

## **Reviewer #1**

Thank you very much for your positive assessment of our work and for recommending it for final publication. We have carefully addressed your remaining minor comments below. We hope that these revisions have satisfactorily addressed your suggestions and that the revised manuscript is now suitable for publication.

**Comment:** Line 9: Is the issue that the data are coarse, or that the totals/intensities do not match? Please specify.

**Reply:** We have changed the sentence as “Key limitations may be include inadequate representation of spatial variability due to coarse resolution precipitation input and inadequate representations of snow-vegetation interactions.”

**Comment:** Line 27: Please rephrase to more clearly articulate the innovation of the study.

**Reply:** Here is the completely revised paragraph:

“Most hydrological models tend to underestimate snow over the mountainous southwestern US. This may result from uncertainties in precipitation forcing and/or inadequate representation of snow–vegetation interactions, particularly canopy interception and shading. Through model experiments, we demonstrate how downscaled meteorological forcing and improved vegetation representation can improve snow simulations over the southwestern US.”

**Comment:** Line 83: The remote sensing products referenced target either vegetation or snow. How can remote sensing techniques be used to assess snow–vegetation interactions? What products are available?

**Reply:** Thank you for this comment. Remote sensing can assess snow–vegetation interactions by jointly analyzing complementary snow and vegetation observations rather than relying on a single product that directly measures the interaction. We have modified the sentence as:

“Tennant et al. (2017) used remotely sensed snow-cover information together with vegetation cover to quantify how vegetation influences seasonal snowpack patterns across western North America, while Wang et al. (2018) combined snow and vegetation observations to investigate how winter snow conditions affect subsequent vegetation activity in northern ecosystems. Thus, spatially and temporally collocated observations of snow cover/SWE and vegetation properties such as greenness, LAI, or vegetation cover can be used to evaluate the relationships and feedbacks between snow and vegetation.”

**Comment:** - Line 110: Because downscaling is a central aspect of the study, it should be introduced more thoroughly. There is also insufficient justification for why the chosen approach was selected.

**Reply:** Dear reviewer, we have included the detail methodology for the downscaling in the revised manuscript under section 4.2.

**Comment:** Fig. 1: Can the river network be added to the map to aid orientation? The bars on the right are not useful as their lengths do not align with percentage coverage—please adjust. Also, why are CDEC stations shown but not CDWR or SNOTEL stations?

**Reply:** We have not discussed runoff routing or streamflow generation in this study; therefore, the streamflow network is not shown. These processes will be considered in future work. The vegetation bars are intended only to illustrate the relative vegetation cover associated with each vegetation class. To improve clarity and provide a more quantitative representation, we have additionally included the vegetation-cover percentages in the revised figure.

**Comment:** Line 355: Why was this downscaling approach selected? Please justify. If PRISM data are included in the downscaling, why do the results differ so much from PRISM?

**Reply:** We selected the WRF-Hydro Meteorological Forcing Engine because it provides a consistent framework for downscaling the coarse-resolution NLDAS-2 meteorological forcing to the 1-km model grid while accounting for important topographic effects associated with elevation, slope, and aspect. This is particularly important for the mountainous Sacramento and San Joaquin basins, where strong terrain gradients can produce substantial spatial variability in precipitation, temperature, and radiation. The Mountain Mapper approach was specifically developed to improve precipitation estimates in complex terrain by incorporating the spatial climatological information from PRISM together with gauge observations.

We would also like to clarify that PRISM is not used as the final precipitation forcing in this procedure. Rather, PRISM provides the climatological spatial distribution used in the Mountain Mapper correction, while hourly gauge observations are used to calculate local bias ratios that are spatially interpolated and applied in producing the high-resolution precipitation field. Therefore, inclusion of PRISM in the downscaling procedure does not constrain the resulting downscaled NLDAS-2 precipitation to reproduce PRISM precipitation magnitudes. Differences between the downscaled NLDAS-2 and PRISM are therefore expected because the products represent different temporal and methodological frameworks: PRISM provides the climatological spatial information used by the correction, whereas the downscaled forcing retains the temporal variability and characteristics of the

underlying forcing and incorporates gauge-based corrections. We have clarified this distinction and the rationale for selecting the downscaling approach in Section 4.2.

