Response to Reviewer Comments

Dear Editor,

Thank you for providing valuable feedback from the reviewers on our manuscript "High-Resolution Snow Water Equivalent Estimation: A Data-Driven Method for Localized Downscaling of Climate Data." We have carefully addressed all of the comments and suggestions to improve the clarity and overall quality of the paper. Below, we provide a point-by-point response to the reviewers' comments.

Reviewer 1

1. Comment: The manuscript is generally well-structured, with a clear flow from the introduction to the methodology, results, and discussion. However, some sections could benefit from further clarification, particularly where the methodology is complex.

Consider adding a diagram or flowchart in the methodology section to visually represent the steps of the proposed downscaling algorithm. This would help readers better understand the process.

1. Response: We would like to thank the reviewer for their constructive feedback. We agree and have reviewed the methodology section, making several adjustments to enhance clarity. Additionally, we have included a flowchart in the methodology section (Section 2.3) to visually represent the steps of the downscaling algorithm, as suggested.

Lines 173-174:

SWE is also affected by preceding meteorological conditions, such as the temperature and precipitation patterns of previous days. For instance, the amount of SWE today may vary depending on the conditions experienced in the preceding days.

Lines 197-205:

In this study, the k-NN algorithm is applied to downscale low-resolution climate data to HR-SWE estimates by selecting learning days with similar climate conditions. The flowchart (Fig. 2) illustrates the proposed downscaling method for estimating HR-SWE, and the "steps" are as follows:

1. Gather the input variables, including the far and near intervals of temperature, precipitation, shortwave radiation, and the LR-SWE for both the target date and the training dates.
2. Calculate the distance between the input vector of the target date and the input vectors of the training dates.
3. Select the K-nearest training dates based on their proximity to the target date in the input space.
4. Retrieve the corresponding HR-SWE images associated with the selected K-nearest training dates.
5. Aggregate the retrieved HR-SWE images to estimate $\widehat{SWE}$ for the target date.

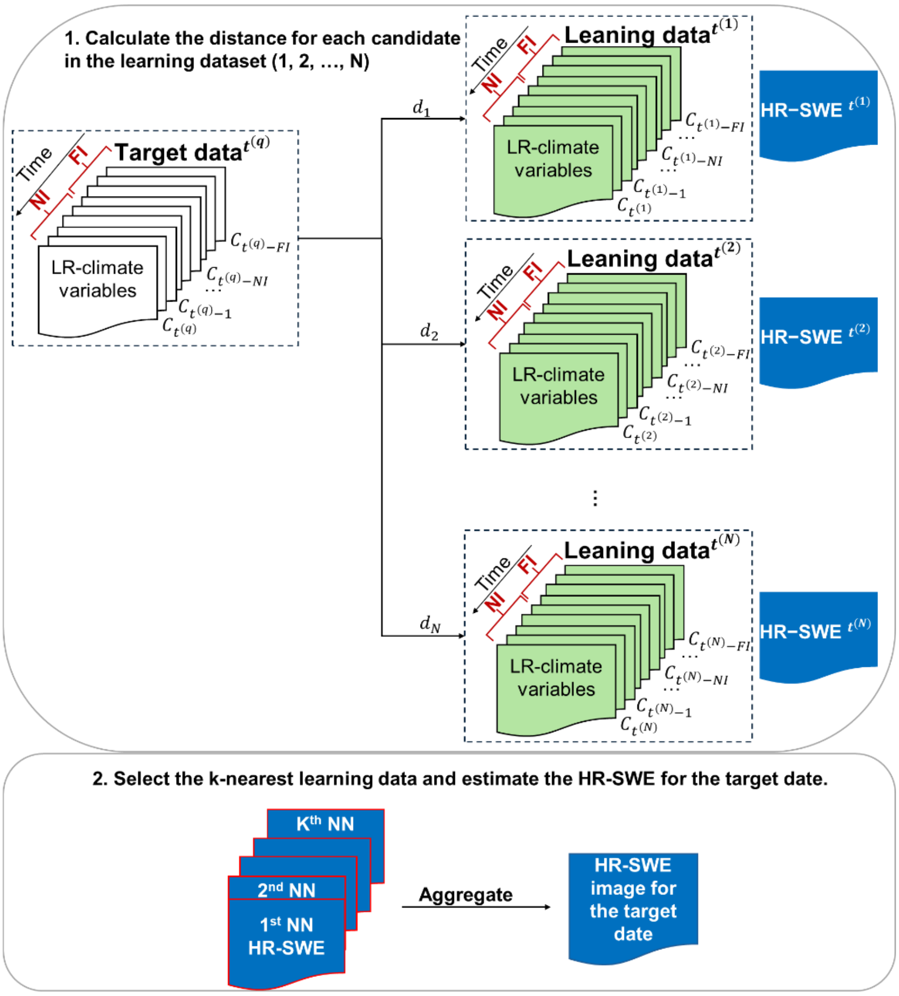


Fig. 2. Flowchart depicting the proposed downscaling algorithm for estimating high-resolution Snow Water Equivalent (HR-SWE) on a target date using low-resolution (LR) climate data: 1. Gather Input Variables: Collect the input variables, including Near (NI) and Far (FI) temporal intervals of climate data (e.g., temperature, precipitation) for both the target date and the training dataset. 2. Calculate Distances: Compute the distance (d) between the input vector of the target date and the input vectors of each candidate in the training dataset using a defined distance metric (Manhattan distance). 3. Identify k-Nearest Neighbors (k-NNs): Rank the training data in order of increasing distances to identify the k-nearest neighbors. 4. Aggregate HR-SWE Images: Retrieve the HR-SWE images corresponding to the k-NNs and aggregate them to estimate the HR-SWE for the target date.

Lines 210-213:

Unlike the Euclidean distance, which calculates the shortest straight-line distance, the Manhattan distance computes the sum of the absolute differences between variables. This makes it more robust against outliers and better suited to handling high-dimensional datasets, such as those containing multiple climate variables. In our method, this distance is used to rank the training dates based on their similarity to the target date.

2. Comment: The novelty of the approach is well articulated, but a more explicit statement of how this method advances the field compared to existing approaches would strengthen the introduction.

2. Response: We agree with this suggestion and have now explicitly highlighted how our method advances existing approaches, as follows:

Lines 79-114:

One of the statistical downscaling methods is bias-correction spatial disaggregation (BCSD) (Wood et al., 2004), which effectively reduces uncertainties in climate model outputs by adjusting biases based on high-resolution observational data. These methods are particularly useful in non-mountainous regions, where data availability is typically higher. Their primary strength lies in their ability to correct model outputs while capturing local variability. However, their reliance on high-quality in-situ data significantly restricts their applicability in remote or data-scarce areas. In contrast, our method overcomes these limitations by utilizing low-resolution climate data without requiring ground-based observations, making it well-suited for a wider range of conditions, including regions with limited data availability.

