November 14, 2024
Evaluation of high resolution snowpack simulations from global datasets and comparison with Sentinel-1 snow depth retrievals in the Sierra Nevada, USA
By: L. Sourp et al.

Dear Editor

We replied below to the third reviewer who provided many relevant suggestions. We also noted that after the second round of review, you suggested to "clearly and more explicitly link the discussion to specific results (...) Overall, try to avoid the reader asking him/herself: "So what?". We believe that the modifications made in response to the reviewers should answer this recommendation as well. In addition, we highlight that we wrote a clear statement in the end of the discussion "*we can conclude that ERA5-SnowModel is promising for water resources applications. This pipeline can be used to simulate SWE in near real time without the need of in situ measurements.*"

Many thanks again for handling our manuscript,

Laura Sourp and Simon Gascoin on behalf the co-authors.

General Comments

In this paper, the authors create a modeling pipeline which leverages global-scale meteorological analyses (ERA5 and ERA5-Land) to force SnowModel and produce 100 m SWE estimates in the Tuolumne River Basin for a seven year study period. They then compare the modeled results to ASO lidar SWE/snow depth and S1-derived snow depths, showing surprisingly performance even with the coarse scale forcing data. The results presented here are novel/robust and provide a baseline to extend this modeling technique to other data sparse regions around the globe, which could provide marked enhancements to SWE/water resource forecasting.
The manuscript is clearly written and the comments from the previous reviewers have all been properly addressed and integrated into the text. I find the discussion of the results well throughout and the figures well presented. The comments provided below are mostly minor, the most important being the adding of numerical error statics to the paragraph between L233–245. Once these are addressed this article is suitable for publication and will be a solid edition to TC. Congrats on a very nice study!

-Jack Tarricone

We would like to sincerely thank Jack Tarricone for the positive comments on our work and the very useful review. We have addressed every specific comment in a revised manuscript as explained below.

Specific Comments

L32: How does Pléiades retrieve snow depth? I know, but a bit more information on stereo photogrammetry would be good for a broader audience. Same thing for ICESat-2.

We have added a sentence for both methods to explain a bit more how they work (L33-35). "*Pléiades* ***very high resolution stereoscopic images can be used to generate snow depth images by differencing two digital elevation models*** (...) *ICEsat-2* ***lidar altimetry*** *has the potential to provide snow depth data at global scale but with a sparse sampling".*

L33: Add a bit of info on co/cross pol S1 retrieval algorithm

We expanded the description of the Sentinel-1 algorithm: *"This method,* ***which is based on an empirical change detection method applied to the cross-polarization ratio*** *(...)"*

L70: "worldwide" is a bit confusing, as it reads on first pass this SWE data is worldwide. Maybe "Globally publicly available from this basin" ?

We agree that it could be confusing. We reformulated as follows : *"The ASO dataset on the Tuolumne catchment is the densest time series of high resolution snow depth (3 m) and SWE (50 m) maps* ***publicly available at this scale (1100 km²) in the world****."*

L113: I like the discussion of the density uncertainty here, but would add a reference to Raleigh & Small (2017) and rework for a more robust statement.

Many thanks for this relevant suggestion which strenghtens our uncertainty estimate. Based on the work of Rayleigh & Small (2017) we have revised our uncertainty estimate from 0.02-0.05 m w.e to 50 kg/m³. Indeed, Rayleigh & Small (2017) estimated an uncertainty in modeled density of 48 kg/m³ in the Tuolumne basin. Therefore, for a 1 m deep snowpack and an uncertainty in snow density of 50 kg/m³, the uncertainty of the 50 m SWE products is 0.05 m w.e. We note however that this uncertainty can be regarded as a conservative estimate as in situ measurements of snow density are also used by the ASO to adjust their density model (Painter et al., 2016). We modified the text accordingly L117-L121.

L149–153: I recommend creating a table in the appendix with all model parameters mentioned in this text so future work attempting to replicate your work knows exactly what you did. While I see they're buried in the Github page, I'm not sure exactly where to find them.

