Response to Reviewer #4 (additional reviewer)

Manuscript ID: ESSD-2025-694 *Earth System Science Data*

Title: A global gridded dataset of significant wave height via fusion of multi-mission altimetry and numerical hindcast

Authors: Haoyu Jiang (Haoyujiang@szu.edu.cn), Hao Su

While the authors present a high-resolution gridded product, the scientific justification is weak. By bypassing spectral information, the product lacks the physical consistency. Furthermore, given that established reanalyses (ERA5) already assimilate this data at the spectral level, and existing climate-standard products (ESA CCI, Ribal/Young) already address temporal consistency, this manuscript offers a redundant statistical patch rather than a significant scientific advancement. The risk of promoting circularity by using this product to train future AI models outweighs the convenience of having a pre-fused grid.

We thank the reviewer for the thorough and critical evaluation of our manuscript. The comments have helped us clarify the motivation, scope, and limitations of the proposed dataset.

We would like to first clarify the fundamental objective of this study. The proposed product is not intended to replace physically based wave re-analyses, modify the underlying numerical wave model, or provide a corrected full-spectrum wave reanalysis. Instead, it aims to construct a globally complete, hourly gridded SWH dataset by combining the complementary advantages of satellite altimeter observations and numerical hindcasts.

Satellite altimetry provides highly accurate local SWH observations but suffers from severe spatiotemporal undersampling. Conversely, numerical wave models provide physically consistent and spatially complete fields but contain model-dependent uncertainties. The purpose of the proposed offline fusion is therefore to reduce the sampling limitations of satellite observations while constraining the model background toward observations. The resulting dataset should be interpreted as an observation-constrained gridded SWH analysis, rather than an improved along-track satellite product or an improved numerical wave model.

We have revised the manuscript to clarify this scope and have added additional discussion regarding the limitations of SWH-only correction, the relationship with existing reanalysis products, and appropriate applications of the dataset.

1. Lack novelty (ERA5 redundancy)

The primary scientific justification for a new data product is that it fills a gap or offers a significant improvement over existing, established datasets.

->ERA5 (and the upcoming ERA6) already assimilates satellite altimetry (using 4D-Var or similar frameworks) to produce a globally gridded SWH product.

->If the authors' "fusion" is simply a statistical post-processing of a hindcast, it is effectively a Reanalysis-lite. The manuscript must demonstrate that their product significantly outperforms ERA5-Wave in a way that isn't just fitting the observations closer (which is trivial with ML). If the improvement is marginal, the community is better served by the physically consistent ERA5 than by a fragmented, offline statistical product.

We thank the reviewer for this important comment. We fully agree that ERA5 and other wave re-analyses represent major achievements in global wave reconstruction and provide valuable, dynamically consistent wave products. However, we respectfully disagree with the premise that the present dataset is redundant unless it universally outperforms ERA5. The two products are designed for different scientific objectives and involve substantially different treatments of satellite observations.

First, we would like to clarify the data-assimilation method used for ocean waves in ERA5. Although the atmospheric component of ERA5 is produced using 4D-Var, altimeter SWH observations in the wave component are assimilated using a sequential optimal interpolation (OI) scheme, rather than 4D-Var. Therefore, the reviewer's reference to variational method does not apply to the assimilation of altimeter SWH observations.

Second, the proposed dataset is not generated using machine learning, and no AI-based fitting is involved at any stage. It is produced using an explicitly defined space-time fusion method that combines a non-assimilative WW3 hindcast background with jointly calibrated multi-mission altimeter observations. Consequently, the characterization of the method as trivially fitting observations using ML is not applicable to the present study.

Third, ERA5 and the proposed dataset have different objectives. ERA5 aims to generate a dynamically consistent wave reanalysis through sequential model integration and assimilation. The proposed dataset instead aims to produce a retrospective, observation-constrained gridded SWH reconstruction. The offline framework can use observations located both before and after the target time within the specified fusion window. In contrast, the sequential OI system does not retrospectively revisit an already produced historical state using observations from later assimilation windows. This difference is important for historical reconstruction because the offline method can more fully exploit the available satellite sampling around each target time, although it is clearly not intended for operational forecasting.

