Responses to Reviewer 2 for "Technical note: Precipitation phase partitioning at landscape-to-regional scales" by Lynn et al.

Reviewer comments are provided in normal text.
Responses are given in blue
Revised text given in italics ***(bold for emphasis)***

Review of "Technical note: Precipitation phase partitioning at landscape-to-regional scales" by Lynn et al. Submitted to Hydrology and Earth system sciences. 2020. General Comments/Overview This paper describes a new approach developed by the California Department of Water Resources to produce a long term (30+years), monthly, high-resolution (4km), rain/snow partitioning dataset over the Western US. The authors use this dataset/method to estimate long-term changes in rain/snow partitioning. With warmer temperatures, more precipitation is falling as rain rather than snow –which will impact snow water storage and water management practices. The authors argue that due to the paucity of snow observational datasets and the complex topography of the western US multiple datasets are needed to monitor and model hydrologic conditions over the Western US. Therefore they combine high-resolution PRISM precipitation data with coarse resolution freezing level and fractional snowfall calculations from NCEP/NCAR reanalysis to generate high-resolution fractional snowfall over California (and the Western US). While I believe this is a novel approach and one that has scientific merits, I have deep concerns about the use of the NCEP/NCAR 1 reanalysis product used in this study. In particular, the fact that precipitation from NCEP/NCAR is used to estimate the fraction of precipitation falling as snow. I also do not think the methods used in this paper are adequately described. As this is a technical paper designed to describe a method I believe this paper could be accepted following major revisions.

We appreciate the constructive comments provided by the reviewer and have majorly revised the paper in order to address their major, minor, and specific concerns. We look forward to submitting a revised manuscript to HESS.

Major Concerns:

1. The NCEP/NCAR reanalysis dataset used in this study is one of the oldest reanalysis products. At the time of its production/publication the authors (Kalany et al, 1996) state that "C" variables (such as precipitation) are completely determined by the model and should be used with caution. As the fraction of precipitation falling as snow is determined from precipitation in NCEP/NCAR reanalysis, I believe this will add significant uncertainties into the study. At a bare minimum this uncertainty/limitation needs to be discussed in section 5.2 "Primary Limitations" and making sure the reader knows this is a limitation of the study. However I suggest the authors consider performing a similar analysis with a new reanalysis product that adjusts model derived precipitation (e.g. MERRA2 or ERA5) and compare the results with NCEP/NCAR reanalysis.

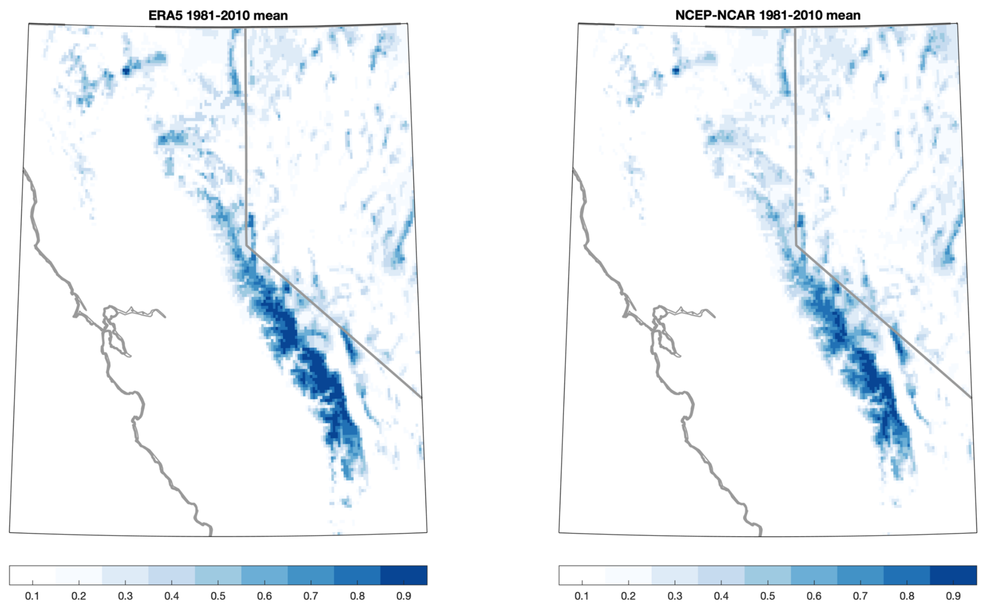
We appreciate the reviewer's concerns about the NCEP/NCAR reanalysis and thank them for requesting additional information regarding the uncertainties as well as the suggestion to perform a similar analysis with a modern reanalysis. Similar concerns with NCEP/NCAR were also brought up by Reviewer 1.

