We thank Referee #1 for his or her timely and constructive comments. Referee #1 and #2 had some similar comments which we will address first.

A shared critique is that the dataset should include hourly SWE. To this end, we will include the hourly snow pillow measurements from CUES (water years 2012 to 2017) and Sesame (water years 2013 to ~ Feb 2017, when the pressure transducer failed). Initially, we focused on providing a complete hourly dataset over the study period, however the referee comments have convinced us to include these incomplete hourly SWE measurements.

To provide a complete record of hourly SWE over the study period, we have also decided to include snow pillow, snow depth, and tipping bucket precipitation from the nearby Mammoth Pass (CDEC code MHP) station. MHP is not located within the ski area, as CUES and Sesame are, but is just outside the boundary and receives very similar precipitation amounts to what is recorded at Sesame. We will present a correlation analysis of the two sites to justify this statement for times when good hourly precipitation was available from Sesame. MHP is at an elevation between CUES and Sesame (2835 m), but is in a forested area similar to the Sesame site, thereby eliminating some of the wind exposure problems at CUES for measuring precipitation. Because it is in a forested area, it likely has (though there are no radiometric measurements to confirm this), a similar radiation budget to Sesame, but not CUES.

As already noted in the manuscript, the precipitation measurements at CUES and Sesame have substantial shortcomings. In fact, upon examination of the Sesame data, we found long periods when the tipping bucket heater was malfunctioning during the 2011 and 2012 water years, eventually requiring replacement from the manufacturer. Likewise, as noted already in the manuscript, the pressure transducer failed at Sesame in 2017, likely due to the exceptional weight of the snowpack.

The MHP measurements are not without their own problems, which we plan to discuss extensively. For example, the tipping bucket data suffer from a number of issues including gaps–likely when the orifice was clogged or the satellite modem could not transmit–causing precipitation to jump and show late timing when compared to the manually weighed precipitation at Sesame. We tried a number of approaches to address these problems, but many could not be fixed. These inherent deficiencies demonstrate the problems with using automated hourly precipitation gauge measurements at a snowy site and reinforce the value of the manual snow measurements from Sesame.

The other main critique of Referee #1, shared with Referee #2, is that a demonstration of a snow model, forced with these measurements is needed. This brings up the validation loop issue brought up by Referee #1 and #2. With the snow pillow measurements provided at CUES, we aim to close the forcing/validation loop using a demonstration with the widely used SNOWPACK model. For WY 2012-2017, we will model the snow mass balance at CUES, validated with the snow pillow measurements. We will force the model with the radiometric measurements at CUES and use 3 different precipitation forcings: a) hourly snow depth measurements over the pillow at CUES using SNOWPACK's empirical new snow density estimate; b) hourly tipping bucket precipitation measurements from MHP; c) daily manual SWE measured at Sesame scaled to hourly measurements, with knowledge of timing based on the tipping bucket measurements and automated snow depth measurements. This approach will address the issues of not just assuming uniform precipitation over 24 hr and of precipitation phase, both brought up by Referee #1.

We provide in-line responses to the individual points below.

Anonymous Referee #1
Received and published: 15 November 2017

Bair et al. present a six year dataset of snow and energy balance measurements at Mammoth Mountain, California. The data include: (1) daily precipitation and hand- weighed SWE at the Sesame Street Snow Study Plot, (2) hourly temperature, relative humidity, and snow depth at Sesame, and (3) hourly uplooking shortwave, longwave, albedo, air temperature, wind speed/direction, relative humidity, air pressure, and snow depth data at the nearby CUES site. The authors describe the data sources, instruments, and processing routines and discuss a subset of variables over the presented record (water years 2011-2017), which include extreme wet and dry years.

Given the notable lack of energy balance measurements in the Sierra Nevada, I think this dataset fills a clear gap and would be useful to the community. I recommend publishing it in the journal after attention to the comments below.

- My most major comment is that I think the dataset may have limited usefulness for evaluating snow models. The authors argue that the dataset is useful for running mod- els (e.g., abstract and introduction) and it is true that they are providing all required data to do so (e.g., temperature, humidity, wind, precipitation, radiation). However, the main problem with the dataset is that it has the bare minimum in evaluation data. Depending on whether the albedo is used as a model input or model evaluation dataset, there are only two or three datasets to check the snow model (e.g., hourly snow depth, event- based SWE from hand measurements), and that will only provide limited insights into model behavior (in my opinion). The hand measurements of SWE and snow density are really only useful at the time of the storm event, and provide no information about what is happening to the existing snowpack in time. Hence, the presented data really provides no direct way of checking the model representation of the mass balance and energy states through time, and other data would be needed (e.g., snow pillow SWE, snow surface temperature, snow pit profile data, etc.). Unless the authors are willing to include the snow pillow data (albeit incomplete) and any other relevant evaluation data, I am not really sure how this problem can be adequately remedied. At a minimum, the authors should at least detail ideas on how the dataset could be used not only to run models but also to evaluate them, given minimal evaluation data. A direct demonstration with a snow model and the dataset would be instructive and would match other snow data papers.

