Overall Comments

This manuscript presents a technically ambitious and potentially valuable global dataset of snowmelt runoff onset derived from Sentinel-1 SAR observations constrained by MODIS-based snow phenology information. The study addresses an important observational gap between coarse passive microwave snowmelt products and sparse in situ measurements, and the large-scale implementation of a multi-orbit SAR processing framework represents a substantial engineering achievement. The manuscript is generally well structured, and the authors provide extensive discussion of environmental controls, validation results, and potential applications.

However, despite the strong technical contribution, the current manuscript contains several conceptual, methodological, and interpretational weaknesses that substantially limit confidence in the physical meaning and global applicability of the dataset. The most important issue is that the study implicitly treats the SAR backscatter minimum as a physically meaningful and globally transferable indicator of "runoff onset," while the manuscript itself simultaneously acknowledges that the backscatter minimum is influenced by multiple interacting processes, including liquid water content, snow surface roughness, canopy interactions, and environmental variability. In its current form, the manuscript demonstrates empirical consistency under selected conditions, but it does not yet establish that the retrieved timing represents a physically consistent runoff onset metric across different snow regimes. In addition, the validation framework is geographically limited relative to the claimed global applicability, several methodological choices appear heuristic and insufficiently justified, and the uncertainty characterization remains incomplete. Overall, the dataset has strong publication potential, but the manuscript requires clearer physical framing, more cautious interpretation, stronger sensitivity analyses, and more rigorous uncertainty quantification.

Major Comments

1. Physical interpretation of "Runoff Onset" needs clearer framing

The central conceptual issue of the manuscript is the interpretation of the SAR backscatter minimum as "snowmelt runoff onset." The manuscript repeatedly presents the backscatter minimum as a direct indicator of runoff initiation, yet the physical basis for this interpretation remains insufficiently demonstrated. As discussed in Sect. 5.2–5.3 and in the studies cited by the authors themselves, particularly Carletti et al. (2025), the SAR backscatter minimum is not a direct measurement of runoff generation, but rather the integrated result of multiple processes, including liquid water absorption, snowpack wetting, evolving surface roughness, vegetation interactions, and changes in scattering mechanisms. Importantly, none of these processes necessarily occur at a fixed temporal offset relative to actual meltwater outflow from the snowpack. The relationship between backscatter minima and true runoff initiation is therefore likely to vary systematically among maritime, continental, tundra, alpine, and shallow snow environments. The manuscript currently does not provide any physically based transfer function or process model capable of quantifying this variability. As a result, the study demonstrates empirical correspondence rather than physical equivalence.

2. Global dataset but regionally limited validation

The dataset is presented as a global product, yet the evaluation is based almost entirely on snow pillow observations from the western United States. While this region provides an exceptional observational network for algorithm development and validation, it represents only a subset of global snow environments.

Snowpack properties, vegetation structure, climate regimes, melt processes, and SAR interactions vary substantially among maritime, continental, alpine, tundra, boreal, Arctic, and High Mountain Asia environments. Therefore, it remains unclear whether the empirical thresholds identified in the manuscript, including the forest cover threshold (~0.5), SWE threshold (~20 cm), and temporal resolution threshold (<14 days), are broadly transferable beyond the western United States. The manuscript currently does not provide sufficient evidence to support the global applicability of these relationships.

3. Strong dependence on MODIS snow phenology constraints

The retrieval framework relies heavily on the MODIS MOD10A2 snow product to define the temporal search window for runoff onset detection. However, the implications of uncertainties in the MODIS-derived snow phenology information are not sufficiently assessed.

Because the runoff onset retrieval is constrained by MODIS-derived snow presence and disappearance dates, any biases or uncertainties in the snow phenology product may directly propagate into the final runoff onset estimates. While the manuscript briefly acknowledges inherited MODIS limitations, it does not quantify the sensitivity of the retrieval results to uncertainties in the MODIS constraints. As a result, the robustness of the retrieval framework to uncertainties in the underlying snow phenology information remains unclear.

A related issue concerns the spatial consistency between the 500 m MODIS product and the 80 m Sentinel-1 retrieval. The manuscript provides limited discussion of how the substantial difference in spatial resolution is reconciled. Potential effects of cross-scale inconsistencies may become important in heterogeneous terrain, patchy snow environments, forest edges, and complex mountain regions where sub-pixel variability is large.

4. Several methodological parameters appear heuristic and insufficiently justified

Several key methodological parameters appear to be selected empirically, yet their robustness is not systematically evaluated. Examples include the 56-day continuous snow criterion, the 16-day search-window extension, the 30-day orbit-gap threshold, and the 95% SWE runoff definition used in the validation framework. Although these choices may be operationally reasonable, the manuscript provides limited evidence demonstrating their applicability across different snow regimes and environmental conditions.

The justification for the selection of the 80 m spatial resolution is also not fully convincing. The manuscript links the chosen resolution to previously reported variance-stabilization scales, yet the logical connection between variance minimization and retrieval optimality is not clearly established. A minimum variance scale does not necessarily correspond to the most physically meaningful or information-rich scale for snowmelt detection. Furthermore, no sensitivity analysis is presented to evaluate how retrieval performance changes with spatial resolution.

In addition, the statistical robustness of the 10-year composite products requires further consideration. In some regions, particularly northern Asia, the composite products are derived from only a small number of valid years. The manuscript does not quantify the uncertainty associated with such limited temporal sampling, making it difficult to determine whether observed spatial patterns reflect meaningful climatic variability or sampling effects.