Another widely used statistical downscaling method in climatology is based on a pattern known as the analog method (Zorita and Von Storch, 1999). These methods identify patterns in historical data that closely match the patterns simulated by atmosphere-ocean general circulation models. The observed surface climate conditions corresponding to these historical matches are then used as downscaled predictions. Analog methods have seen extensive application, as highlighted in studies such as those by Abatzoglou and Brown (2012), who demonstrated their effectiveness in wildfire assessments through the multivariate adapted constructed analog, and outperformed traditional spatial downscaling methods. Similarly, Pons et al. (2010) utilized analog-based downscaling to analyze snow trends in Northern Spain, successfully replicating observed variability and trends for seasonal and climate change projections. The study by Caillouet et al. (2016) demonstrated the utility of probabilistic downscaling in reconstructing high-resolution precipitation and temperature fields over France, effectively addressing seasonal biases. The AtmoSwing software by Horton (2019) demonstrates the flexibility of analog methods for operational forecasting and climate impact studies.

Additionally, an analog-type method, named nearest neighbor resampling, also used by Lall and Sharma (1996), relies on identifying patterns in historical point-based time series data and resampling them using a nearest-neighbor approach to preserve the serial dependence structure of the data. The k-nearest neighbors (k-NN) algorithm is a simple, non-parametric machine learning technique commonly used for classification and regression (Cover and Hart, 1967). It works by identifying the 'k' most similar data points (neighbors) to a target data point based on a chosen distance metric, such as the Manhattan distance. The k-NN downscaling by Gangopadhyay et al. (2005) extends the analog method by weighting several similar historical analogs to create predictive ensembles, adding further flexibility to this approach. Building on this concept, Rajagopalan and Lall (1999) extended the methodology to multivariate weather simulations, incorporating variables such as precipitation, temperature, and wind speed to simulate daily weather sequences. Later, Yates et al. (2003) used an adapted version of these methods to generate daily weather sequences and alternative climate scenarios. In the weather generator models based on k-NN, in general, the day directly succeeding the identified analog day is selected as the next day in the generated sequence, and this process continues iteratively (Gangopadhyay et al., 2005). Similarly, recent advancements, such as the study by Yiou and Déandréis (2019), have extended analog methods to ensemble-based probabilistic forecasts, demonstrating skill in predicting variables like the NAO index and temperature measurements at European stations. These innovations highlight the adaptability and growing utility of analog-based and k-NN approaches in climate and hydrological modeling.

Lines 121-136:

This method offers several key advancements over existing k-NN downscaling techniques. First, the adaptation made to the k-NN downscaling method, specifically by introducing far and near temporal intervals of climate data, is highly flexible to dynamic variables undergoing significant changes due to climate variability. Second, unlike most analog methods that restrict analog candidates to a specific temporal window near the query date, this approach does not impose such limitations. This flexibility is crucial for three reasons: (1) the inclusion of far and near temporal intervals makes such restrictions unnecessary, as the most suitable candidates are selected based on their match within the temporal window; (2) it is essential for preserving extreme events, as restricting candidates to a narrow date range risks losing matches that represent rare but important extreme events; and (3) it enables downscaling for future periods where exact analogs may not exist in the historical record within a specific date range. However, suitable analogs may still be found in historical observations but on different dates. For example, with climate change, a specific snow day in winter may no longer match the query day, but an analog might be found in another season, such as fall or spring, due to warmer climatic conditions.

Third, the method can reconstruct HR-SWE data for historical periods where only low-resolution climate data is available, providing valuable insights into past snow conditions. Additionally, the method excels at capturing fine-scale SWE patterns in complex terrains, such as mountainous regions, significantly improving upon traditional statistical downscaling models that often struggle in such environments. Finally, our method does not require high-resolution climate input data. This substantially reduces computational demands while maintaining the high quality of the downscaled SWE data.

3. Comment: The methodology is thorough, but some sections are densely packed with technical details, making it challenging to follow. The description of the Manhattan distance metric, in particular, could be expanded to enhance reader understanding. Explain how the Manhattan distance operates within the context of your downscaling approach. This will help make the concept more accessible, especially for readers less familiar with the metric.

3. Response: We have expanded the explanation of the Manhattan distance metric in Section 2.3 to provide greater clarity. Specifically, we explain how this metric is used to rank dates in the training dataset and why it is preferred over the Euclidean distance for this specific problem. We also added more details and made several adjustments to enhance clarity (please refer to Response 1).

4. Comment: The choice of parameters (e.g., FI and NI intervals) is justified through sensitivity analysis, which is appropriate. However, the manuscript could benefit from a brief discussion on the potential limitations or assumptions made in the parameter selection process.

4. Response: We have added a discussion in Section 2.4.1 to highlight potential limitations related to the choice of parameters, as follows:

Lines 240-246:

The parameters are selected through a sensitivity analysis to minimize ε, but they do not necessarily correspond a global optimum. This means that while they perform well in the studied regions, they may not generalize well to other locations or climatic conditions. Therefore, we recommend performing a sensitivity analysis for each region. Additionally, the sensitivity analysis assumes that the influence of these intervals is consistent across different temporal scales, which may not always be valid, particularly in regions with highly variable climate patterns. Despite these assumptions, the chosen parameters strike a balance between computational efficiency and accuracy for the downscaling task. Moreover, the subsequent weight optimization further mitigates the impact of non-global optimal parameter selection.

5. Comment: The results are comprehensive and well-presented with relevant figures and tables. However, the discussion on the comparison between different models and resolutions could be expanded to provide more insights into why certain models perform better in specific regions.

5. Response: We agree and have now referenced Kouki et al. (2022), who analyzed CMIP6 models and found significant variations in model performance based on regional climatic and geographical conditions. Additionally, we have expanded the discussion to include the contributions of temperature and precipitation to SWE biases. Specifically, we emphasize that precipitation plays a dominant role in winter SWE biases, while temperature becomes more influential during the spring snowmelt season. These insights have been added to Section 5 to highlight important factors that affect model performance, although they were not the central focus of this study.

Lines 441-450:

Several factors explain why certain models outperform others in specific locations. While a detailed comparison between model performances across various regions is beyond the scope of this study, other studies have explored this area. For example, Kouki et al. (2022) evaluated the ability of CMIP6 models to estimate SWE across the Northern Hemisphere and found that different models perform variably in specific regions based on their ability to simulate particular climatic and geographical conditions. In terms of the contribution of temperature and precipitation to SWE biases, precipitation plays a more dominant role, especially in winter. However, temperature becomes more significant during spring, when snowmelt occurs. In regions where temperatures are closer to 0°C, biases in temperature can substantially affect snowmelt. Overall, the results underscore that precipitation is the primary driver of SWE biases in winter. However, temperature plays a crucial role during the snowmelt season in spring, particularly in regions with more temperate climates, such as the southern parts of North America and Europe.