**Comment:** Line 380: Why is a warm-up period required?

**Reply:** The warm-up (spin-up) period is required to minimize the influence of model initial conditions on the simulated land-surface states. Noah-MP contains prognostic variables, including soil moisture, soil temperature, and snow-related states, that require time to adjust to the meteorological forcing and reach dynamically consistent conditions. Without an adequate spin-up period, simulations during the initial analysis period may reflect initialization artifacts rather than the modeled land-surface processes.

**Comment:** - Line 382: Please include more details on the snow layering and add them to the snow routing description.

**Reply:** Noah-MP represents the snowpack using a dynamic multi-layer scheme with up to three snow layers, with the number and thickness of layers varying according to snow depth. The multilayer representation accounts for the vertical evolution of snow temperature, density, liquid water, compaction, and melt/refreezing processes. We have already detailed these processes in section 4.1.

**Comment:** - Fig. 3: Please revise the color scheme. Rainbow palettes are not recommended for scientific publications.

**Reply:** Dear reviewer, we have already changed the colors from the first version based on the second reviewer comment.

**Comment:** Line 457: Does SNOWDAS account for snow–canopy interactions? Why was this model chosen as the reference? Please provide a stronger justification.

**Reply:** SNODAS is not a purely observational snow product; it combines a physically based energy- and mass-balance snow model with assimilation of available ground-based, airborne, and satellite snow observations to provide spatially continuous estimates of snowpack conditions at approximately 1-km resolution. The underlying NOHRSC snow-modeling framework includes parameterizations for forest-cover effects; however, its treatment of snow–canopy processes differ from the explicit vegetation-dependent processes represented in Noah-MP. Therefore, we do not consider SNODAS an independent benchmark specifically for evaluating snow–canopy interactions. Rather, SNODAS was selected as a reference because it provides spatially continuous SWE estimates at a resolution comparable to our 1-km simulations and integrates multiple observational sources, making it particularly useful for evaluating the spatial and temporal evolution of SWE across the Sacramento and San Joaquin basins.

**Comment:** Figs. 5, 6, and Table 3: Emphasizing improved results when switching the vegetation model may be misleading, as there is no clear reason to expect better performance at stations in open areas. Often, the dynamic model appears to reduce underestimation by decreasing shading effects; however, this can represent improvement for the wrong reasons.

**Reply:** We thank the reviewer for this important comment. We agree that the effects of meteorological forcing and vegetation representation need to be more clearly distinguished, and we have revised the manuscript accordingly. In particular, we now evaluate the effect of dynamic vegetation separately within each meteorological forcing by comparing DynVeg\_Pred against DynVeg\_Off under NLDAS-2 and AORC independently. This paired comparison keeps the meteorological forcing unchanged and therefore provides a more direct assessment of the changes associated with the LAI/GVF representation, rather than attributing differences between the NLDAS-2 and AORC simulations to vegetation processes.

To strengthen the station-scale analysis, we have revised Figures 5 and 6 so that the CDWR stations are now ordered according to the magnitude of the mean LAI difference ( $\Delta\text{LAI}$ ) between DynVeg\_Pred and DynVeg\_Off rather than by elevation. We additionally report the change in SWE performance at each station as  $R^2$  ( $\Delta R^2 = R^2_{\text{DynVeg\_Pred}} - R^2_{\text{DynVeg\_Off}}$ ) separately for NLDAS-2 and AORC. This analysis shows that the vegetation effect is spatially heterogeneous and forcing dependent. Under NLDAS-2, several stations with substantial LAI changes show improved SWE performance, whereas comparable LAI changes do not consistently lead to improvements under AORC. We have revised the discussion to explicitly acknowledge this behavior and to avoid implying that all differences between the simulations can be attributed solely to improved snow–vegetation interactions.

Here are the revised figures (please see reviewer #2 reply as well):