Following both reviewer's suggestions, we repeated the analysis with ERA-5 for the four aggregated DWR watersheds to provide some estimates of how well NCEP/NCAR performs. We found encouraging results (figure below, added as a supplementary figure), with ERA-5 and NCEP/NCAR being very well-correlated over the overlapping time period (correlations exceeding 0.9). ERA-5 was a bit colder (more %SNOW), which is likely related to a number of improvements in the ERA-5 model compared to NCEP/NCAR (data assimilation, spatial/vertical resolution, terrain, physical process representation). We added a paragraph to the limitations section highlighting our use of an older model (which was state-of-the-art at the time the NAFLT was developed in ~2008) and showing that it still performs relatively well. All in all, this comparison suggests that the method we are showing is valid (despite limitations in NCEP/NCAR) and thus the method represents a useful way to evaluate precipitation partitioning in models and distribute this partitioning across landscapes.

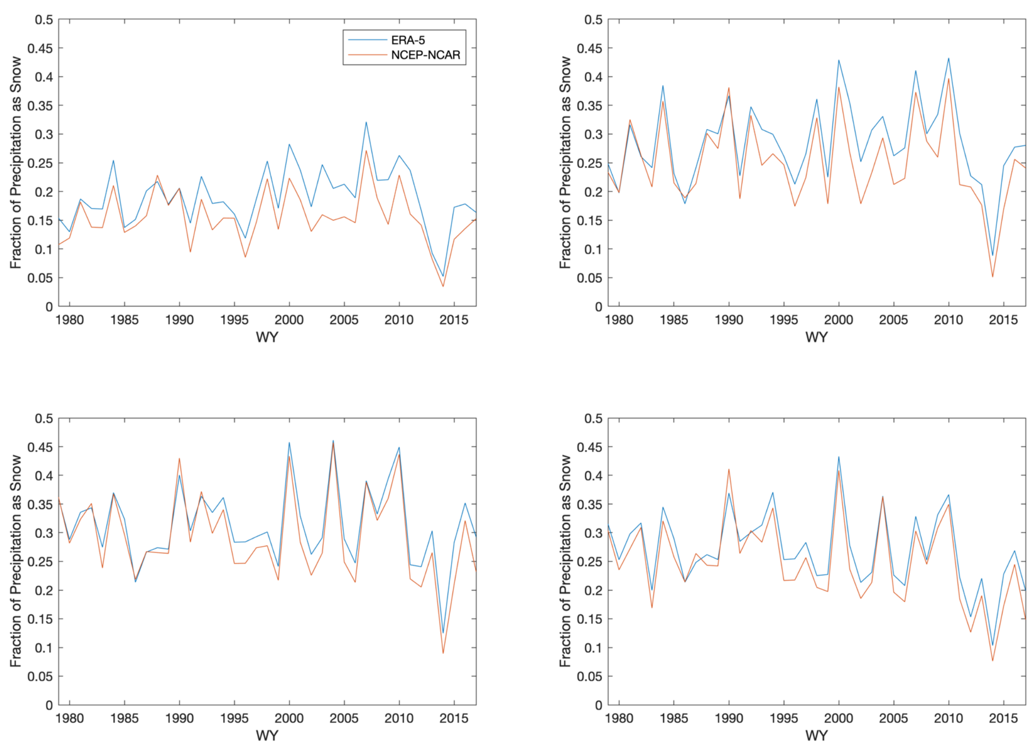
Added text:

"The NCEP/NCAR reanalysis, which the NAFLT uses to identify freezing levels and partition precipitation, is an older generation model. Recent advances in atmospheric reanalysis products such as ERA-5 (Hersbach et al., 2020) may provide additional benefits given their advances in data assimilation procedures, greater resolution in time and space (both horizontal and vertical), improved representation of terrain gradients and physical processes, and direct production of 0°C height as standard products. Comparisons of the NCEP/NCAR approach to ERA-5 showed that NCEP/NCAR performs reasonably well in terms of the mean spatial distribution (Supplementary Figure 3), with high correlations ($0.9<R<0.99$) of interannual $\%_{SNOW}$ variability between the models (Supplementary Figure 4). ERA-5 was consistently colder, with approximately 5% more $\%_{SNOW}$, supporting the notion that NCEP/NCAR is a conservative estimate of the freezing level. However, because it begins in 1979, ERA-5 precludes longer analyses periods. The method for partitioning precipitation described herein shows promise despite its use of an older reanlaysis model. As the NAFLT is updated, we anticipate including all available global reanalyses to provide an ensemble perspective of historic freezing level and precipitation partitioning."

New Supplementary Figures that will be added to the revised manuscript:



Supplementary Figure 3: Comparison of 1981-2010 mean water year fraction of precipitation falling as snow (multiply by 100 to yield $\%_{SNOW}$*) for northern California and western Nevada produced using ERA-5 (left) with NCEP-NCAR (right).*



Supplementary Figure 4: Comparison of fraction of precipitation falling as snow for ERA-5 (blue line) and NCEP-NCAR (red line) for the period 1979-2018 for the four DWR analysis zones, ordered clockwise from upper left: Southern Cascades, Northern Sierra Nevada, Central Sierra Nevada, and Southern Sierra Nevada.

2.PRISM also provides daily precipitation and surface temperature at 4km resolution. One could estimate daily snowfall using a surface temperature threshold (as in the UA University of Arizona 4km SWE product estimations (https://nsidc.org/data/nsidc-0719; Broxton et al, 2019 and as is done in many land-surface models). The authors need to better explain the science behind why it is more useful or credible to use coarse freezing level data from NCEP/NCAR reanalysis (or another reanalysis product in 1.) to estimate snowfall percentages, than directly calculating these from PRSIM data. (I don't disagree that this surface temperature is not a great indicator of snow level, its just not explained/justified in the paper). An interesting comparison could be to look at snowfall estimated from daily T/P vs their method of approximating rain/snow partitioning.

We appreciate the reviewer's request for better justification about the use of freezing level data instead of surface-based temperatures. There are a number of issues with surface temperatures (data sparseness especially in complex terrain, inadequacies in resolving near-surface lapse rates (e.g., Lute and Abatzoglou, 2020)). These errors can be translated into significant errors in snow models. Another issue with daily-based gridded products are the daily time step that may miss the dynamics of change in snow level throughout the day (such as abrupt rises or falls with frontal passage). The 6-hourly approach with reanalysis may help address this issue somewhat. As a synoptic scale phenomenon, the freezing level is generally well-resolved by models (admittedly there can be substantial variation at finer scales due to microphysics/latent heating/hydrometeor dragging). We added additional text to the introduction to further motivate the study and the use of freezing elevations. We will continue thinking about this issue as we revise the paper, as this is a very important consideration to incorporate into the manuscript well.