Discussion Section, Lines 537-540:

Accordingly, achieving accurate HR-SWE estimation relies significantly on the choice and accuracy of the climate model inputs, such as precipitation and temperature data, which may introduce biases into the SWE estimates. For example, precipitation biases are a dominant factor influencing SWE estimation errors, while temperature biases become more significant during transitional periods, such as the spring melt season.

6. Comment: Consider including a more detailed comparison of the proposed method with existing downscaling techniques. This could be done in the discussion section, where you could highlight the strengths and limitations of your approach relative to others.

6. Response: We agree and have now added more detail in the Discussion section, lines 542–568, as follows:

The following comparison provides a broader perspective on how our proposed method compares with other statistical downscaling techniques. BCSD methods are effective in reducing uncertainties in climate model outputs by adjusting model biases using high-resolution observations. These methods are particularly valuable for ensuring that model outputs align with observed climatology and capture local variability. However, they depend heavily on the availability of high-quality in-situ data, which limits their application in remote or data-sparse regions. In contrast, our method excels in areas with sparse data, as it uses low-resolution climate data without requiring ground observations, making it adaptable to a broader range of conditions.

Analog-type statistical downscaling approaches offer a simple, computationally efficient way to project high-resolution data based on historical relationships between large-scale climate patterns and local climate variables. These methods are useful in regions where historical climate patterns are stable and well-documented. On the other hand, our method introduces several key improvements over traditional k-NN downscaling techniques. First, the adaptation of the k-NN approach through the incorporation of far and near temporal intervals of climate data enhances its ability to handle dynamic variables, such as snow, which are subject to significant changes due to climate variability. Unlike conventional analog methods, which constrain analog candidates to a specific temporal window near the query date, this method eliminates such restrictions.

This flexibility is crucial for three reasons: (1) the inclusion of far and near temporal intervals allows the selection of the most suitable candidates across a broader temporal range, eliminating the need for narrow constraints; (2) it ensures the preservation of extreme events, as restricting candidates to a narrow temporal window may exclude matches that represent rare but important extreme events; and (3) it facilitates downscaling for future periods where no exact analogs exist in the historical record within a specific date range. Instead, suitable analogs may still be found in historical data but during different periods. For instance, as climate change progresses, a future day in winter may no longer have a match on the same calendar day in the past. Instead, an analog might be found on another calendar day, such as in a warmer season like fall or spring.

Using very low-resolution climate data as input reduces both memory requirements and computation time compared to physical snow models, which also require high-resolution climate data as input to estimate HR-SWE information. This efficiency enables the generation of HR-SWE estimates over large spatial domains with reduced computational overhead. This is particularly beneficial when applying the method to large areas, long temporal scales, or ensembles of climate data.

7. Comment: The conclusion could be more definitive in summarizing the key findings and contributions of the study. The main quantities results should be added in the conclusion.

7. Response: We have now divided the conclusion and discussion section into two separate sections: Section 6, Discussion, which includes more detailed analysis, and Section 7, Conclusion, to more definitively summarize the key findings and contributions of the study.

Below, we provide a portion focused on the results.

6. Discussion, Lines 523-540

In general, using the 'cnrm-esm2-1' model as an estimator results in better accuracy in Colorado at both 100 km and 9 km resolutions compared to other models. For instance, in Colorado, the use of the 'cnrm-esm2-1' model at a 9 km resolution demonstrated close agreement with observed SWE, with an average RMSE of 0.07 meters. This performance highlights the model's strong compatibility with the climatic and geographical complexities of Colorado. Conversely, in California, the 'ec-earth3-veg' model excels at a 9 km resolution, providing the most accurate results with an average RMSE of 0.13 meters compared to the reference datasets. This suggests that its higher resolution better captures the region's complex environmental and topographical variations.

It also appears that a finer resolution of 9 km provides slightly better accuracy than a 100 km resolution across all models, although the difference is not substantial. This underscores the importance of selecting the appropriate climate model for SWE estimation, which can have a more significant impact than merely choosing a higher-resolution model. Moreover, CMIP6 models are designed for long-term climate projections and capture broad climate trends rather than predicting specific weather events. Despite this, the downscaled SWE using the proposed approach based on CMIP6 is comparable to that of WRF-CMIP6, which dynamically downscales CMIP6. This is largely because the proposed methodology relies on long-term climate data through the use of far and near temporal intervals, and CMIP6 effectively captures broad climatic trends and seasonality, including changes in temperature and precipitation patterns. Accordingly, achieving accurate HR-SWE estimation relies significantly on the choice and accuracy of the climate model inputs, such as precipitation and temperature data, which may introduce biases into the SWE estimates. For example, precipitation biases are a dominant factor influencing SWE estimation errors, while temperature biases become more significant during transitional periods, such as the spring melt season.

Additional details have also been incorporated into this Section, as highlighted in Response 5, Discussion Section, Lines 548–551.

7. Conclusion, Lines 587-591:

Overall, the 'cnrm-esm2-1' model tends to provide higher accuracy in Colorado when used as an estimator, outperforming other models at both 100 km and 9 km resolutions. In contrast, the 'ec-earth3-veg' model performs best in California at a 9 km resolution. A finer resolution (9 km) generally offers slightly better accuracy than a 100 km resolution across models, though the difference is modest, emphasizing the importance of model selection over merely increasing resolution.

8. Comment: Given the limitations of ground observations in data-scarce high mountain regions, to what extent do you believe that the method used in your study can provide accurate and comprehensive insights into SWE? Are there any potential sources of uncertainty or bias that need to be addressed?

8. Response: The proposed methodology does not use ground observation data as input. Instead, it focuses on downscaling low-resolution climate data to generate high-resolution SWE information. However, we have now added more details to Section 6, Discussion, and Section 7, Conclusion, addressing potential sources of uncertainty. Specifically, we discuss how the choice and accuracy of climate model inputs may introduce biases into the SWE estimates (please refer to Response 7).

9. Comment: Abstract (Lines 1-20): The abstract is informative but could be condensed to focus more on the results and implications rather than the background. The key findings should be more prominently highlighted.

9. Response: The abstract has been revised to highlight key results as follows:

Lines 15-24:

To evaluate the performance of our approach, we conduct tests in California's Sierra Nevada and Colorado's Upper Colorado River Basin in the western United States using different low-resolution climate models ('ec-earth3-veg', 'mpi-esm1-2', and 'cnrm-esm2-1') at both 100 km and 9 km scales. A cross-validation analysis is performed, and comparisons are made with commonly used gridded SWE datasets as well as through point-scale time-series comparisons. The results demonstrate that our approach enables the generation of downscaled SWE, which closely matches reanalysis data in terms of statistical properties. The outputs demonstrate that, for each region, performance depends on the choice and accuracy of the climate model inputs, such as precipitation and temperature data. Overall, the 'cnrm-esm2-1' model demonstrates superior accuracy in Colorado, outperforming other models at both 100 km and 9 km resolutions. Conversely, the 'ec-earth3-veg' model achieves the best performance in California with 9 km climate data. Across models, a 9 km resolution typically provides slightly better accuracy compared to a 100 km resolution.