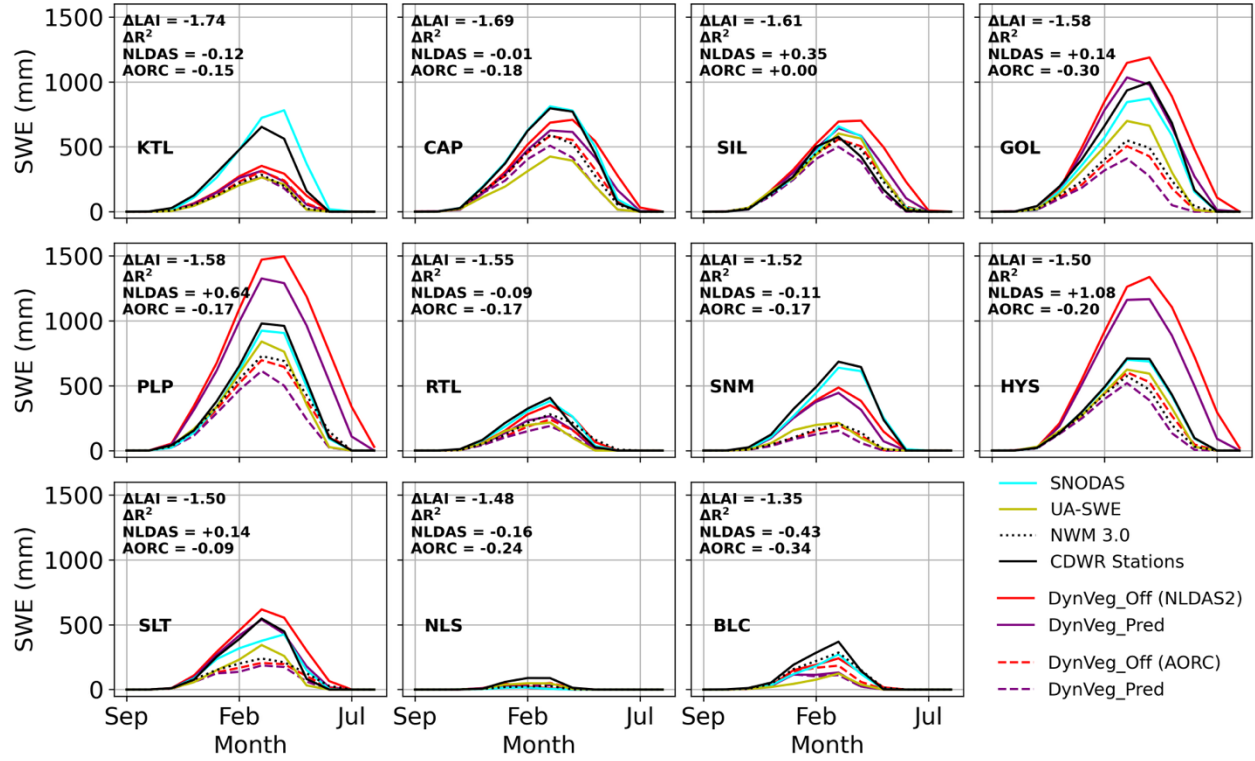


**Figure 5:** Snow water equivalent (SWE) comparison at CDWR stations in the Sacramento River Basin as represented in study area figure (Fig. 1). Monthly mean of control and dynamic vegetation modeled SWE with NLDAS-2 (solid lines) and AORC (dotted lines) forcing are plotted against the gridded observation (UA-SWE, SNODAS and NWM 3.0), and CDWR station (solid black line). The stations are sorted according to the difference in the LAI between both simulations, improvement in  $R^2$  is also reported for each plot.