Fourth, the two approaches treat satellite biases differently. In the ECMWF wave-assimilation system, bias correction is applied to harmonize altimeter SWH observations with their model counterparts before the OI analysis. This procedure is appropriate for maintaining a stable and dynamically consistent assimilative system. However, because the correction is defined relative to the model background, it may, in principle, attenuate observational signals that systematically differ from the model baseline. In the present study, the altimeter observations are jointly calibrated across satellite missions independently of the WW3 background and are not regressed toward the background model climatology before fusion. This distinction is particularly relevant when constructing a long-term SWH record intended to retain observational variability.

We therefore do not consider a universal reduction in errors relative to ERA5 to be the sole criterion for the scientific value of the dataset. ERA5 prioritizes dynamical and spectral consistency, whereas the present product prioritizes the direct incorporation of jointly calibrated altimeter SWH information into a spatially and temporally continuous gridded record. These are complementary rather than competing objectives.

Accordingly, the proposed dataset should not be interpreted as a simplified replacement for ERA5 or as a “Reanalysis-lite”. It is an alternative observation-constrained SWH product that bridges the gap between accurate but sparsely sampled along-track altimeter observations and continuous but model-dependent numerical hindcasts.

2. Physical inconsistency

The authors explicitly state that their method bypasses the need for wave spectral information. To expert users, this might be a major issue.

->SWH is an integrated parameter of the 2D wave spectrum

->By correcting only the SWH using a black box statistical method, the authors create a physical mismatch between SWH and all other spectral parameters (Mean Period, Peak Direction, etc.). A dataset where the SWH is "corrected" but the underlying spectrum remains tied to the original erroneous hindcast could be dangerous… users might perform secondary analyses (like wave power or stress calculations) that are physically

impossible. There is not much discussion about this or analysis to show that they thought it through.

Yes, SWH is an integral parameter and a corrected SWH should not be arbitrarily combined with uncorrected spectral parameters and interpreted as a physically self-consistent wave state. However, we would like to clarify two important points regarding the present dataset.

First, the proposed method is not a black-box statistical correction. It is an explicitly formulated space-time optimal interpolation method, and no machine-learning or AI model is involved in generating the dataset at all.

Second, the released product contains only total SWH. It does not provide a corrected directional spectrum, wave period, wave direction, or spectral partitions. Therefore, the dataset itself does not contain an internally inconsistent combination of corrected SWH and uncorrected spectral variables. Such an inconsistency would arise only if a user combined the fused SWH with spectral parameters from the original hindcast and interpreted them as a jointly corrected wave state, which is outside the intended scope of the product.

The decision not to modify the spectrum is deliberate. Altimeter observations provide total SWH but do not uniquely determine the underlying two-dimensional wave spectrum. Reconstructing or adjusting the full spectrum from SWH alone is an underdetermined problem and necessarily requires additional assumptions that may introduce further uncertainty.

In principle, the fused SWH field could subsequently be used to adjust the model spectrum using an approach similar to that adopted in the ECMWF wave assimilation system. In that system, an analyzed SWH field is first obtained through optimal interpolation, after which the background spectrum is rescaled in energy and frequency using different assumptions for wind-sea and swell components. However, such spectral adjustment requires additional assumptions regarding spectral shape and wave-system composition and may introduce uncertainties beyond those associated with the SWH analysis itself. Developing and evaluating a physically consistent spectral correction is therefore beyond the scope of the present SWH dataset.

We nevertheless agree that the intended scope of the product and the implications for multi-parameter applications should be stated more clearly. We have therefore added a caution in the revised manuscript. The revised text clarifies that the fused SWH may be used jointly with wave parameters from other products, as is common in multi-source analyses, but that these variables have not been jointly adjusted within the present fusion framework. Users should therefore evaluate the relevant consistency assumptions and uncertainties when their applications require strict internal consistency among SWH, wave period, wave direction, and the underlying directional spectrum:

“Moreover, this dataset is intended for analyses based only on total SWH, thus, other wave parameters from WW3 are not jointly adjusted within the present fusion framework, caution is warranted in applications that require strict internal spectral consistency among SWH and other wave parameters. For such applications, the consistency assumptions and associated uncertainties should be evaluated according to the objective.”

3. Circularity

One of the stated goals is to provide a multi-sat version for training future AI wave models.

->ML should be trained on the best possible representation of reality (observations) or high-fidelity physics

->If we use an AI-corrected product to train other AI models, we are creating an endless loop. We risk training future models on the artifacts and biases of the authors' specific statistical correction rather than on true ocean physics. This data incest can lead to overconfident models that fail when faced with non-linear extreme events that were not in the training fusion logic (they are smoothed out or highly distorted).