10. Comment: Please describe the study regions in the abstract to provide context for where the proposed method was tested. This addition will give readers a better understanding of the geographical relevance of the study, as follows:

10. Response: We have added a brief description of the study regions (California and Colorado) to the abstract to provide context. Please refer to Response 9.

11. Comment: Section 2.3 (Lines 115-130): The explanation of the K-nearest neighbor algorithm could be clarified. Please further illustrate how the algorithm works in the context of downscaling SWE.

11. Response: We have revised the paper to include a more detailed explanation of how the K-nearest neighbor algorithm operates in the context of SWE downscaling (please refer to Response 2). Additionally, we have incorporated more details and made several adjustments to improve clarity (please refer to Response 1).

12. Comment: Ensure consistency in the use of terms such as "high-resolution" and "low-resolution" throughout the manuscript. Some sections use these terms interchangeably with "HR-SWE" and "LR-SWE," which could confuse readers.

12. Response: We have carefully reviewed the manuscript to ensure consistency in the use of the terms "high-resolution SWE" (HR-SWE) and "low-resolution SWE" (LR-SWE).

13. Comment: There seems to be an issue with Table 3, as some information might be missing or incomplete.

13. Response: It is corrected now and also based on the second reviewer's comment that Table 3 contains commonly used metrics, we have moved Table 3 to the Supplementary.

14. Comment: Lines 268-269: The reference to "Evaluation criteria Table 1" appears to be incorrect. Please check this sentence and ensure that the correct table is referenced.

14. Response: This issue has been corrected.

15. Comment: The legend and boxplot in Figure 6 require revision, as they currently do not appear to be properly formatted. Please ensure that the legend is clear and correctly labeled, and that the boxplot is visually consistent with the rest of the figure.

15. Response: We use a combination of color and line style to avoid repetitions.To make it clear, now we have added an explanation in the figure caption as follows:

Fig. 7. The average of the UCLA SWE (black dotted line, reference) and the downscaled SWEs ($\widehat{SWE}$; blue and green lines) for each area over the six-year period from 2005 to 2010 are shown. The three climate models used in the downscaling are represented by different colors: light blue for 'ec-earth3-veg', green for 'cnrm-esm2-1', and dark blue for 'mpi-esm1-2'. Line styles indicate the spatial resolution of the climate data: solid lines correspond to results based on a 9 km resolution, and dashed lines represent results based on a 100 km resolution. The dotted black line represents the average UCLA SWE, which is the reference.

16. Comment: Figures 9, 11, 13, and supplementary figures: The Y-axis values in Figures 9, 11, 13, and the supplementary figures need to be revised. Please check that the scales are appropriate and clearly labeled to ensure that they accurately represent the data being presented.

16. Response: We have revised the Y-axis values in Figures 9, 11, 13, and the supplementary figures.

17. Comment: Tables and Figures Captions: Please remove the bold formatting from all table and figure captions. The current formatting detracts from the overall appearance of the manuscript. Standardize the captions to match the style used throughout the manuscript.

17. Response: We have removed the bold formatting from all table and figure captions and standardized the style.

18. Comment: There are a few minor grammatical errors and typos that should be corrected. For example, in line 30, "Snow Eater equivalent" should be corrected to "Snow Water Equivalent."

18. Response: We have corrected all identified grammatical errors and typos, including changing "Snow Eater equivalent" to "Snow Water Equivalent."

We hope these revisions address all of your concerns, and we thank you once again for your valuable feedback.

References:

Abatzoglou, J. T. and Brown, T. J.: A comparison of statistical downscaling methods suited for wildfire applications, International journal of climatology, 32, 772-780, 2012.

Caillouet, L., Vidal, J.-P., Sauquet, E., and Graff, B.: Probabilistic precipitation and temperature downscaling of the Twentieth Century Reanalysis over France, Climate of the Past, 12, 635-662, 2016.

Cover, T. and Hart, P.: Nearest neighbor pattern classification, IEEE transactions on information theory, 13, 21-27, 1967.

Gangopadhyay, S., Clark, M., and Rajagopalan, B.: Statistical downscaling using K‐nearest neighbors, Water Resources Research, 41, 2005.

Horton, P.: AtmoSwing: Analog technique model for statistical weather forecastING and downscalING (v2. 1.0), Geoscientific Model Development, 12, 2915-2940, 2019.

Kouki, K., Räisänen, P., Luojus, K., Luomaranta, A., and Riihelä, A.: Evaluation of Northern Hemisphere snow water equivalent in CMIP6 models during 1982–2014, The Cryosphere, 16, 1007-1030, 2022.

Lall, U. and Sharma, A.: A nearest neighbor bootstrap for resampling hydrologic time series, Water resources research, 32, 679-693, 1996.

Pons, M., San-Martín, D., Herrera, S., and Gutiérrez, J. M.: Snow trends in Northern Spain: analysis and simulation with statistical downscaling methods, international journal of climatology, 2010.

Rajagopalan, B. and Lall, U.: A k‐nearest‐neighbor simulator for daily precipitation and other weather variables, Water resources research, 35, 3089-3101, 1999.

Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D.: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, Climatic change, 62, 189-216, 2004.

Yates, D., Gangopadhyay, S., Rajagopalan, B., and Strzepek, K.: A technique for generating regional climate scenarios using a nearest‐neighbor algorithm, Water resources research, 39, 2003.

Yiou, P. and Déandréis, C.: Stochastic ensemble climate forecast with an analogue model, Geoscientific Model Development, 12, 723-734, 2019.

Zorita, E. and Von Storch, H.: The analog method as a simple statistical downscaling technique: Comparison with more complicated methods, Journal of climate, 12, 2474-2489, 1999.

Response to Reviewer Comments

Dear Editor,

Thank you for providing valuable feedback from the reviewers on our manuscript "High-Resolution Snow Water Equivalent Estimation: A Data-Driven Method for Localized Downscaling of Climate Data." We have carefully addressed all of the comments and suggestions to improve the clarity and overall quality of the paper. Below, we provide a point-by-point response to the reviewers' comments.

Reviewer 2

1. Comment: This study by Zakeri et al. presents a method for downscaling Snow Water Equivalent (SWE) to 500 m, from a low-resolution SWE dataset and low-resolution climate reanalysis data. The model is trained using a high-resolution SWE dataset (SWE reanalysis from UCLA) and low-resolution climate data. The downscaling algorithm uses a K-nearest neighbors' method, for which parameters are determined through sensitivity analysis and optimization. Results are shown in two mountainous regions of the Western United States, in California and Colorado, and compared to other SWE spatial datasets.