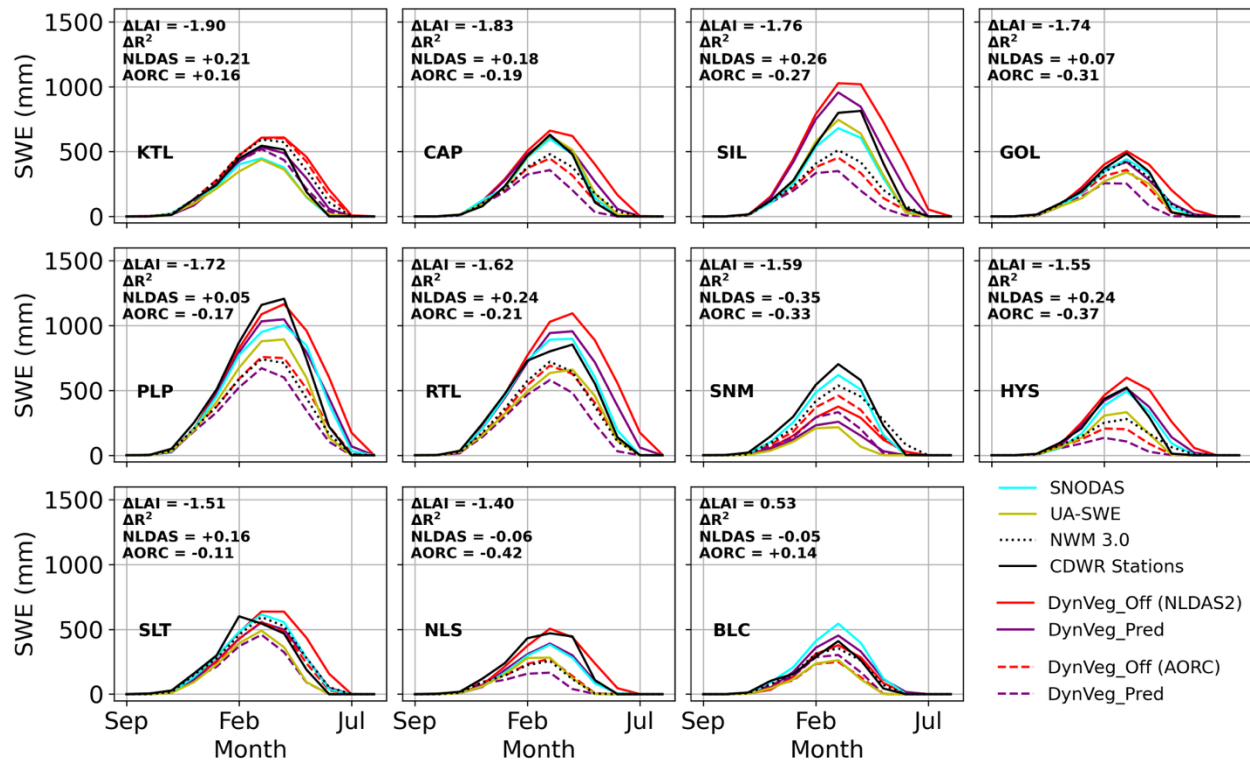


**Figure 6:** Snow water equivalent (SWE) comparison at CDWR stations in the San Joaquin River Basin as represented in study area figure (Fig. 1). Monthly mean of control and dynamic vegetation modeled SWE with NLDAS-2 (solid lines) and AORC (dotted lines) forcing are plotted against the gridded observation (UA-SWE, SNODAS and NWM 3.0), and CDWR station (solid black line). The stations are sorted according to the difference in the LAI between both simulations, improvement in  $R^2$  is also reported for each plot.

**Comment:** - Lines 574–575: Where do you show this? As I understand it, radiation is reduced because LAI is higher. Where do you demonstrate that snow, in addition to higher LAI, decreases incoming shortwave radiation?

**Reply:** We have completely removed this sentence from the manuscript in the revised version.

**Comment:** Line 580: If static LAI parameters are well calibrated, they may outperform the dynamic LAI model, especially since LAI changes little in winter (limited plant activity). Would adjusting LAI parameters to match observed values improve the modeling? Please comment.

**Reply:** We agree that a well-calibrated static LAI may outperform dynamic LAI, particularly during winter when LAI variability is relatively small. Adjusting LAI to better match observations could improve SWE simulations by providing a more realistic representation of canopy conditions. We have clarified that dynamic LAI does not systematically improve SWE performance; rather, it provides LAI values that are closer to the observed vegetation

conditions. Evaluating and further constraining LAI with observations remains an important direction for future work.

**Comment:** Line 581: “...better simulate SWE dynamics.” Where do you show this? Without measurements of snow beneath canopy, how can you conclude that the simulation of sub-canopy snow is improved?

**Reply:** We have revised the sentence as:

“With limited gridded estimates, we have shown that dynamic vegetation produces improved LAI estimates, which can influence the simulated SWE dynamics.”