We thank the reviewer for raising the general concern that uncertainties in a training dataset may be inherited by downstream AI wave models. However, the concern expressed here is based on an incorrect characterization of the proposed product as an “AI-corrected product”. We reiterate that no AI method is used at any stage in generating the fused dataset. The product is constructed using an explicitly defined space-time optimal interpolation framework that combines jointly calibrated altimeter observations with a non-assimilative WW3 hindcast. Therefore, using this dataset to train an AI wave model does not constitute an AI-to-AI feedback loop or the recursive use of an AI-generated product.

More fundamentally, even if an AI method were used for the data-fusion step, which is not the case in the present study, this would not by itself constitute an AI-to-AI feedback loop. An observation-fusion model and an AI wave model solve fundamentally different problems. A fusion model estimates an analysis field from sparse observations and a background field, whereas an AI wave model predicts the wave field from atmospheric forcing and, depending on the model formulation, sometimes an initial wave state. Their inputs, objectives, and learned mappings are therefore different. The use of an observation-constrained analysis as a target for a prognostic AI model is not inherently circular merely because both procedures may employ AI.

The more relevant scientific concern is instead whether residual errors or artifacts in the fused product may be inherited by the downstream model. We agree that this possibility must be considered. However, the same issue applies to any AI model trained using reanalysis, analysis, remotely sensed Level-4 products, or numerical hindcasts; none of these products constitutes an error-free representation of reality.

Training AI wave models often requires spatially and temporally complete target fields. Direct along-track altimeter observations are highly valuable but are too sparse to provide complete target fields for supervised global wave-field modeling. The fused dataset is intended to bridge this gap by providing an observation-constrained gridded SWH field. This is conceptually similar to the widespread use of reanalysis or analysis products as training targets, although the present product is more directly constrained by jointly calibrated altimeter observations. The supplied distance-related variable R also indicates the proximity of each grid point to the nearest altimeter observation and can be used to assign greater training weight to samples that are more directly constrained by observations.

We acknowledge that the fused dataset should not be regarded as an error-free representation of reality. Like any observation-constrained analysis or reanalysis, it may retain residual uncertainties associated with the background field, observational sampling, and the assumptions of the fusion method. AI models trained using this product should therefore be independently evaluated against withheld satellite and buoy observations, particularly for extreme sea states and conditions that are poorly represented in the training record.

4. Time-space issues

The authors mention that their offline approach allows retrospective correction of past model outputs using future observations.

->While mathematically sound for gap-filling, this is physically problematic for wave climate analysis.

->Waves propagate… using future satellite passes to correct a past state in a non-causal way can smear out the temporal evolution of swell events. For wave climate trends (their two-sat version), this smoothing might hide the very variability or subtleties the researchers are looking for…

We thank the reviewer for raising the concern regarding the preservation of the temporal evolution of swell events. However, we respectfully disagree that the use of observations acquired after the target time is inherently non-physical or necessarily causes temporal smoothing in an offline historical reconstruction. Temporal smoothing may occur when observations representing different stages of an evolving event are blended using excessively broad or insufficiently constrained weights. The relevant issue is therefore how the observational influence is formulated, rather than whether an observation is acquired before or after the target time.

The use of later observations to constrain an earlier state is well established in retrospective estimation. For example, 4D-Var adjusts the model state at the beginning of an assimilation window so that the resulting trajectory is consistent with observations distributed throughout that window. Consequently, observations acquired later in the window can constrain an earlier model state through the model dynamics. We do not suggest that the present method is equivalent to 4D-Var. This example simply demonstrates that the use of later observations is not inherently physically invalid when their influence is appropriately constrained.

In the present method, temporal separation is explicitly incorporated into the distance metric. Observational influence progressively decreases with increasing temporal separation, and observations separated from the target time by more than 24 h are excluded. The influence is further constrained by spatial separation, the difference in modelled SWH between the target and observation locations, and the correlation between their background SWH time series. These constraints are intended to reduce the contribution of observations representing unrelated wave systems or substantially different stages of the same event.

Moreover, the dynamically evolving WW3 hindcast provides the background spatial and temporal structure, including the modelled generation, propagation, and decay of swell systems. The altimeter observations provide constrained correction increments to this evolving background rather than replacing its temporal evolution through direct interpolation of satellite measurements. Therefore, the potential for smoothing would arise if temporally distant or dynamically inconsistent observations were assigned excessive weights, which the present formulation is specifically designed to avoid.