We agree that reproducibility is important. Apart from the parameters that we modified for this study and that we explicitly mentioned in the manuscript, there are many default parameters which would make a very large table. Therefore, we included the SnowModel configuration file which includes all the parameters and their description in our Github repository and modified the text as follows:

We set all SnowModel parameters (the curvature length scale, curvature and wind slope weights, minimum wind speed, precipitations schemes for downscaling or for rain-snow fractions, subcanopy radiations schemes, various thresholds for wind transport calculations) to the default values (see the parameter file snowmodel.par in the code availability section)

L196–202: A bit confused on how you're referencing "interquartiles" here. Also the text states, "...for the wet year 2017, they peak respectively at 0.64 and 0.82 m." Yet, when I look at the boxplots it seems 06-04 has the max value, which barely extends below –0.5.
Please check this paragraph for clarity and correct numbers.

We are referring to the interquartile as the difference between the 75th percentile (upper quartile) and the 25th percentile (lower quartile). This means that we are not looking at the maximum value of a lower or upper quartile but the distance between these two quartiles. It is used as an indicator of the dispersion of the datasets. The numbers are correct : maximum interquartile with ERA5 in the 2017-05-02 with 0.64 m and with ERA5-Land it is the 2017-04-01 with 0.82 m. To clarify, we added this sentence: "*The spread of the residuals are shown with the interquartile (i.e., the difference between the 25 and 75th percentiles) inside the colored boxes, and with the 5-95th percentiles inside the whiskers.*"

L214: Note TC date formatting requirements https://www.the-cryosphere.net/submission.html): "1 April 2016", I won't ask you to do this but will likely need to be updated in the copy editing stage!

Thanks, we corrected it.

L225: Why is the resampling procedure set up this way? It seems like you're losing valuable information if you're throwing away a whole 1 km pixel if 1 of 400 50 m pixels is missing. Not saying it's incorrect but some justification of why this is the proper method should then be added in Section 2.2.3 then.

We chose this resampling method because we do not know if the missing data inside a 1km cell are evenly distributed among the values of the valid ASO HS 3 m pixels. We preferred losing some data that induce an incorrect bias. Also, most of the missing data are located at low elevation where there is almost no snow so this should not affect our statistics too much.

We propose this reformulation in section 2.2.4 : *We applied another validity mask for the cells where the snow depth is not always available to all three snow depth datasets (here representing 8.5% of missing data in the catchment). The missing values in the 3 m resolution ASO dataset are propagated at the 1 km resolution validity mask. This decreases the number of observations but ensures that the resampled 1 km snow depths maps are not biased by the spatial distribution of non-valid pixels in the 3 m ASO snow depth dataset.*

L233–245: Add values to bias, SD, R^2, and RMSE when referenced, this will likely require some tweaking of the language as well. You've performed solid analysis that is not being communicated clearly in this paragraph!
We propose the following reformulation :
Figure 7 shows the Sentinel-1 observed and SnowModel simulated snow depth compared to the ASO observed snow depth, resampled to a 1 km resolution. On the 2017-03-03, Sentinel-1 has the lower bias (-0.43 m), standard deviation (0.86 m) and RMSE (0.96 m). These statistics are close to the ERA5-SnowModel simulations (respectively -0.49 m, 0.9 m, 1.02 m) while ERA5-Land-SnowModel simulations have a greater bias (-0.83 m) and RMSE

(1.2 m) with a comparable standard deviation (0.86 m). On the second date, the 2018-05-01, Sentinel-1 still performs the best with a bias of -0.05 m, and standard deviation and RMSE both equals to 0.21 m,. On this date, ERA5-Land-SnowModel simulations are similar to Sentinel-1 with a bias of -0.09 m, standard deviation of 0.26 m and RMSE of 0.27 m; while ERA5-SnowModel simulations underperform with a 0.16 m bias, a 0.41 m standard deviation and a 0.44 m RMSE.. Finally on the 2019-03-24, the closer data to the ASO snow depths is the ERA5-SnowModel simulations with an bias of -0.65 m, a standard deviation of 0.81 m and an RMSE of 1.04 m. Sentinel-1 data have the highest bias (-1.24 m) and RMSE (1.38 m), but the lowest standard deviation (0.61 m). ERA5-Land-SnowModel simulations also have a high bias (-0.92 m) and RMSE (1.17 m), with a standard deviation of 0.73 m.