New text:
"*Daily gridded products based on sparse observational networks in mountainous areas have their own suite of limitations, such as capturing sub-daily fluctuations in temperature or resolving lapse rates (Lute and Abatzoglou 2020).*"

The purpose of this technical note is to describe the development of this diagnostic indicator aimed at quantifying how rain and snow are partitioned ***based upon the elevation of the atmospheric freezing (0°C) isotherm, which has been found to be well-resolved by global models in complex terrain (Abatzoglou 2011).***"

Added reference:
Lute, AC, Abatzoglou, JT. Best practices for estimating near-surface air temperature lapse rates. *Int J Climatol*. 2020; 1– 16. https://doi.org/10.1002/joc.6668

We agree it would be interesting to compare this approach to a surface-based approach and appreciate the suggestion. While beyond the scope of this methods paper to dive into comprehensive comparisons, we have added text to encourage this type of study using snow reanalyses and surface based data (especially if the surface data includes RH or other necessary parameters to calculate wet bulb temperatures):

“Both sources of uncertainty may result in substantial biases in rain/snow partitioning estimates as a function of individual storms, particularly during frontal passage and when the magnitude and spatial distribution of precipitation is also considered. ***To begin to address these limitations, other high temporal (daily) and spatial resolution (>6 km) snow datasets can be utilized in future work, such as snow reanalyses that incorporate satellite and/or in situ data (e.g., Margulis et al., 2016; Zeng et al. 2018). Further, comparisons with approaches that include relative humidity or wet bulb temperatures are recommended to further improve the methodology, as these have been shown to improve the quality of rain-snow partitioning (Harpold et al., 2017, Wang et al., 2019).”***

3.The sections describing the NAFLT methods and the DWR approach to rain/snow partitioning are not clear.

We appreciate the request for improving our description, and provide specific responses below.

It seems to me the lowest to the ground freezing level would matter most for snowfall and surface conditions. Why then does the NAFLT method use the uppermost level in areas where there may be a temperature inversion? Please justify this method.

Perhaps surprisingly, the surface temperature (what I am interpreting by ‘ground freezing level’), does not have the ‘final say’ in precipitation phase. A multitude of physical processes, such as latent heating, hydrometeor fall speed, and others control the melting of a snowflake (see Minder et al., 2011 and Jennings et al., 2018 for further descriptions of processes; note the other reviewer suggested the additional Jennings reference). This means that the freezing level elevation is a maximum estimate of where snow may turn to rain, but it is often well below (100’s of meters) that elevation, hence why snow can be experienced at surface temperatures above freezing. Further, soundings may indicate freezing rain in overrunning types of situations (cold air pooling at the surface and being overrun by warmer air; this would produce an inversion in a sounding), which is not snow. We are avoiding this situation by following standard NWP definitions of freezing level. We approached this issue in the discussion, but we agree with the reviewer that some up-front justification would help readers when introducing the method.

With regards to the inversion issue, our original text used the standard NWP calculation (also used in ERA-5 but we slightly revised it for clarity: “The uppermost atmospheric level below which the 0°C isotherm occurs is considered for cases in which the vertical temperature profile includes inversion conditions *with* multiple incursions of the 0°C isotherm”

We have revised the text to improve the clarity of the description of the method:
“*The* NAFLT calculates the freezing level as the ***highest elevation in the troposphere (200-1000 hPa)*** above mean sea level where ***free-air temperatures are 0°C.*** If the entire atmosphere is at or below freezing on a given 6-hr period, a value of zero meters above mean sea level is provided. The uppermost atmospheric level below which the 0°C isotherm occurs is considered for cases in which the vertical temperature profile includes inversion conditions ***with*** multiple incursions of the 0°C isotherm. In addition to providing estimates of the elevation of the 0°C isotherm, the

NAFLT calculates the ***monthly*** percent of precipitation that falls ***as snow (%SNOW) at 200 m elevational increments from 0-4000 m. This is done by assigning all 6-hourly modelled precipitation from the NCEP/NCAR reanalysis as snow for elevations above the corresponding freezing level and all precipitation in a 6-hour increment as rain for elevations below the freezing level.*** The freezing level is a very conservative estimate of the snow level as precipitation can often persist as snow below the freezing elevation due to latent heat fluxes (e.g., snow falling in a sub-saturated atmosphere, deep isothermal temperature profiles, or during heavy precipitation episodes that entrains colder air to lower levels in the atmosphere; ***Minder et al., 2011; Jennings et al., 2018***). However, accumulations of snow below the elevation of the 0°C isotherm may be transient due to nominal cold content of snow."