Finally, the fused product is evaluated against instantaneous independent buoy measurements and withheld altimeter observations, rather than only against monthly means or climatological statistics. These independent evaluations show that the fusion improves instantaneous SWH accuracy relative to the original hindcast. The results therefore do not support the characterization of the method as producing an excessively smoothed reconstruction of evolving wave fields.

5. Climate trend justification

The authors claim their "two-sat" version is for climate studies to ensure temporal consistency.

->The ESA CCI Sea State dataset already exists for this exact purpose, using a very rigorous, inter-calibrated multi-mission approach.

->The manuscript should address: Why is this better than the ESA CCI product? If the answer is just "it's on a grid," that is a processing step, not a significant contribution… Gridding can be done by anyone with a basic interpolation script; it doesn't necessitate a new global dataset publication.

We fully agree that the ESA CCI Sea State dataset is a rigorous, inter-calibrated multi-mission observational record developed specifically to support climate applications. Indeed, the present fused dataset is built directly upon the CCI Sea State observations and should be regarded as a complementary product rather than a replacement for the original CCI record.

The term "temporal consistency" was not sufficiently defined in the original manuscript. In the CCI Sea State dataset, temporal consistency primarily refers to the calibration and harmonization of measurements from different satellite missions. In the present two-satellite product, it refers instead to the consistency of the sampling configuration and observational influence used in the reconstruction. Inter-calibration can reduce systematic differences between missions, but it cannot eliminate the temporal inhomogeneity caused by changes in the number of operating satellites, their orbital configurations, and the associated probability of sampling individual storms and extreme sea states. This distinction is particularly relevant for extreme-wave statistics. As the number of concurrent altimeters increases over time, the probability that short-lived extreme events are sampled also increases. Consequently, even a perfectly inter-calibrated observational record may exhibit time-dependent sampling characteristics. The effects of this temporally uneven sampling have been systematically investigated in a previous study by the first author (Jiang 2020).

The two-satellite configuration is designed to reduce this source of inhomogeneity by maintaining a relatively stable number of observations and a broadly consistent combination of orbital sampling patterns throughout the record. Specifically, one approximately 66°-inclination mission is paired with one near-polar mission, with replacements selected from satellites having similar orbital characteristics when necessary.

The two-satellite configuration is designed to reduce this source of inhomogeneity by maintaining a relatively stable number of observations and broadly consistent orbital sampling characteristics throughout the record. Specifically, one mission with an orbital inclination of approximately 66° is paired with one near-polar mission, and replacements are preferentially selected from satellites with similar orbital characteristics when necessary. We agree, however, that the original wording "ensuring temporal consistency" was too strong. The two-satellite configuration cannot guarantee that the resulting dataset is free from all temporal inhomogeneities or that estimated climate trends are unbiased. Residual uncertainties associated with cross-mission calibration, observational noise, and the numerical background remain. We have therefore revised the wording to state that the two-satellite product is designed to reduce sampling-related temporal inhomogeneity:

Abstract:

"A 'two-sat' version that incorporates data from only two satellites at any given time, designed for wave climate studies. This configuration is designed to reduce sampling-related temporal inhomogeneity by maintaining a relatively stable number of observations and broadly consistent orbital sampling characteristics throughout the record."

Section 2.3 Two Types of Fused Data:

"The 'all-sat' product prioritizes accuracy by maximizing observational coverage, whereas the 'two-sat' product is designed to reduce sampling-related temporal inhomogeneity by maintaining a relatively stable number of observations and broadly consistent orbital sampling characteristics. Here, sampling consistency refers to reducing temporal changes in observation density and orbital sampling characteristics and is distinct from the cross-mission calibration consistency already provided by the CCI Sea State dataset."

"For the 'two-sat' product, the selection of satellites is guided by sampling characteristics: one satellite with an orbital inclination of approximately 66° is paired with another in a near-polar orbit to maintain broadly consistent orbital sampling characteristics over time"

Summary:

"… a 'two-sat' version, which incorporates data from only two satellites at any given time and is optimized for SWH-based wave-climate studies, ensuring broadly consistent orbital sampling characteristics throughout the record by reducing sampling-related temporal inhomogeneity"

We also agree that simply converting along-track observations onto a grid would not, by itself, constitute the principal scientific contribution of this study. However, the proposed dataset is not produced through basic geometric interpolation or coordinate remapping. The central difficulty is not the grid format, but the reconstruction of wave conditions that were never sampled by an altimeter. Because ocean waves vary rapidly in both space and time, direct interpolation of sparse satellite tracks cannot recover storms or swell systems that fall between tracks or between satellite overpasses. Increasing the sophistication of the interpolation algorithm does not create the missing observational information.