L237: Remove "seems to be" – no need for subjective language when you've conducted numerical analysis. Use the error metrics you calculated and state the performance of each Dataset!
We corrected it with "is", cf previous answer.

L240: How do we know S1 it underestimates? State specific metrics used to support this Sentence.
We change the sentence to : *In 2018, both the ASO and Sentinel-1 observed really low snow depths (<1 m) but there is still a negative bias (-0.05 m) in the Sentinel snow depth distribution*

Figure 7: Provide number of values in each scatter plot (n = xx). I only say this because you said there are different numbers in each, so the reader should know how much that varies.

That is right, here is the new version that we included in the revised manuscript.

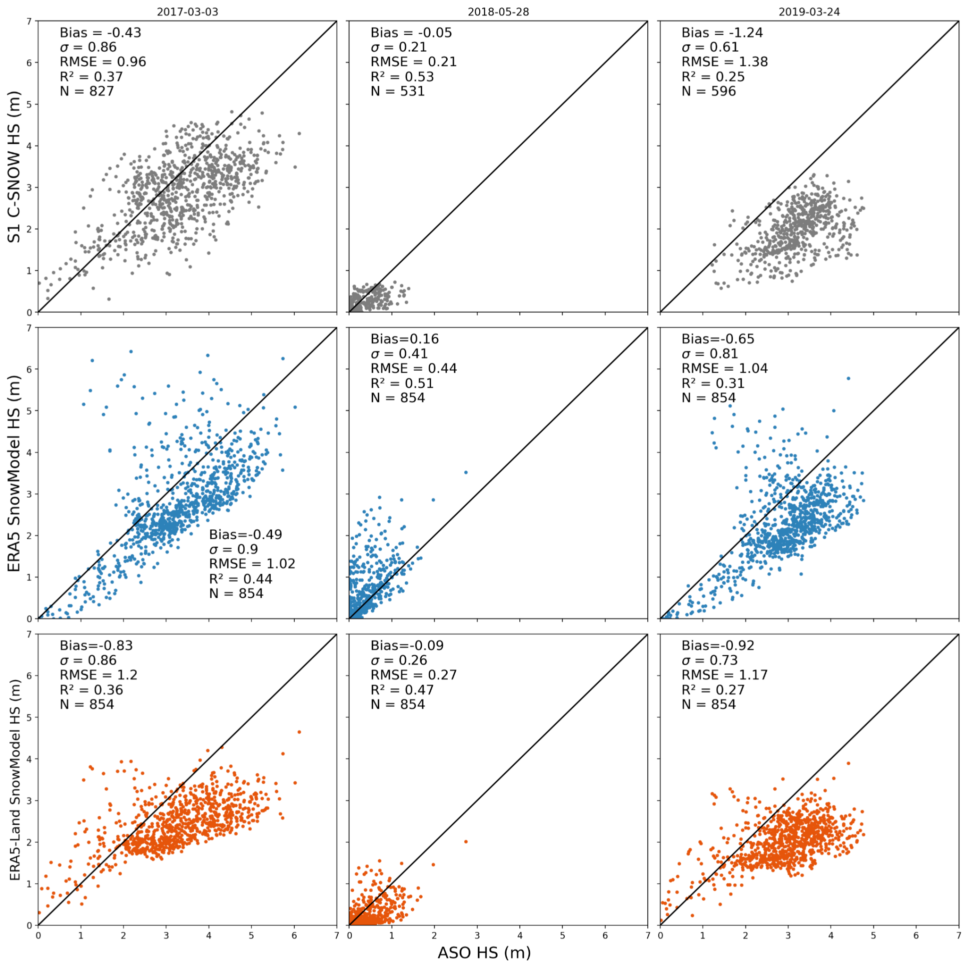


Figure 7: Scatter plots representing the observed and SnowModel simulated snow depth data as a function of ASO snow depth data, with a one to one line in black. All data are resampled at 1 km resolution. N is the number of values in each plot.

L273: I would add a figure in the appendix of this analysis, as you're referencing something you did but provide no data/figure to back it up.