We also revised the DWR approach text in section 3.1:
The DWR approach calculates $\%_{SNOW}$ by first bilinearly interpolating of $\%_{SNOW}$ horizontally for each 200 m elevational increment from NAFLT and then assigning $\%_{SNOW}$ to each fine-scale grid point per the smallest elevational difference between fine-scale elevation (e.g., 4-km DEM) and the 200 m elevational bins. If the freezing level elevation is below the terrain elevation, all precipitation falls as snow ($\%_{SNOW}$ = 100). Given the known inadequacies of coarse-scale reanalysis precipitation fields in complex terrain, we multiplied estimates of monthly PRISM precipitation by $\%_{SNOW}$ to partition precipitation between total frozen ($\%_{SNOW}$) and liquid ($\%_{RAIN}$) components similar to Abatzoglou (2011).

It needs to be made clear that with this method –the freezing level can be below the surface topography.

Thank you for the suggestion, we added a sentence to make this clear:
"If the freezing level is below the terrain, all precipitation falls as snow ($\%_{SNOW}$ = 100)."

Is the freezing level calculated independently for each 2.5º grid box?

Correct, we revised and added text to reflect this:
"*free-air temperatures are 0°C **for each 2.5° NCEP/NCAR grid point***"

"The DWR approach calculates $\%_{SNOW}$ by first bilinearly interpolating the 2.5° grid point estimates of $\%_{SNOW}$ horizontally for each 200 m elevational increment from the NAFLT. Then it assigns $\%_{SNOW}$ to each fine-scale PRISM grid point per the smallest elevational difference between fine-scale elevation (e.g., 4 km DEM) and the 200 m elevational bins."

Most critical: This statement on Page 3 Line 20 "percent of precipitation that falls at elevations above the 0ºC isotherm at 200m increments from 0-4000m" does not make sense to me. Why wouldn't all elevations above the 0º isotherm also be below freezing, and therefore all precipitation would fall as snow? Is there an equation being applied to estimate the fraction of precip falling as snow?

This sentence has been revised for clarity:

"*This is done by assigning all 6-hourly modeled precipitation from NCEP/NCAR* ***reanalysis as snow for elevations above the corresponding freezing level***"

Also, does the method start at the freezing level elevation and work up from there (is the reference point for the 0-4000m) then from the freezing level. Or does it start from the surface? Does the method really estimate the % of snowfall below the 0º isotherm (where you could have mixed precipitation). Similar language is used in section 3.1 to describe how you apply this method to the high-resolution precip data from prism (Page 3 Lines 29-31).

Per standard approaches with Numerical Weather Prediction models and freezing-level outputs that are now part of modern reanalyses (e.g., ERA-5), the 0C level is defined as the highest atmospheric level in the troposphere where the 0C level is crossed. This is done to avoid inversions in the case of freezing rain/sleet that might otherwise assign the precipitation type as snow. We further clarified that the approach used in binary in the sense that for a given 6-hour block elevations above the freezing level are assigned 100% snow, and those below 0% snow. Please see major revisions to the text provided above.

In the Primarily Limitations section you state the assumption was made that %snow linearly relates to the NAFTL –but that was not actually stated in the discussion of the methods, it needs to be. My suggestion is to think about how to describe this method to someone who does not know what the freezing level is, or how %snow is calculated and really step through the process–if this is too much detail you might put some of this in supplemental (a diagram could be helpful as well!).