The present method therefore uses a dynamically evolving WW3 hindcast as a spatially and temporally complete background field and applies observation-derived SWH increments through a constrained space-time optimal interpolation framework. The observational weights depend not only on geographical distance, but also on temporal separation, modelled-SWH similarity, and the correlation between the target and observation locations. The resulting product is therefore a model-assisted, observation-constrained reconstruction rather than a conventional gridding of the CCI observations.

The importance of this distinction is supported by the validation results. In the buoy comparison, the altimeter-only gridded product considered in the manuscript had an RMSE of 0.37 m, which was higher than the 0.31 m RMSE of the non-assimilative WW3 hindcast, while the fused dataset had an RMSE of 0.27 m and performed better for extreme cases. These results show that gridding sparse altimeter observations alone does not necessarily produce a more accurate spatially complete SWH field.

We also recognize that CCI provides a monthly Level-4 gridded product, and we have clarified this more explicitly in the manuscript. The CCI Level-4 product and the present fused dataset nevertheless remain fundamentally different. The former represents statistics derived from the satellite samples available within each grid cell and time interval, whereas the latter provides a spatially complete time series constrained by both the numerical background and the satellite observations. In the comparison presented in the manuscript, residual ground-track patterns remain visible in the monthly CCI Level-4 field, and the differences become particularly pronounced for upper-percentile SWH because of the limited probability of sampling individual storms

The CCI Sea State record remains the fundamental observational source of this dataset, and is more appropriate when direct measurement traceability and along-track variability are required. The added value of the present product is that it uses these high-quality CCI observations to construct and independently validate a spatially complete, temporally resolved SWH reconstruction. The two-satellite version specifically aims to reduce changes in the degree and geometry of observational constraint over the multi-decadal record.

Ref:
Jiang, H. (2020). Evaluation of altimeter undersampling in estimating global wind and wave climate using virtual observation. *Remote Sensing of Environment*, *245*, 111840.

6. More broadly...

->Scientific progress comes from fixing the source terms (Sin, Sds) in models like WAVEWATCH3.

->This work treats the model as a "black box" that is "wrong" and needs a statistical patch. In a data journal, the goal should be to provide data that reveals new truths. If the corrected data simply masks model bias without explaining it, it serves as a band-aid that discourages researchers from doing the hard work of improving model physics.

We fully agree that the continued development of source terms, including wind-input and wave-dissipation parameterizations, is important for advancing numerical wave modelling. However, we respectfully disagree with the premise that progress in wave science and engineering can arise only from fixing the source terms of numerical wave models, or that observation-model fusion and data assimilation are inherently opposed to improvements in model physics.

Numerical model development, satellite observations, in situ measurements, data assimilation, data reconstruction, and model evaluation are complementary rather than competing scientific activities. In particular, reliable observational datasets are indispensable for identifying systematic model deficiencies, evaluating alternative physical parameterizations, and determining whether a proposed change actually improves model performance. Improving model physics and producing more accurate reconstructions of the historical ocean state therefore address different scientific objectives and can support one another.

The characterization of WW3 as a “black box” or a “wrong” model in the present study is incorrect. The WW3 hindcast is used precisely because its physics-based integration provides a spatially and temporally complete representation of the modelled generation, propagation, and dissipation of wave systems. The manuscript explicitly describes the source-term configuration, atmospheric forcing, spatial and temporal resolution, and validation accuracy of the WW3-ST6 hindcast. We also conduct a background-sensitivity experiment using the substantially different WW3-ST4 hindcast forced by CFSR winds. As stated in the manuscript, the accurate but sparse altimeter observations and the dynamically evolving numerical background are both indispensable components of the fusion framework.

More fundamentally, the fusion method relies on a physically credible and skillful background field. In regions and periods where direct observations are sparse or distant, the fused estimate necessarily retains a larger contribution from the numerical hindcast. The accuracy and physical realism of the background therefore directly affect the quality of the fused product: a more accurate background will generally lead to a more accurate reconstruction, whereas a random, physically implausible, or severely biased field could not serve as a meaningful background. Our motivation is therefore to adopt a background field that is as “right” as possible, i.e., physically credible, skillful, and independently validated, rather than to assume that a “wrong” field can simply be repaired.