We agree, here is the figure that we added to the appendix :

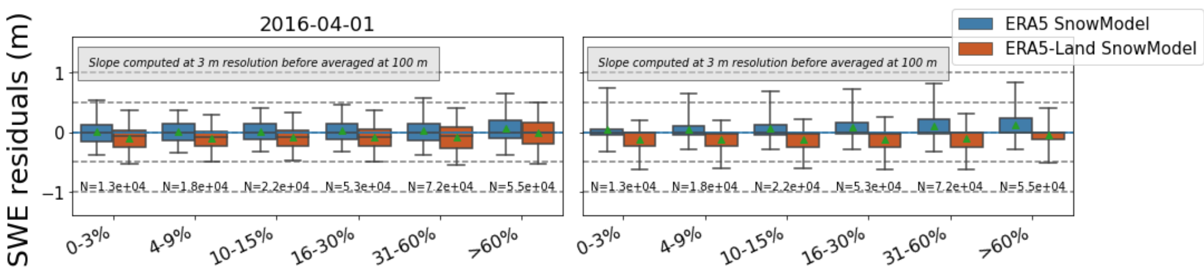


Figure A2: Distribution of the residuals between the SnowModel simulated SWE and the ASO SWE at 100 m resolution in the Tuolumne river catchment (in m w.e.) on the 1st of April 2016 (left) and the 27th of May (right), stratified by slope. Whiskers show the 5-95 percentile, the line in each box represents the median of the distribution and the green triangle shows the mean. Slope has been calculated using the DEM at 3 m resolution and has been resampled with an average algorithm at 100 m.

L300: I would add some context to the Shao et al. RMSE of 0.04 m for ERA5-Land. What are some of the uncertainties associated with validating a 9 km pixel against point-based in situ observations? Would these be magnified in complex mountain terrain?

This is a relevant question that we cannot answer easily. However, we can refer to a recent work by Mortimer et al. (2024) who have discussed this issue in the case of several reanalysis products including ERA5-Land.

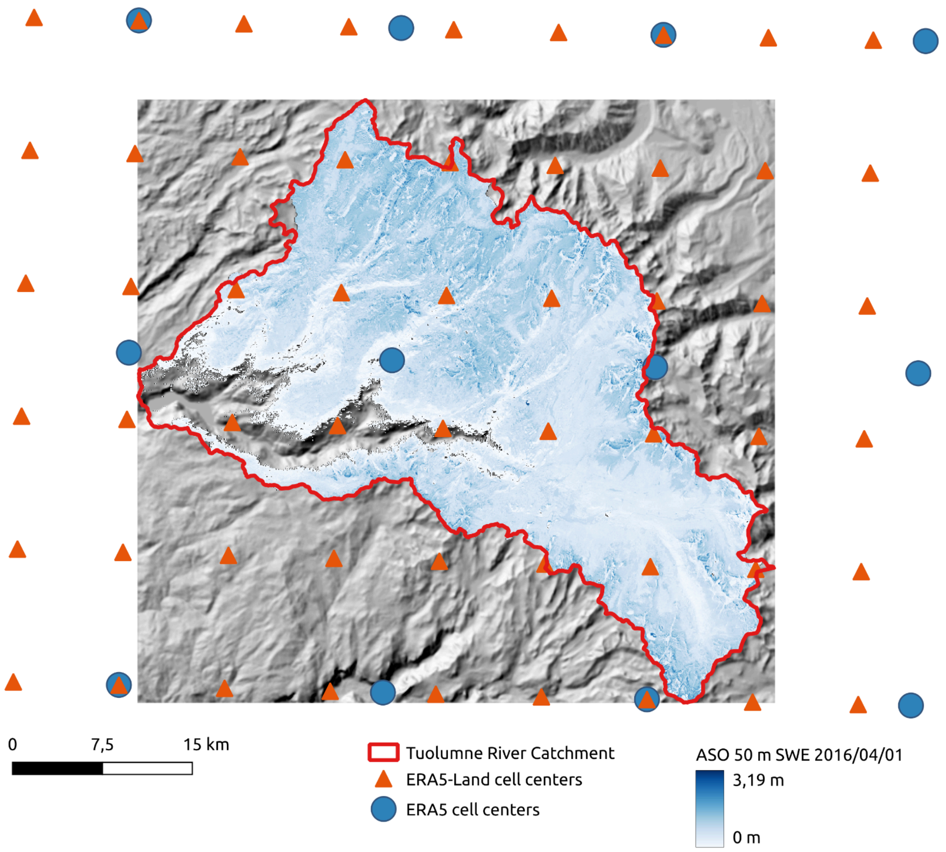
*Shao et al. (2022) found a **similar** accuracy of the ERA5-Land SWE dataset with an RMSE below 0.04 m w.e. in regions north of 45°N. This evaluation was performed using point-scale in situ measurements over large flat regions but not in complex mountain terrain like the Tuolumne Basin where the high spatial variability of SWE makes such evaluation more challenging **(Mortimer et al. 2024).***