We appreciate the suggestions to describe the method in a step-by-step manner and will include a schematic diagram in the revised manuscript. This will likely help a number of readers understand the method better (especially those more visually inclined). We will update the response to reviewer comments once this diagram has been completed and included.

Minor Concerns

One thing that is missing from the introduction and could be a nice addition is an understanding of how this indicator could actually be used for water resource planning. This is touched upon in the Discussion (Page 5), but I think some of this type of information needs to be included in the introduction to further motivate the need for this method. You do say that changes in rain-snow partition are important for water storage –however it wasn't until the Discussion that I could really see why this indicator might actually be used for planning purposes.

This is a valuable suggestion, especially for a methods/technical note type paper intended to help other practitioners/managers as well as the science community. We have revised and added text to the introduction to better describe some of the specific ways that DWR is applying this information. For the most part, they use it as an indicator of change for situational awareness of where expected impacts of climate change are occurring (and how fast). We also added a final sentence noting the approach can be used to look forward as well (longer-term planning).

New text:
"***Since 2015, DWR has documented this indicator in its annual Hydroclimate Report (DWR, 2019). Though not used directly in operational forecasts, the indicator provides DWR with situational awareness regarding the location of changes in precipitation phase and the rates of these changes.*** Because the method uses publicly available gridded data sets, the indicator is scalable to regional-to-continental scales and therefore could be an informative diagnostic for water resources management and model development in snowmelt dependent regions. While we focus on California watersheds, an example application to the western United States is provided. ***Last, instead of historic data, it can also use model projections as input to help inform the development of adaptation strategies to achieve water resource management goals amidst a changing climate.****"*

Page 1, Line 23: "components has been used as a foundation"

Thank you for suggesting a fix, in light of other reviewer comments we revised the sentence as follows:

"*The partitioning of precipitation into liquid (rain) and, in particular, frozen (snow) components along with climatic stationarity were foundational assumptions in the development of water management infrastructure and practices in California and other mountain environments in the western United States (US) since the mid-1800s (Milly et al., 2008).*"

Page 1, Line 27: The use of the word "fate" is a little awkward here, do you mean "phase"?

We agree 'fate' is awkward, though the original intent was a meaning of fate as "destiny", where the destiny of cool season precip is either rain that runs off or snow that accumulates. We changed to use the suggestion of "phase" as this is less awkward and is correct, as the phase does ultimately drive the management strategy.

New text:
"The ***phase*** of cool season precipitation ultimately drives"

Page 1, Line 36: Unclear what you mean by "winter snow levels" here. Do you mean the freezing level in the atmosphere or do you mean increase snow pack (which would be counter intuitive). The jargon of 'winter snow levels" is confusing.

We understand the reviewer's confusion and appreciate them pointing this out as they are not the first to be confused by the nomenclature. We changed the text to winter snow line elevation (which is close to the freezing level, but usually several hundred meters below due to melting times/other processes that influence melting rates).

New text: "winter snow ***line elevation***"

Page 2, Lines 9-11: The first sentence of this paragraph is not complete. It is unclear what are you incorporating multiple data sources and model outputs into.

We appreciate the suggestion to correct the structure of this sentence. We have revised it accordingly to be clearer that the problem is little data availability to apply hydrologic models or to evaluate change. Incorporation of multiple data sources or other model output can often provide the necessary pieces to perform the study of interest.

New text:
"The sparse observational networks and complex topography of the western US introduces challenges into basin-scale hydrologic monitoring and modelling***. To address these challenges when applying hydrologic models or for monitoring long-term change,*** the incorporation of multiple sources of data (Bales et al., 2006) and/or model output (e.g., Wrzesian et al., 2019) ***is often required.***"

You mention on Page 2, Line 14 that DWR developed a methodology to study historical rain/snow trends at fine spatial resolutions and then on Page 2, Line 16 that the purpose of the note is to provide an updated approach and detail the methods of this indicator. It is unclear to me what in this paper is from the original DWR method and what is the "updated" approach. Is the only difference between the DWR method and the method described in the paper the resolution of the PRISM model data? If the goal of the paper is just to outline the DWR approach –then state that and remove the confusing "updated approach" language. However if the DWR approach is documented elsewhere, and you are documenting changes here in this paper, those differences need to be more explicitly stated.