At the same time, no numerical wave simulation is expected to reproduce observations perfectly. Model-observation discrepancies may arise from multiple sources, including atmospheric forcing errors, source-term parameterizations, spatial resolution, numerical propagation, bathymetric representation, and omitted processes such as wave–current interactions. Determining the physical origin of each discrepancy and improving the corresponding model physics are important research objectives, but they are fundamentally different from reconstructing the most accurate historical SWH field possible. The present study addresses the latter objective.

The proposed product should also not be characterized as a simple statistical correction that merely masks model bias. Before fusion, we apply a one-dimensional look-up-table adjustment to remove the mean systematic difference between WW3 and altimeter SWH. The manuscript evaluates this step separately. It reduces the RMSE by only approximately 0.01 m and increases the correlation by only 0.001, demonstrating that simple bias adjustment contributes negligibly to the overall improvement. In contrast, the full space–time fusion reduces the RMSE against withheld Jason-2 observations from approximately 0.34 to 0.23 m and increases the correlation to 0.985. The improvement therefore cannot be explained as merely masking a mean model bias.

Furthermore, the fusion does not erase or conceal the original model information. The original WW3 hindcast and CCI Sea State observations remain independently available, while the fused field is provided as an additional derived product. Differences between the original hindcast and the observation-constrained reconstruction can themselves be analysed to identify where and when the numerical model systematically departs from observations. The dataset may therefore facilitate, rather than discourage, subsequent investigations of atmospheric forcing, source terms, and other model deficiencies.

The background-sensitivity experiment further demonstrates that the method is not an empirical correction tailored to a particular model configuration. WW3-ST4 and WW3-ST6 differ in both source-term parameterizations and atmospheric forcing. Although the fused result remains dependent on the background, as any model-assisted reconstruction must, the discrepancies between the two fused products are substantially smaller than the undersampling errors obtained when no physically meaningful background is used. This shows that the method requires a physically reasonable background but is not critically tied to one specific source-term package.

We respectfully disagree that the scientific value of a data paper depends on whether the dataset independently explains the physical origin of every corrected discrepancy. The purpose of the present dataset is to provide a spatially complete, observation-constrained reconstruction of historical SWH that mitigates both errors inherited from the numerical background and errors due to satellite undersampling. Such a product has direct value for wave-climate analysis, extreme-value assessment, model evaluation, training AI wave model, and applications that cannot be adequately supported by either sparse along-track observations or a standalone numerical hindcast. It is not presented as a replacement for physics-based model development, nor as evidence that further improvements to WW3 source terms are unnecessary.

Finally, we acknowledge that observation-model fusion and data assimilation only provides a pragmatic means of improving historical reconstructions and wave forecasting while numerical wave models remain imperfect. However, we see no basis for the suggestion that such products discourage improvements in model physics. On the contrary, the quality of the fused product depends positively on the quality of its numerical background, and improvements in source terms, atmospheric forcing, propagation, and other physical processes would directly improve future versions of the reconstruction.

Overall, we appreciate the reviewer's concerns regarding the physical consistency, potential applications, and limitations of the proposed dataset. However, several of the criticisms appear to arise from a different interpretation of the purpose and implementation of the study from that intended in the manuscript. The proposed product is neither generated by AI nor intended to replace physics-based wave modelling, the original CCI Sea State observations. Rather, it is an offline, model-assisted and observation-constrained reconstruction of total SWH, designed to address the complementary limitations of sparse altimeter sampling and imperfect but spatiotemporally complete numerical hindcasts. The fusion method is explicitly formulated, its dependence on the numerical background is evaluated, and its performance is assessed against independent buoy and withheld satellite observations.

We have carefully reconsidered all of the reviewer's comments and have revised several expressions where the original wording could have been overly broad or insufficiently precise, particularly regarding the interpretation of temporal consistency and the intended use of the two-satellite product. At the same time, we believe that the main methodological and scientific concerns raised in the comments do not reveal fundamental flaws in the dataset, but largely reflect misunderstandings of its scope, methodology, and intended applications. We hope that the detailed responses above clarify these distinctions and provide a more accurate basis for evaluating the scientific contribution and appropriate use of the dataset.