## **Reviewer #2**

**Comment:** Review comments – “Improved snow simulations through improved representation of vegetation conditions ...” by Gupta et al. This paper was sent to me after a first round of reviews, to which the authors have responded. It’s unfortunate that I didn’t see the paper earlier but so be it. The “bottom line” from Reviewer 1 was “[the paper] currently does not sufficiently highlight its main contributions, and substantial revisions are required”, and from Reviewer 2 it was “I am not sure whether the study delivers on what was promised in the title”. Unfortunately, there were then pages and pages of detailed comments from each reviewer, which the authors have dutifully attempted to address. However, in my view the paper still suffers from Reviewer 2’s assessment -- it doesn’t offer insights into the performance of the Noah-MP snow model with respect to snow-vegetation interactions that can be deconvolved from model forcings. Frankly I think the authors need to go back to the drawing board. The stated aim certainly merits a good paper -- the interaction of snow and vegetation (primarily forests) is key to understanding the hydrologic response of watersheds in the Western U.S. over time scales from hours to days (e.g., floods) to seasons to decades (e.g., climate change). In its present form though, the paper doesn’t address that need. One basic problem with the paper is that it convolves the effects of uncertainty in the model forcings at the large scale of the Sacramento-San Joaquin River basins with the interaction of snow and vegetation. There is no relevant data at that scale that get at snow-vegetation interactions, so they quickly are into evaluating the quality of the different data sets used for model forcing and assessment. In my view, that should be the topic of a separate paper. I will say that our own limited experience with AORC was discouraging – my impression is that the developers were overly focused on high spatial resolution, and didn’t adequately recognize that the actual data sources they are forced to use (and this is critical for precipitation) are gauges with average separation 10-20 km – not close to the 1 km spatial resolution of the data set. NLDAS-2 has a more realistic spatial resolution (1/8 degree), but it has a host of its own problems, including a major shift in the stations that it ingests in the early 2000s when it became semi-operational. In my view, PRISM is probably better than the other two, primarily because it makes direct use of high elevation stations (primarily SNOTEL). Exactly what AORC does in terms of orographic adjustments to precipitation and temperature, I’m not sure. Again, while these issues need to be addressed, they’re really the topic of a separate paper. But as it is, I don’t take away a message as to how well or poorly Noah-MP is handling snow-vegetation interactions, as there’s no attempt to segregate model and data error. Given the absence of large-scale observations, the obvious alternative is to use what’s available at the process scale. The best alternative in my view would be to run the model at SNOTEL sites, using the precipitation and temperature observed at the sites. This takes away most of the problem of convolving forcing and model errors. I should note that this strategy

is not without challenges, as SnoTel sites are generally in clearings, although some of the clearings may be small enough to reduce the magnitude of this problem. The other issue though is that the authors probably were poorly advised in their choice of watersheds, as there are relatively few SnoTel sites in the Sacramento-San Joaquin watersheds. California DWR operates its own snow pillow network, which is somewhat similar to SnoTel, but I don't believe most of those sites observe precipitation. That said, I am surprised that the paper has no reference to the work of John Pomeroy or Pascal Storck, both of whom collected data on snow interception in and under forest canopies (basically by weighing the trees). Those data are reported in a 2009 WRR paper by Kostas Andreadis (also not referenced). By now, there likely are other data sets that could be used to understand performance of the Noah-MP snow-vegetation module. The other major problem I have with the paper has to do with their results from the NoahMP dynamic vegetation model. They say little about the model itself, but it appears that they're just reporting results with that switch on or off in the model. In any case, what they report makes no sense to me (their table 3 in particular, and the big differences in SWE for dynamic vegetation on and off). I Googled "What fraction of high elevation forests in California are conifers?" What I got back (sourced from USFS) was "At higher altitudes and montane belts (like the Sierra Nevada, Cascades, and Klamath ranges), vegetation is highly dominated by needle-leaved, cone-bearing trees." That matches with my observation that deciduous forests (e.g., oak) are almost entirely at lower elevations, where snow is rare. I don't know what's in the Noah-MP parameterization for snow, but I suspect that it's based on LAI, which interacts with snow in two major ways. The first is snow interception, which is strongly related to LAI. The second is attenuation through the canopy of wind and solar radiation that reach surface snowpacks; both are strongly related to LAI. So my guess is, the dynamic vegetation model somehow is changing LAI, either seasonally or over longer time scales. But why? Seasonally, conifer LAI is essentially constant. Over longer time scales, yes trees do grow and die, and there's fire. They don't say anything about this though (and in particular how well the model matches observed variations in vegetation properties over their domain). I strongly doubt that whatever the model is doing matches even remotely what vegetation observations show. My overall recommendation that the paper be rejected could be based on the dynamic vegetation results alone, notwithstanding the other critical problems I've identified above. I do want to re-emphasize that the topic of the paper is a worthy one. Noah-MP is the hydrologic core of the National Water Model, which increasingly is finding its way into both research and operations, at least in the U.S. In the Western U.S. in particular, watershed hydrology is mostly strongly linked to snow, and in turn to vegetation. This has long been recognized and was the basis for a lot of work years ago at USFS experimental watersheds, and later pre-2000 (especially in the Pacific Northwest) motivated by concerns about effects of logging on flooding. The authors appear to be unfamiliar with that literature, and I'd urge them to take some time to bring themselves

up to speed. Then, I think they need to rethink their experimental design, so that the core of the paper is in fact linked to the title of the paper.