The paper is interesting for the snow and hydrology community as it offers a new method to produce useful spatial SWE datasets over large areas, without using physical modelling of the snowpack which can be computationally expensive. The method is quite clearly exposed.

However, the paper generally lacks a solid discussion of the results. Results and metrics of comparison are exposed in the result section but are not further discussed since the Discussion and Conclusion section is more a summary. A dedicated Discussion section would be welcome, in particular to convince the reader of the quality of the evaluated method against other methods, provide more physically based explanations of differences (e.g. can the chosen method explain certain biases?), provide a critical point of view on the benefits and limitations of the method, etc.

1. Response: We would like to thank the reviewer for their constructive feedback. To address the reviewer's concern we have added a new Section (6. Discussion) to include a dedicated discussion. This section now provides an in-depth analysis of the results, a comparison with different downscaling methodologies, an exploration of physical explanations, and an evaluation of the method's strengths and limitations.

Lines 523-568:

6. Discussion

In general, using the 'cnrm-esm2-1' model as an estimator results in better accuracy in Colorado at both 100 km and 9 km resolutions compared to other models. For instance, in Colorado, the use of the 'cnrm-esm2-1' model at a 9 km resolution demonstrated close agreement with observed SWE, with an average RMSE of 0.07 meters. This performance highlights the model's strong compatibility with the climatic and geographical complexities of Colorado. Conversely, in California, the 'ec-earth3-veg' model excels at a 9 km resolution, providing the most accurate results with an average RMSE of 0.13 meters compared to the reference datasets. This suggests that its higher resolution better captures the region's complex environmental and topographical variations.

It also appears that a finer resolution of 9 km provides slightly better accuracy than a 100 km resolution across all models, although the difference is not substantial. This underscores the importance of selecting the appropriate climate model for SWE estimation, which can have a more significant impact than merely choosing a higher-resolution model. Moreover, CMIP6 models are designed for long-term climate projections and capture broad climate trends rather than predicting specific weather events. Despite this, the downscaled SWE using the proposed approach based on CMIP6 is comparable to that of WRF-CMIP6, which dynamically downscales CMIP6. This is largely because the proposed methodology relies on long-term climate data through the use of far and near temporal intervals, and CMIP6 effectively captures broad climatic trends and seasonality, including changes in temperature and precipitation patterns. Accordingly, achieving accurate HR-SWE estimation relies significantly on the choice and accuracy of the climate model inputs, such as precipitation and temperature data, which may introduce biases into the SWE estimates. For example, precipitation biases are a dominant factor influencing SWE estimation errors, while temperature biases become more significant during transitional periods, such as the spring melt season.

The following comparison provides a broader perspective on how our proposed method compares with other statistical downscaling techniques. BCSD methods are effective in reducing uncertainties in climate model outputs by adjusting model biases using high-resolution observations. These methods are particularly valuable for ensuring that model outputs align with observed climatology and capture local variability. However, they depend heavily on the availability of high-quality in-situ data, which limits their application in remote or data-sparse regions. In contrast, our method excels in areas with sparse data, as it uses low-resolution climate data without requiring ground observations, making it adaptable to a broader range of conditions.

Analog-type statistical downscaling approaches offer a simple, computationally efficient way to project high-resolution data based on historical relationships between large-scale climate patterns and local climate variables. These methods are useful in regions where historical climate patterns are stable and well-documented. On the other hand, our method introduces several key improvements over traditional k-NN downscaling techniques. First, the adaptation of the k-NN approach through the incorporation of far and near temporal intervals of climate data enhances its ability to handle dynamic variables, such as snow, which are subject to significant changes due to climate variability. Unlike conventional analog methods, which constrain analog candidates to a specific temporal window near the query date, this method eliminates such restrictions.

This flexibility is crucial for three reasons: (1) the inclusion of far and near temporal intervals allows the selection of the most suitable candidates across a broader temporal range, eliminating the need for narrow constraints; (2) it ensures the preservation of extreme events, as restricting candidates to a narrow temporal window may exclude matches that represent rare but important extreme events; and (3) it facilitates downscaling for future periods where no exact analogs exist in the historical record within a specific date range. Instead, suitable analogs may still be found in historical data but during different periods. For instance, as climate change progresses, a future day in winter may no longer have a match on the same calendar day in the past. Instead, an analog might be found on another calendar day, such as in a warmer season like fall or spring.

Using very low-resolution climate data as input reduces both memory requirements and computation time compared to physical snow models, which also require high-resolution climate data as input to estimate HR-SWE information. This efficiency enables the generation of HR-SWE estimates over large spatial domains with reduced computational overhead. This is particularly beneficial when applying the method to large areas, long temporal scales, or ensembles of climate data.

2. Comment: Bales et al. (2006) is cited three times (l. 33), more references (particularly more recent ones) could be used in the beginning of the introduction. You could also provide a few more references about physically-based snow models (l. 34).

2. Response: To address this comment, the Introduction is now revised to include recent references, and additional references on physically based snow models to provide a more balanced, and comprehensive context, as follows: Introduction Section, lines 35–53:

Therefore, accurate and detailed information on Snow Water Equivalent (SWE) with high temporal and spatial resolution is crucial for effective water resource management and decision-making (Bales et al., 2006; Fiddes et al., 2019; Siirila-Woodburn et al., 2021b).

Although ground stations are valuable for collecting SWE data, their limited presence in certain regions affects their representativeness. Moreover, variations in topography, land cover, and environmental conditions in mountainous areas make point-scale data insufficient for capturing the overall spatial characteristics of a watershed (Bales et al., 2006; Alonso-González et al., 2023). To address this lack of data, physically based snow models utilize an energy balance approach to estimate snowmelt. These models range in complexity, with more advanced models integrating detailed processes such as wind-induced snow transport, interactions with topography, and vegetation impacts. While complex models, such as those incorporating advection-diffusion equations or three-dimensional wind fields, provide more accurate representations of snow properties, they often require extensive input data, which may not always be available (Liston and Sturm, 1998; Lehning et al., 2006; Vionnet et al., 2014). Simpler models, on the other hand, may fail to capture critical aspects of snow dynamics (Bair et al., 2016; Clow et al., 2012). Moreover, to achieve high-resolution SWE (HR-SWE) estimates using these models, it is necessary to use meteorological and land cover-related data that match the desired output resolution of the SWE. However, obtaining high-resolution data in mountainous regions remains challenging (Wundram and Löffler, 2008). Although generating HR-SWE with physical models can be time-consuming, recent studies have focused on reducing this computational burden. For instance, advances in SWE modeling have been achieved by implementing parallelized versions of snow models. This approach maintains the integrity of physical processes while utilizing parallelization to manage the computational demands of fine-resolution datasets over large domains (Mower et al., 2024).