See above

- Introduction: It would be useful to identify other similar snow datasets available (and cite the data papers), for example at Reynold's Mountain, Senator Beck Basin, and others. Highlighting the unique attributes of Mammoth relative to these areas would be helpful to the community.

Ok, we will do this

- I think the paper would be more useful if more specific guidance/recommendations were provided to scale the daily precipitation data to hourly. This is not trivial, given that mixed precipitation and rain are possible and hence assuming uniform precipitation over all 24 hours is not necessarily a robust approach.

Ok, and see initial paragraph

- While the snow pillow measurements do not span the entire period, I think they still hold enough value that they should be included in the dataset, without having to request from the authors. For long-term purposes, it would be more ideal if researchers ten years from now do not have to track down the authors to obtain these snow pillow data.

Ok, and see initial paragraph

TECHNICAL CORRECTIONS
- P2.L4-6: While interesting, this would be more relevant if you actually detected any such events in the snow albedo dataset. Please comment.

Even with very large grain sizes and a nadir solar zenith angle, clean snow albedo does not drop below 0.6 (Dozier et al., 2009). We show albedos below 0.6 every year, meaning surface impurities are be present. Sterle et al. (2013) confirm that the surface impurities at CUES are dust and black carbon.

- P.2, L.7-9: Given the winter recreation, please comment on what measures (if any) were in place to prevent humans from impacting measurements (e.g., skiing underneath the snow depth sensor).

Ok, we will mention that both sides are roped off and signed

- P.2, L.20: Awkward wording here because the phrase “to accurately weigh” splits the phrase “snow falls”. Please rephrase.

Ok, we will fix.

- P.3, L.1: “on as” reads oddly to me. Delete one word? - P.3, L.7: Recommend starting a new sentence at the semicolon: “one minute readings. The measurements from these gauges. . .”. -

Ok, we will fix.

P.9, L18-25: It is not clear what “peak base depth” or even “base depth” means.
Please clarify.

Ok, will change to “peak snow depth”.

- P.12, L1-4: This is a long sentence that really would be better framed as two sen- tences.

Ok, we will fix.

TABLE AND FIGURE COMMENTS
- Figure 10: Please confirm these are hourly values and the period of record included in this figure.

There is no figure 10. Figure 7 maybe? We will add that these are hourly values for the period of record.

DATASET COMMENTS

- In the daily precipitation table, please include measurement units with each variable name in the header.

Ok, we will fix, good point. We should note that we've kept only these hand weighed measurements in Imperial units, since they were taken this way, and the notes often refer to the measurements in these units.

Also, it may help to have a metadata file describing what each of the columns means, as there are some that I think are not necessary self-evident.

Ok, we will create a file to describe the column headers.

It would be useful to have some guidance on how to use the precipitation data, as only days with precipitation appear to be recorded in the table. Is it safe to assume these are all 24 hour measurements and days not in this table have no precipitation?

This is already addressed in the text on p 2, l 20

"We provide all the manual Sesame Snow Study Plot measurements (Table 1) for days with precipitation, based on the morning daily weather observations, posted on as the "Storm Summaries" on http://patrol.mammothmountain.com."

- There is a value of 90% snow density on October 19, 2015 which is physically unlikely, as it is close to the density of pure ice. Please check.

The reported "density" is simply the water equivalent (WE) / New Snow (HN), in this case (0.45" WE/0.5" HN). In the previously suggested metadata, we will clarify this.

- At both sites, there are RH values exceeding 1.0. Please provide additional quality control.

RH can and should exceed 1.0 regularly at this site.

- At Sesame, there are wildly varying RH values in July 2017 and early August 2017. Please provide additional quality control.

Ok, we will fix those values.

Dozier, J., Green, R.O., Nolin, A.W. and Painter, T.H., 2009. Interpretation of snow properties from imaging spectrometry. Remote Sensing of Environment, 113: S25-S37.

Sterle, K.M., McConnell, J.R., Dozier, J., Edwards, R. and Flanner, M.G., 2013. Retention and radiative forcing of black carbon in eastern Sierra Nevada snow. The Cryosphere, 7: 365-374.