We apologize for the confusing language. This document is intended to provide the first peer-reviewed documentation of the approach originally described in the 2014 DWR report. We have removed 'updated' to avoid confusion.

New text:
"The purpose of this technical note is to *describe the development of this diagnostic indicator aimed at quantifying how rain and snow are partitioned.*"

The reviewer is correct that the only difference described here is with regards to the PRISM spatial resolution. Because this is already described in the methods (section 2.2; the DWR approach uses 800 m but because of data accessibility (free vs. pay) we use the 4km product), we chose not to further explain this difference in the introduction.

Page 2, Line 16: You say "detail the methods of this indicator" but at this point in the paper it is not clear what "indicator" you are talking about. A sentence about the "indicator" before this one is needed. (e.g. Page 6, Line 4 could also be stated here in the text).

Thank you for the request for additional clarity. We revised the text as follows:
"…indicator ***of how rain and snow are partitioned***."

Page 2, Line 19: "and be an important" a modal verb is need before "be" as in "may be" or "can be" etc.
We appreciate the grammatical correction (and example!). The text now reads:
"and ***therefore could*** be an *informative*"

Page 3, Line 20: ...the percent of precipitation that falls (as snow??) at elevations above the 0ºisotherm ...
Thank you for pointing out our omission. Text has been changed:
"…falls ***as snow*** at…"

2.1 the Study Area Did DWR create this method exclusively for studying trends in rain/snow partitioning, or is this data used in operational forecasts?

Correct, DWR did develop the method exclusively for studying trends. We revised to make it clear in the introduction that the indicator is not used operationally but to inform about trends:
"***Though not used directly in operational forecasts***, *the indicator provides DWR with situational awareness regarding the location of changes in precipitation phase and the rates of these changes.*"

Figure 4 has a number of problems:
What season is being plotted? Entire water year? Cold season etc?

Thank you for pointing out this omission. These are water year plots. The text has been changed to specify this:

"Figure 4: (a) Aggregated trends in $\%_{SNOW}$ (% decade^{-1}) by latitude and elevation ***for the water year.*** Dot size is scaled by area of watershed occupying each elevation and latitude bin. (b) Elevation-based trends (aggregated over all latitudes) of $\%_{SNOW}$ (% decade^{-1}) showing median (black line), the interquartile range (dark grey shading), and 90% confidence intervals (light grey shading) on the left y-axis. Right y-axis shows the total watershed area occupied by each elevation bin (red line; km^2). Aggregations were performed on gridpoints within the subset of California Department of Water Resources analysis zones (see Figure 1a)."

You discuss Figure 4b before 4a, they should be flipped in the panel.
Thank you for pointing this out, we flipped the panels (and adjusted the caption order as well). New figure and caption:

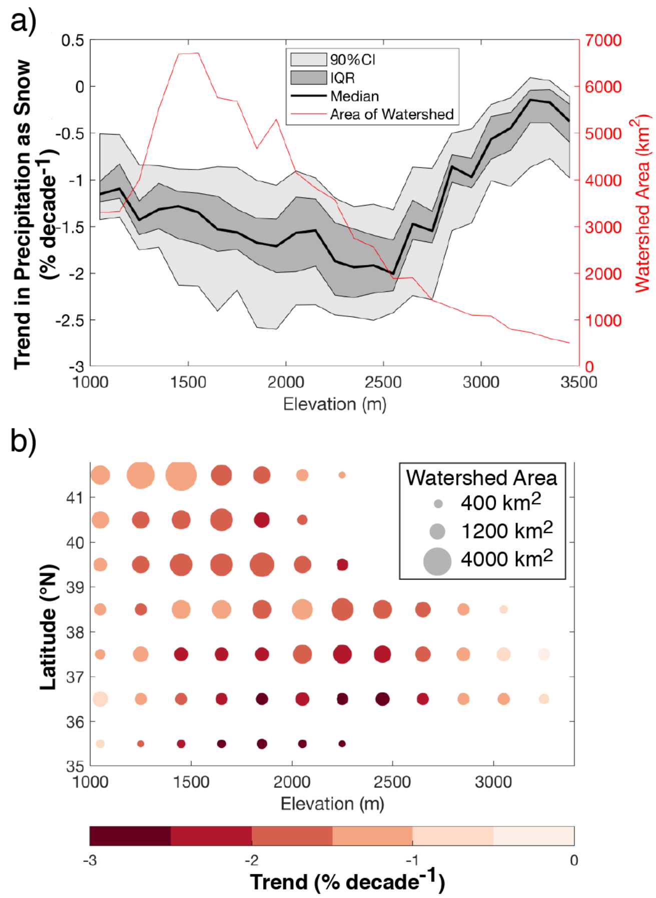


Figure 4: ***(a) Elevation-based trends (aggregated over all latitudes) of $\%_{SNOW}$ (% decade^{-1}) showing median (black line), the interquartile range (dark grey shading), and 90% confidence intervals (light grey shading) on the left y-axis. Right y-axis shows the total watershed area occupied by each elevation bin (red line; km^2). (b) Aggregated trends in $\%_{SNOW}$ (% decade^{-1}) by latitude and elevation for the water year. Dot size is scaled by area of watershed occupying each elevation and latitude bin.*** Aggregations were performed on gridpoints within the subset of California Department of Water Resources analysis zones (see Figure 1a) and sorted by elevation. The interquartile range (IQR) and 90% confidence interval (CI) were estimated using all grid points within each elevation band and analysis zone.

It is unclear from the text how Figure 4b is calculated –what is the IQR and 90% CI based on (is this covering every grid point within that elevation band?).

Correct, and we apologize for the oversight to include this detail in our original submission. We added a sentence to the caption to describe how the IQR and CI were calculated (please also see complete revised caption above):

"The interquartile range (IQR) and 90% confidence interval (CI) were estimated using all grid points within each elevation band and analysis zone."

Does Figure A represent the values shown in Figure 3 but sorted by elevation?

Yes, the difference being values were aggregated by latitude and elevation and not by watershed. We added a note in the Figure 4 caption that the aggregations were *"sorted by elevation"*.

Page 5, Line 37 –what is a flood pool? This should be stated in a way that non-flood forecasters/water managers can understand.

Good suggestion. We added the definition and some text clarifying why it matters. The flood pool exists to prevent downstream flooding when inflows are high (such as during storms). The states that a flood pool must be maintained during certain parts of the year, meaning that inflows into the flood pool must be released as soon as possible. Instead of water being stored upstream in the snowpack to flow into the reservoir in July (a resource), now this water can be lost for later consumptive use as it flows downstream in February (managed as a hazard).

New text:
"More precipitation falling as rain during storms, especially in regions with large watershed areas in lower elevations, increases *midwinter* inflow into reservoirs. Many current multipurpose reservoir management paradigms require the maintenance of a flood pool, ***which is reservoir storage space allocated to attenuate periods of heavy inflow and reduce flood hazard during cool season storms. Water captured during the flood is later released to maintain the flood pool storage capabilities during the next possible event. Flood pool releases*** mean this water cannot be stored for later beneficial use and must be managed as a hazard rather than a resource."

Page 6, Line 4 –this description of the goal of this paper needs to be moved up into the introduction.

Thank you for the suggestion. We moved the sentence to the introduction.