3. Comment: l. 61: specify "spatial snow patterns" (or do you also mean temporal patterns?)

3. Response: We have clarified the text to specify that it refers exclusively to spatiotemporal snow patterns that are repeatable between days with similar climatological data.

Recognizing the potential for snow spatiotemporal patterns to repeat on days with similar climatological characteristics.

4. Comment: l. 85-86: "physical models are computationally expensive". This assertion is not true by itself, see several recent publications where physical snow models are being applied over large domains, e.g. Mower et al. (2024). Please specify in a few words what makes them too computationally expensive for the application (domain, resolution, solved processes...?), or what computational benefit your method offers.

4. Response:

As suggested, we have modified the introduction to include a mention of recent studies that focus on reducing computational burdens, including the approach taken by Mower et al. (2024). Additionally, we have added details about what makes physical models computationally expensive and the benefits of the proposed methodology in the Introduction, as follows:

Introduction lines 49-57:

Although generating HR-SWE with physical models can be time-consuming, recent studies have focused on reducing this computational burden. For instance, advances in SWE modeling have been achieved by implementing parallelized versions of snow models. This approach maintains the integrity of physical processes while utilizing parallelization to manage the computational demands of fine-resolution datasets over large domains (Mower et al., 2024). However, it is important to clarify that the computational expense becomes significant when operating at high spatial and temporal resolutions, particularly in an ensemble context, which is often required for robust climate predictions. Additionally, regardless of the computational aspects, generating high-resolution snow data using physical models typically necessitates high-resolution climate inputs, such as temperature and precipitation fields, to ensure the quality of the downscaled outputs.

And Introduction lines 135-136:

Finally, our method does not require high-resolution climate input data. This substantially reduces computational demands while maintaining the high quality of the downscaled SWE data.

5. Comment: l. 92: "climatic variables". Perhaps rather say "meteorological variables".

5. Response: We agree and have revised the manuscript to replace 'climatic variables' with 'meteorological variables' for greater clarity.

6. Comment: l. 94-95: "not subject to significant temporal variations within the specified regions". Note that terrain shading depends on the sun position, so varies throughout the year.

6. Response: We have revised the text to acknowledge the variation in terrain shading throughout the year due to the sun's position.

Lines 166-172:

Although the effects of environmental variables like terrain shading may vary, driven by the sun's position throughout the year, these can be captured in the baseline SWE models used to generate training data. In contrast, climatic variables exhibit significant temporal and spatial variability, making them the primary drivers of SWE dynamics. Consequently, the downscaled SWE estimation focuses primarily on climatic variables. Their dominant influence ensures that the proposed model can capture the spatiotemporal variability essential for SWE estimation. The effects of environmental variables, while excluded from the direct downscaling process, can be accounted for in the baseline SWE models that provide the training data.

7. Comment: l. 97: "SWE is also affected by conditions in the preceding periods". This sentence is a bit unclear: do you mean SWE conditions? Meteorological conditions? Please clarify "preceding periods" too.

7. Response: We have clarified the sentence in the text:

Lines 173-174:

SWE is also affected by preceding meteorological conditions, such as the temperature and precipitation patterns of previous days. For instance, the amount of SWE today may vary depending on the conditions experienced in the preceding days.

8. Comment: l. 98-99: "These intervals consider climate variables such as minimum temperature, maximum temperature, precipitation, and surface downwelling shortwave radiation". The use of "such as" makes it unclear whether it is an exhaustive list of considered variables. This section needs more justification of why and how these specific variables were chosen, and why other meteorological variables were discarded.

8. Response: We have revised the manuscript to clearly indicate that this is an exhaustive list of considered variables. Additionally, we have provided justification for the selection of these variables and explained why others were not included.

Lines 191-192:

Details on the experiments conducted to identify the most effective meteorological predictors within these datasets are available in the Supplementary (Table S1 and Fig. S1).

Table S1. A summary of the selected features in each scenario.

Scenarios	Features
1st scenario	Minimum and maximum temperatures, precipitation, downwelling shortwave radiation
2nd scenario	Minimum and maximum temperatures, precipitation, downwelling shortwave radiation, LR-SWE
3rd scenario	Minimum and maximum temperatures, precipitation, LR-SWE, MODIS snow cover
4th scenario	Minimum and maximum temperatures, precipitation, LR-SWE, MODIS snow cover, monthly terrain shadow
5th scenario	Minimum and maximum temperatures, precipitation, LR-SWE, downwelling shortwave radiation, MODIS snow cover
6th scenario	Minimum and maximum temperatures, precipitation, LR-SWE, downwelling shortwave radiation, MODIS snow cover, monthly terrain shadow

To evaluate the efficacy of the chosen features, we synthesize SWE images of Colorado for six consecutive years (2005–2010) using ERA5-Land climate data, MODIS snow cover, and monthly terrain shadow with different combinations as explained in Table S1. We analyze the months of January, February, March, April, November, and December across these years. The sum of the RMSE, (1-correlation), and Mean Difference is determined using cross-validation, and the average is calculated to identify which parameters consistently outperform the others (Fig. S1). As shown in Fig. S1, the second scenario generally demonstrates superior performance.

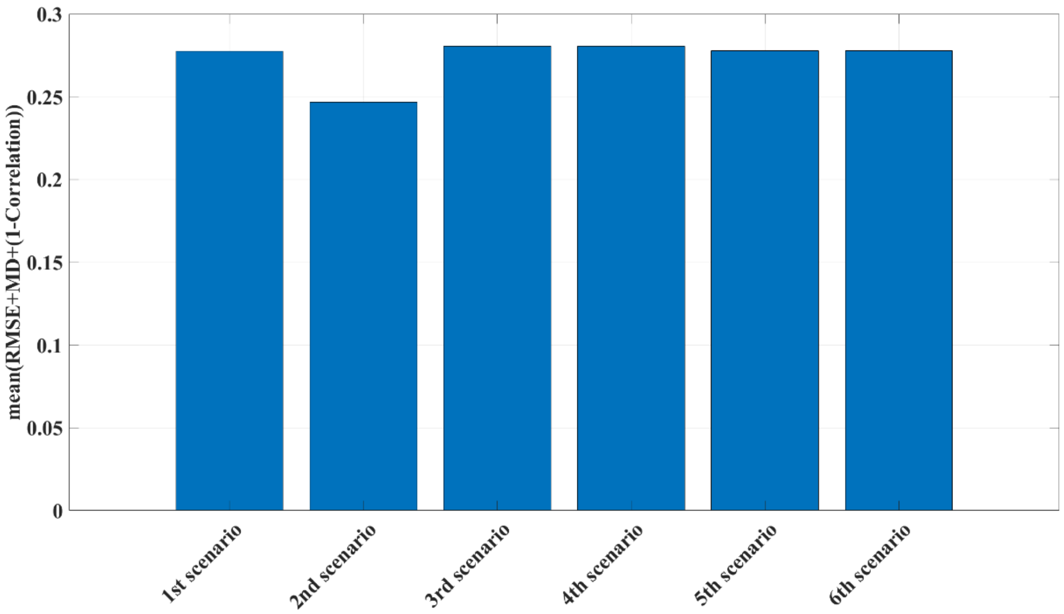


Fig. S1. The average error for the scenarios presented in Table S1.