**Reply:** We agree that the original manuscript placed too much emphasis on attributing improvements in SWE directly to snow–vegetation interactions, while uncertainty in meteorological forcing also exerts a strong control on the simulations. We have substantially revised the manuscript to distinguish the effects associated with meteorological forcing from those associated with vegetation representation. Specifically, NLDAS-2 and AORC are now treated as separate forcing experiments, while the effect of vegetation representation is evaluated by comparing paired DynVeg\_Off and DynVeg\_Pred simulations within each forcing dataset. Because the meteorological forcing remains unchanged within each paired comparison, this provides a more direct assessment of SWE changes associated with differences in LAI/GVF representation. We therefore no longer interpret differences between the NLDAS-2 and AORC simulations as evidence of improved snow–vegetation processes.

To strengthen the station-scale analysis, we have revised Figures 5 and 6 so that the CDWR stations are now ordered according to the magnitude of the mean LAI difference ( $\Delta\text{LAI}$ ) between DynVeg\_Pred and DynVeg\_Off rather than by elevation. We additionally report the change in SWE performance at each station as  $R^2$  ( $\Delta R^2 = R^2_{\text{DynVeg\_Pred}} - R^2_{\text{DynVeg\_Off}}$ ) separately for NLDAS-2 and AORC. This analysis shows that the vegetation effect is spatially heterogeneous and forcing dependent. Under NLDAS-2, several stations with substantial LAI changes show improved SWE performance, whereas comparable LAI changes do not consistently lead to improvements under AORC. We have revised the discussion to explicitly acknowledge this behavior and to avoid implying that all differences between the simulations can be attributed solely to improved snow–vegetation interactions.

Here are the revised figures:

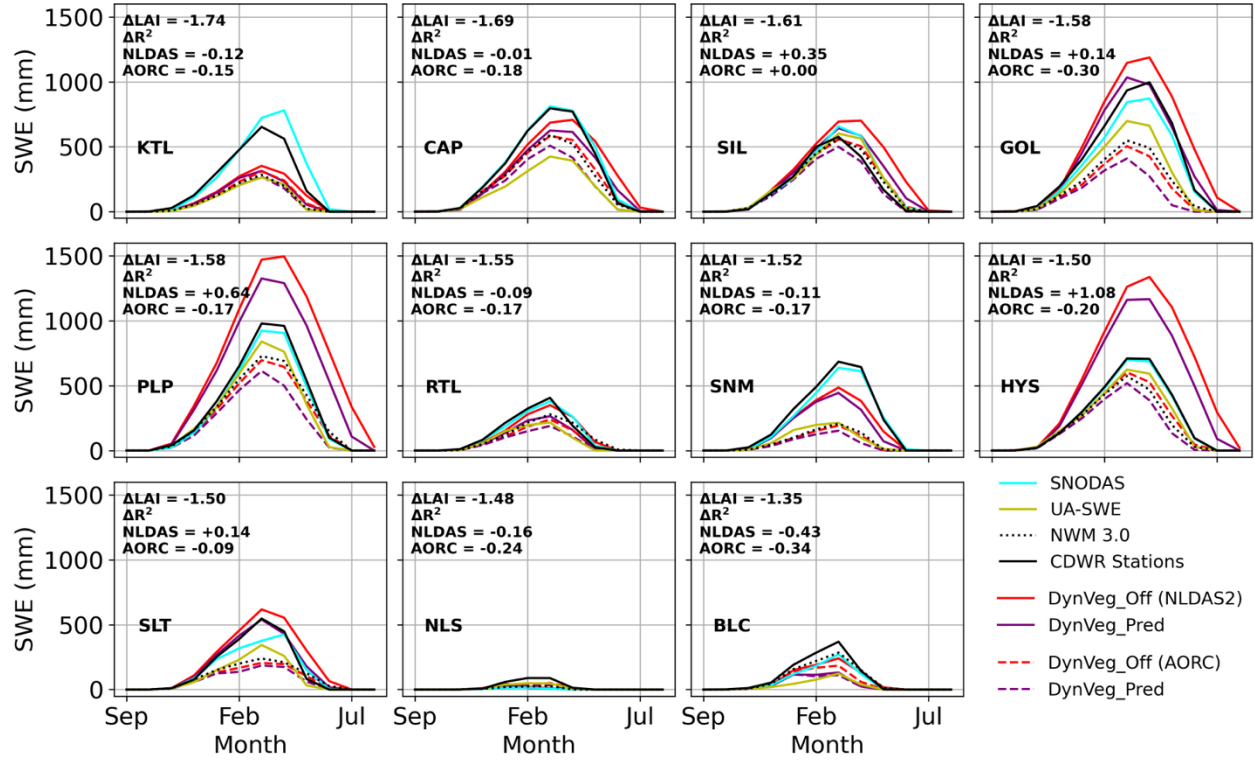


**Figure 5:** Snow water equivalent (SWE) comparison at CDWR stations in the Sacramento River Basin as represented in study area figure (Fig. 1). Monthly mean of control and dynamic vegetation modeled SWE with NLDAS-2 (solid lines) and AORC (dotted lines) forcing are plotted against the gridded observation (UA-SWE, SNODAS and NWM 3.0), and CDWR station (solid black line). The stations are sorted according to the difference in the LAI between both simulations, improvement in  $R^2$  is also reported for each plot.

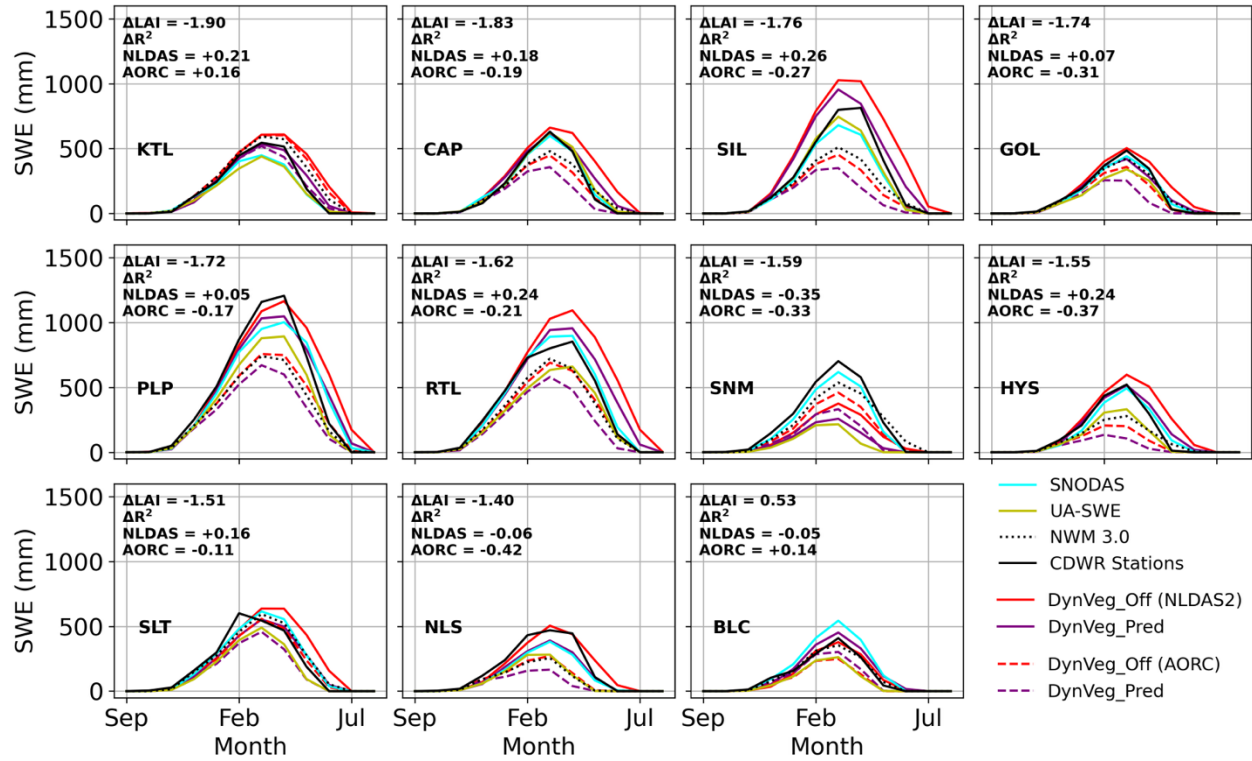


**Figure 6:** Snow water equivalent (SWE) comparison at CDWR stations in the San Joaquin River Basin as represented in study area figure (Fig. 1). Monthly mean of control and dynamic vegetation modeled SWE with NLDAS-2 (solid lines) and AORC (dotted lines) forcing are plotted against the gridded observation (UA-SWE, SNODAS and NWM 3.0), and CDWR station (solid black line). The stations are sorted according to the difference in the LAI between both simulations, improvement in  $R^2$  is also reported for each plot.

