describe the design behind these installations.

The collection of all data plots is not very informative, as most of the time series are too long (and not plotted on the same graph with other regional time series) to detect outliers.

Actual photographs of the sites, actual instrument heights at specific sites, and actual information on how QC was conducted would be helpful. How were "nearby stations" selected (how close was nearby)? How much data were removed in the higher level data? When was a site determined to have too much shading of the radiometer?

What is fuel moisture temperature and how is it measured? (Is this standard? I saw an instrument placed in a stick at Dana Meadows but am not sure how such a stick was selected and/or what exactly was being measured there.)

Thank you!

[Printer-friendly version](#)[Discussion paper](#)

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2018-102>, 2018.

**ESSDD**

---

Interactive  
comment

Printer-friendly version

Discussion paper