9. Comment: L. 143: The reference HR-SWE is not an actual "observation", so please simply use reference instead of observation.

9. Response: We have revised the text to replace 'observation' with 'reference', as suggested.

10. Comment: l. 147: please provide the units (days?).

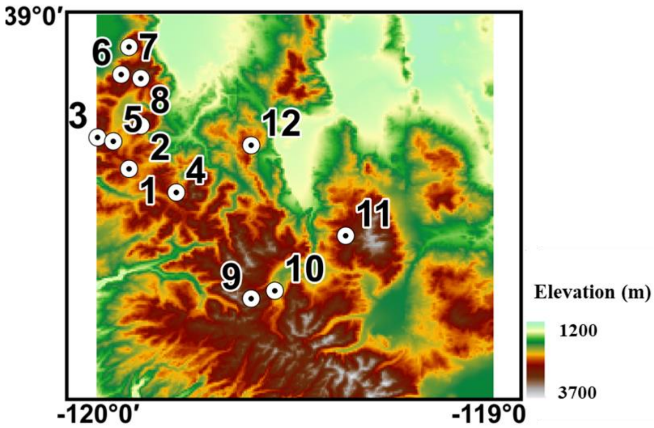
10. Response: We have specified the units as 'days' in the revised manuscript.

Line 230-231:

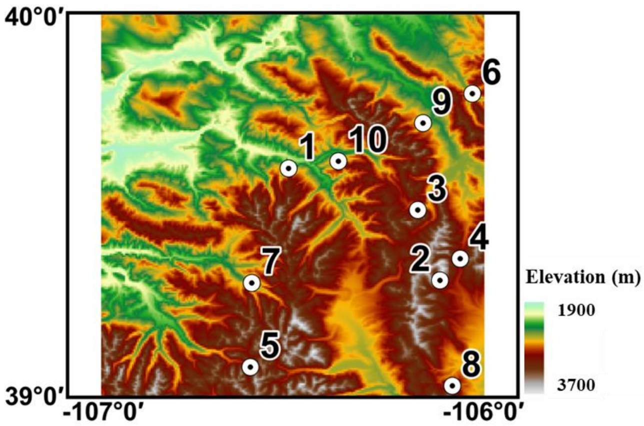
we conduct a sensitivity analysis within predefined ranges, set by NI ∈ [1,7] days and FI ∈ [1,90] days, aiming to minimize ε .

11. Comment: Fig. 2: Maybe a better colour than blue can be chosen for visibility of the numbers on the map.

11. Response: Thank you for this suggestion. We have revised Fig. 2 and updated the color scheme to improve the visibility of the numbers on the map. A contrasting color was selected to enhance clarity for the readers.



(b) California



(c) Colorado

12. Comment: Section 3.1: you could add a few words on the snow climatology of these two regions.

12. Response: We have added a brief description of the snow climatology for the California and Colorado regions in lines 260-274 as follows:

In the California region, elevation ranges from a minimum of 1200 m to a maximum of 3700 m, with an average elevation of 2200 m. In the Colorado region, elevation ranges from a minimum of 1900 m to a maximum of 4300 m, with an average elevation of 3000 m. By focusing on these specific regions, we can assess the performance and applicability of the proposed methodology in other areas where snow water resources are an important component of the hydrology system.

The California Sierra Nevada snowpack plays a critical role in water resource management, contributing approximately 30% of the state's water supply through snowmelt. However, its sensitivity to warming temperatures is evident, with the 2015 SWE on April 1 dropping to just 5% of historical averages. This dramatic decline underscores the combined effects of reduced precipitation and higher temperatures, which exacerbate drought severity and shift the timing of water availability (Belmecheri et al., 2016).

In the Upper Colorado River Basin, snowmelt accounts for 70–90% of annual streamflow, making snowpack dynamics essential to hydrological processes and water management.

Heldmyer et al. (2023) identify three distinct snow drought types: 'warm', 'dry', and 'warm-and-dry', which differentially impact SWE and streamflow timing. Warm droughts tend to reduce SWE at lower elevations, while dry conditions cause uniform SWE reductions across elevations. These droughts advance peak streamflow timing by 7–13 days, emphasizing the region's sensitivity to climatic changes in temperature and precipitation (Heldmyer et al., 2023).

13. Comment: l. 175: A bit more details (one or two sentences) on what this dataset is based would be appreciated for the reader's understanding.

13. Response: We agree and have expanded the description of the dataset to include details on its basis, as outlined in lines 278–282:

It combines high-resolution remotely-sensed data with a Bayesian data assimilation (Margulis et al., 2019; Margulis et al., 2015; Margulis et al., 2016) framework. The dataset is derived from Landsat-based fractional snow-covered area observations, updated daily, and incorporates a land surface model to estimate SWE and snow depth. This approach enables spatially and temporally continuous SWE estimates, which are verified against in situ and lidar-derived SWE measurements for accuracy.

14. Comment: l. 180: Please don't use "such as" if it is the exhaustive list of considered variables.

14. Response: We have revised the text to remove 'such as' and explicitly state that the listed variables are included.

15. Comment: l. 227-228: it would be interesting to discuss the relative weights of each variables from a physical perspective.

15. Response: We have added a discussion in lines 336-340 about the relative weights of each variable, including an analysis of their physical significance and contribution to the downscaling process as follows:

From a physical perspective, this weighting aligns with the processes governing SWE distribution and dynamics. Minimum temperature ($\boldsymbol{T^{NI}_{min,LR}}$) significantly influences freezing and melting thresholds, which are critical for snowpack accumulation or reduction. Precipitation variables, both near and far intervals ($\boldsymbol{P^{NI}_{LR}}$, $\boldsymbol{P^{FI}_{LR}}$), directly contribute to SWE through their impact on the volume of snowfall. Meanwhile, the inclusion of $\boldsymbol{SWE_{LR}}$ as a highly weighted variable underscores highlights its importance role as a baseline indicator of existing snowpack conditions.

16. Comment: l. 240: Please stick to mm as SWE unit, which corresponds to the standard unit kg/m2. Potentially m for high values, but cm introduces confusion.

16. Response: We have updated the manuscript to consistently use millimeters (mm) as the standard unit for low SWE values. For exceptionally high values, meters (m) are used to maintain clarity and avoid confusion.