We have also incorporated your suggestion regarding canopy interception. In the revised Figure 8, canopy-intercepted snowfall is normalized by the corresponding snowfall forcing and is presented as the fraction of snowfall intercepted by the canopy. This normalization minimizes the direct influence of differences in snowfall magnitude between NLDAS-2 and AORC, allowing the

effects of vegetation representation on canopy interception to be compared more directly within and across the two forcing datasets. The revised Figure 8 is provided below:

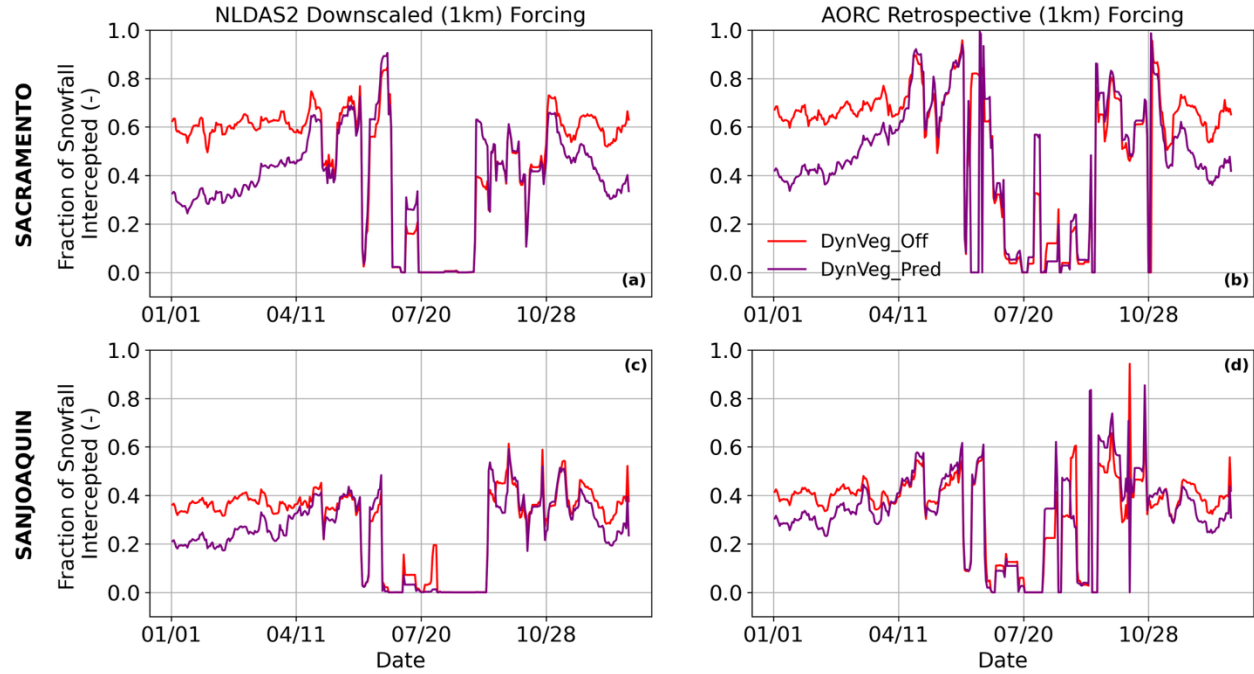


**Figure 8.** Modeled (normalised) canopy intercepted snow rate using different atmospheric forcing data (left panels: NLDAS-2; right panels: AORC) under prescribed vegetation (DynVeg\_Off) and dynamic vegetation (DynVeg\_Pred) in the Sacramento (upper two panels) and San Joaquin (lower two panels) River Basins. The intercepted snow was normalised using the snowfall from each forcing datasets. A SWE threshold of 0.1 mm has been applied to mask the region and perform spatial average.

Finally, we acknowledge that these paired experiments do not completely eliminate uncertainty in the meteorological forcing or provide a direct validation of sub-canopy snow processes. Rather, they allow us to isolate the model sensitivity to vegetation representation under a given forcing dataset. Direct process-level evaluation using collocated meteorological, vegetation, and sub-canopy snow observations would provide a stronger test of Noah-MP snow-vegetation processes and is identified as an important direction for future work.