5. Uncertainty characterization and quality information are insufficient

As a global Earth system dataset, uncertainty characterization represents a critical component of the product description. However, the current manuscript provides limited information regarding retrieval uncertainty.

The dataset includes annual runoff onset estimates, temporal resolution information, and composite statistics, but it does not provide a direct characterization of retrieval uncertainty. In particular, the reported median absolute deviation primarily reflects variability among years rather than uncertainty associated with the retrieval itself. Consequently, users may have difficulty distinguishing between genuine environmental variability and methodological uncertainty.

The manuscript also identifies several factors that strongly influence retrieval performance, including forest cover fraction, SWE, temporal resolution, and observation availability. However, these relationships are not translated into a quantitative assessment of spatially varying data quality. As a result, the reliability of the product remains difficult to evaluate in regions where environmental conditions differ substantially from those represented in the validation dataset

6. Interpretation of temporal resolution effects may be confounded by environmental heterogeneity

The manuscript concludes that finer temporal resolution improves agreement between SAR-derived runoff onset estimates and snow pillow observations. While such a relationship is plausible, the current analysis does not clearly isolate temporal resolution from other environmental factors.

Temporal resolution is not randomly distributed across the study domain. Regions with finer revisit frequency often differ systematically from regions with coarser revisit intervals in terms of snow climate, vegetation structure, topography, and snowpack characteristics. Therefore, part of the observed relationship between temporal resolution and retrieval performance may reflect environmental differences rather than temporal sampling effects alone.

Because these factors are not explicitly disentangled, the current evidence is insufficient to establish a causal relationship between revisit frequency and retrieval accuracy.

7. Coverage statistics and evaluation strategy may lead to overly optimistic interpretations

The presentation of spatial coverage may lead readers to overestimate the effective completeness of the dataset. The reported coverage percentages are calculated relative to "observable seasonal snow extent," which already excludes regions lacking Sentinel-1 IW coverage, including portions of Greenland, Arctic islands, and Antarctica. Because these exclusions occur within the denominator itself, the reported coverage values may appear substantially closer to complete global coverage than is actually the case.

A related concern involves the evaluation strategy. Several empirical filtering criteria are applied prior to reporting station-level agreement statistics, including restrictions based on forest cover fraction, SWE, and temporal resolution. While such filtering may be useful for identifying favorable retrieval conditions, it also removes many of the most challenging environments before performance metrics are reported. Consequently, it remains difficult to determine how the product performs across the full range of environmental conditions represented within the dataset.

Minor Comments

1. Line 97-98 contains multiple hyphenated "-ed" adjectives that are grammatically acceptable but stylistically awkward. The sentence could be improved by rephrasing to "available as radiometrically terrain-corrected, cloud-optimized GeoTIFFs" or similar wording.
2. The rationale linking the variance analysis of Manickam and Barros (2020) to the selection of 80 m resolution in Lines 103-106 should be revised. The current interpretation incorrectly implies that the minimum variance scale identifies the optimal scale for capturing variability.
3. In Sect. 2.1, the manuscript should provide direct URLs or DOIs for all primary input datasets, particularly the MODIS snow products and Sentinel-1 RTC products, to improve reproducibility and accessibility.
4. The manuscript should more clearly explain whether the MOD10A2 8-day maximum snow extent product is appropriate for identifying snow phenology metrics such as onset and disappearance timing, given the compositing nature of the product.
5. The spatial matching strategy between 500 m MODIS data and 80 m Sentinel-1 observations should be described in greater detail, including any resampling, masking, or aggregation procedures.
6. Figure 1 is visually cluttered and somewhat redundant. The illustration of the 56-day snow duration criterion is also difficult to interpret and may not accurately communicate the intended logic of the filtering process.
6. Lines 177-178 mention the removal of "border noise artifacts," but the specific filtering approach is not described. A brief explanation should be added for reproducibility.

7. Line 181 should briefly define “Sentinel-1 relative orbit” for readers outside the SAR community.
8. Lines 185-186 introduce the 30-day orbit-gap filtering criterion, but the manuscript does not quantify how many relative orbits were removed by this requirement. This information would help readers assess the practical impact of the filtering.
9. The temporal resolution metric defined in Lines 200-202 is not equivalent to a true revisit interval and more closely represents an effective local sampling interval. The manuscript should clarify this distinction.
10. The manuscript states that 4,453 tiles containing both seasonal snow and Sentinel-1 RTC data were processed, but it does not indicate how many tiles contained seasonal snow without corresponding RTC coverage. This information would help clarify the true spatial limitations of the dataset.
11. The expressions “7,916 water years of snow pillow time series” (Line 245) and “4,763 water years” (Line 268) are awkward and technically imprecise. Terms such as “station-years” or “water-year observations” would be more appropriate.
12. Lines 283-284 contain somewhat repetitive wording regarding annual and composite products and could likely be simplified.
13. In Line 346, the phrase “of residuals” appears unnecessary and may be redundant.
14. The Discussion section is substantially longer than typical for a dataset-oriented paper. Sections 5.2 and 5.3 in particular contain considerable overlap between detailed interpretation of controls and evaluation discussion, making the narrative somewhat repetitive. Condensing these sections would improve readability and strengthen the overall focus of the manuscript.
15. The term “unprecedented” appears repeatedly throughout the manuscript (e.g., unprecedented spatial detail, unprecedented coverage, unprecedented characterization). While the dataset clearly represents a significant advance, repeated use of superlative language may overstate the contribution. Consider using more neutral and precise descriptions where appropriate.