17. Comment: l. 246-250, Table 3: It is probably unnecessary to define very common metrics like mean difference, correlation and RMSE.

17. Response: We agree with this comment and have moved Table 3 to the Supplementary.

18. Comment: Why aren't in-situ SNOTEL SWE measurements not shown? Metrics of comparison to these point measurements could be presented and compared to metrics of the other products.

18. Response: In our study, we utilized the spatial locations of in-situ SNOTEL SWE measurement sites for pixel-based evaluations, comparing our downscaled SWE estimates with well-established SWE datasets in the United States, including the UCLA SWE dataset, SNODAS, Daymet, and the University of Arizona dataset. This approach demonstrates that the proposed downscaling method effectively preserves SWE information at the pixel level while analyzing the values of each pixel over multiple years.

It is worth noting that the UCLA SWE dataset has been previously validated against in-situ observations, including SNOTEL SWE measurements, establishing its robustness as a state-of-the-art dataset for SWE estimation in the Western United States (Fang et al., 2022). Moreover, the snow reanalysis framework has been successfully utilized in the past to generate datasets for the Sierra Nevada (Margulis et al., 2016).

Additionally, the UCLA SWE dataset has been widely adopted in subsequent studies as a benchmark for SWE validation. For instance:

Ma et al. (2023) used the UCLA SWE dataset as a ground truth reference to evaluate machine learning approaches for spatiotemporal SWE estimation.

Fang et al. (2023) employed the UCLA SWE dataset as a baseline reference to quantify snow water storage uncertainty in the midlatitude American Cordillera.

By utilizing this extensively validated reanalysis dataset and incorporating SNOTEL site locations for pixel-based assessments, we ensured a rigorous evaluation of our methodology. The manuscript has been revised to emphasize these points and improve clarity, as follows in lines 356-362:

While the SNOTEL network provided the spatial locations for these comparisons, we did not use direct in-situ SNOTEL SWE measurements in this analysis. The UCLA SWE dataset has been previously validated against in-situ observations, including SNOTEL, and has been widely recognized as a benchmark for SWE validation in subsequent studies. For example, Ma et al. (2023) used the UCLA SWE dataset as a reference to assess the performance of machine learning approaches for estimating spatiotemporally continuous SWE. Similarly, Fang et al. (2023) used the UCLA SWE dataset as a baseline for evaluating snow water storage uncertainty. These applications highlight the dataset's reliability and its critical role in advancing snow hydrology research.

19. Comment: See the general comment about a more in-depth discussion of the results.

19. Response: We have addressed the general comment. Please refer to our response to Comment 1.

We hope these revisions address all of your concerns, and we thank you once again for your valuable feedback.

References:

Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R., and Dozier, J.: Mountain hydrology of the western United States, Water Resources Research, 42, 2006.

Fiddes, J., Aalstad, K., and Westermann, S.: Hyper-resolution ensemble-based snow reanalysis in mountain regions using clustering, Hydrology Earth System Sciences, 23, 4717-4736, 2019.

Alonso-González, E., Aalstad, K., Pirk, N., Mazzolini, M., Treichler, D., Leclercq, P., Westermann, S., López-Moreno, J. I., and Gascoin, S.: Spatio-temporal information propagation using sparse observations in hyper-resolution ensemble-based snow data assimilation, Hydrology Earth System Sciences, 27, 4637-4659, 2023.

Liston, G. E. and Sturm, M.: A snow-transport model for complex terrain, Journal of Glaciology, 44, 498-516, 1998.

Vionnet, V., Martin, E., Masson, V., Guyomarc'h, G., Naaim-Bouvet, F., Prokop, A., Durand, Y., and Lac, C.: Simulation of wind-induced snow transport and sublimation in alpine terrain using a fully coupled snowpack/atmosphere model, The Cryosphere, 8, 395-415, 2014.

Bair, E. H., Rittger, K., Davis, R. E., Painter, T. H., and Dozier, J.: Validating reconstruction of snow water equivalent in California's Sierra Nevada using measurements from the NASA Airborne Snow Observatory, Water Resources Research, 52, 8437-8460, 2016.

Clow, D. W., Nanus, L., Verdin, K. L., and Schmidt, J.: Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA, Hydrological Processes, 26, 2583-2591, 2012.

Wundram, D. and Löffler, J.: High-resolution spatial analysis of mountain landscapes using a low-altitude remote sensing approach, International Journal of Remote Sensing, 29, 961-974, 2008.

Mower, R., Gutmann, E. D., Liston, G. E., Lundquist, J., and Rasmussen, S.: Parallel SnowModel (v1. 0): a parallel implementation of a distributed snow-evolution modeling system (SnowModel), Geoscientific Model Development, 17, 4135-4154, 2024.

Belmecheri, S., Babst, F., Wahl, E. R., Stahle, D. W., and Trouet, V.: Multi-century evaluation of Sierra Nevada snowpack, Nature Climate Change, 6, 2-3, 2016.

Heldmyer, A. J., Bjarke, N. R., and Livneh, B.: A 21st-Century perspective on snow drought in the Upper Colorado River Basin, Journal of the American Water Resources Association, 59, 396-415, 2023.

Ma, X., Li, D., Fang, Y., Margulis, S. A., and Lettenmaier, D. P.: Estimating spatiotemporally continuous snow water equivalent from intermittent satellite observations: an evaluation using synthetic data, Hydrology Earth System Sciences, 27, 21-38, 2023.

Fang, Y., Liu, Y., Li, D., Sun, H., and Margulis, S. A.: Spatiotemporal snow water storage uncertainty in the midlatitude American Cordillera, The Cryosphere, 17, 5175-5195, 2023.

Siirila-Woodburn, E. R., Rhoades, A. M., Hatchett, B. J., Huning, L. S., Szinai, J., Tague, C., Nico, P. S., Feldman, D. R., Jones, A. D., and Collins, W. D.: A low-to-no snow future and its impacts on water resources in the western United States, Nature Reviews Earth and Environmental, 2, 800-819, 2021.

Fang, Y., Liu, Y., and Margulis, S. A.: A western United States snow reanalysis dataset over the Landsat era from water years 1985 to 2021, Scientific Data, 9, 1-17, 2022.

Margulis, S. A., Cortés, G., Girotto, M., and Durand, M.: A Landsat-era Sierra Nevada snow reanalysis (1985–2015), Journal of Hydrometeorology, 17, 1203-1221, 2016.

Margulis, S. A., Liu, Y., and Baldo, E.: A joint landsat-and modis-based reanalysis approach for midlatitude montane seasonal snow characterization, Frontiers in Earth Science, 7, 272, 2019.

Margulis, S. A., Girotto, M., Cortés, G., and Durand, M.: A particle batch smoother approach to snow water equivalent estimation, Journal of Hydrometeorology, 16, 1752-1772, 2015.