L309: Not sure I totally agree here, "seems to represent quite well" yet R^2 0.25–0.53. Maybe "agrees moderately"?
Yes agreed and changed accordingly.

L319: Recent work has shown S1 struggles in shallow snow (<1.5 m), as there is almost no physical co/cross pol backscattering signal detectable (Broxton et al., 2024; Hoppinen et al., 2024). The technique has been shown to work well in the Alps and moderately well here as they both have deeper snowpacks. I would caution against recommending it for operational use as (1) Many snowpacks are not deep and therefore not well suited, (2) No one has been able to produce the Lievens method to anywhere near the quality of the closed-source code he has. This supports that your modeling pipeline is superior!
Indeed these studies are relevant and could advise against an operational use in this context. We removed the sentence "*There are different error sources in the three methods which are neither insignificant nor prohibitive for an operational use*" and added the following sentence above in the same paragraph "*Other studies highlighted that the C-SNOW algorithm is not adapted to retrieve snow depth of shallower snowpack (<1.5 m) (Broxton et al., 2024; Hoppinen et al.,2024) which could be a significant obstacle for an operational use of this product.* "

Technical Comments:

Fig 1: Change color scale to 0.
Changed to :



L97: Replace '(see below)' with the specific section you're referencing.
We referenced the section 2.2.1 (Method/SnowModel) instead.

L100: Add link to C-SNOW website here.

L113: What does "w.e." mean here and throughout the manuscript? Excuse my ignorance if this is a common phrase.
It means meters of water equivalent. It has been added in the manuscript for clarification.

L132: DEM already defined.
We use DEM directly instead.

L199: Noting "w.e." again. Found a few examples of what it could but still unsure.

L210: State two dates.
The new sentence is Figure 5 shows the distribution of the residuals for two dates (2016-04-01 and 2016-05-27) by slope, elevation and aspect.

Figure 5: Added (left) and (right) in caption after corresponding dates.
Ok

L212: Remove double period.
Ok

L263: “most probably” -> likely
Ok

L308: “modelisation of snow density” -> modeled snow density
Ok

L341: “global datasets only” -> global publicly available atmospheric reanalysis datasets only
We should probably keep “global dataset only” because it is also referring to the Copernicus Land Cover and DEM, as stated in the next sentence.

L349: “near real time” -> near-real-time
Ok

Bibliography:

Broxton, P., Ehsani, M. R., & Behrangi, A. (2024). Improving Mountain Snowpack Estimation Using Machine Learning With Sentinel-1, the Airborne Snow Observatory, and University of Arizona Snowpack Data. Earth and Space Science, 11(3), e2023EA002964. https://doi.org/10.1029/2023EA002964

Hoppinen, Z., Palomaki, R. T., Brencher, G., Dunmire, D., Gagliano, E., Marziliano, A., et al. (2024). Evaluating Snow Depth Retrievals from Sentinel-1 Volume Scattering over NASA SnowEx Sites. EGUsphere, 1–35. https://doi.org/10.5194/egusphere-2024-1018

(^ published article should be coming out soon so be on the lookout for that)

Raleigh, M. S., & Small, E. E. (2017). Snowpack density modeling is the primary source of uncertainty when mapping basin-wide SWE with lidar: Uncertainties in SWE Mapping With Lidar. Geophysical Research Letters, 44(8), 3700–3709. https://doi.org/10.1002/2016GL071999