**Dear Handling Topic Editor (Yuntao Zhou) and Editor-in-Chief:**

**Re: essd-2026-341**

We sincerely appreciate your time and effort in reviewing our manuscript. Your critical comments and insightful suggestions have been invaluable in improving the quality and clarity of our work. We have carefully addressed all the concerns raised during this round of review. In the revised manuscript, all modifications made in response to the reviewers' comments are highlighted in blue for easy identification. Additional experiments, analyses, or clarifications have been incorporated as suggested. The manuscript has been thoroughly proofread to enhance its rigor and readability. Below, we provide point-by-point responses to the reviewers' specific questions. Should further revisions be required, we are committed to addressing them promptly.

**Reviewer 1**

**This manuscript uses a physics-informed deep learning approach to reconstruct a global daily sea surface ageostrophic current dataset for 1993-2023. The topic is meaningful, since many existing surface current products are mainly based on geostrophic currents, or geostrophic currents with Ekman corrections, while ageostrophic processes, especially nonlinear ageostrophic motions, are less well represented. The attempt to introduce momentum-balance constraints into a neural-network framework and generate a long-term global dataset is valuable.**

**Major Comments**

**Q1: In Section 4.1, the observational validation is based on only one TAO mooring site. This example is helpful, but its representativeness is limited. Additional observational sites from different regions (such as polar regions, equatorial regions, and coastal areas), or other independent observational data, would better demonstrate the applicability of the dataset under different dynamical regimes. Whether the errors exhibit systematic differences across different ocean regions should be clarified.**

**Response to Q1:**

We sincerely thank the reviewer for this constructive comment. Following your suggestion, we have systematically expanded our validation against 54 in-situ observations from the TAO, PIRATA, and RAMA

arrays (https://climatedataguide.ucar.edu/climate-data/tropical-moored-buoy-system-tao-triton-pirata-rama-toga). Besides, the widely used OSCAR dataset was also included in the comparison.

**1. Overall error analysis and expanded validation**

As shown in Figure R1(a), those 54 in-situ observations stations are widely distributed across the Pacific (21 stations), Atlantic (14 stations), and Indian Oceans (19 stations). Figure R1(b)–(d) display the Root Mean Square Error (RMSE) of the AC-PIDNN-corrected total currents, OSCAR currents, geostrophic currents (Geo), and Ekman corrected currents (Ekman) against the in-situ measurements. The statistical results are as follows:

- **Pacific Ocean:** The RMSE of the AC-PIDNN-corrected total current is 0.16 m/s, which outperforms both the Ekman corrected (0.24 m/s) and Geo (0.22 m/s) currents. Furthermore, compared to the standard OSCAR product's error of 0.19 m/s against in-situ data, the AC-PIDNN product reduces the error by approximately **15.8%**.
- **Indian Ocean:** The RMSE of our corrected current is 0.10 m/s, outperforming Ekman corrected (0.13 m/s) and Geo (0.15 m/s) currents. Compared to the OSCAR product's error of 0.13 m/s, the AC-PIDNN product reduces the error by approximately **23.1%**.
- **Atlantic Ocean:** The RMSE of our corrected current is 0.15 m/s, outperforming Ekman corrected (0.22 m/s) and Geo (0.17 m/s) currents. Compared to the OSCAR product's error of 0.17 m/s, the AC-PIDNN product reduces the error by approximately **11.8%**.

Overall, across all three major ocean basins, the corrected surface circulations are closer to the in-situ observations compared to the traditional geostrophic and Ekman corrected components, and they consistently demonstrate improved accuracy compared with OSCAR products.

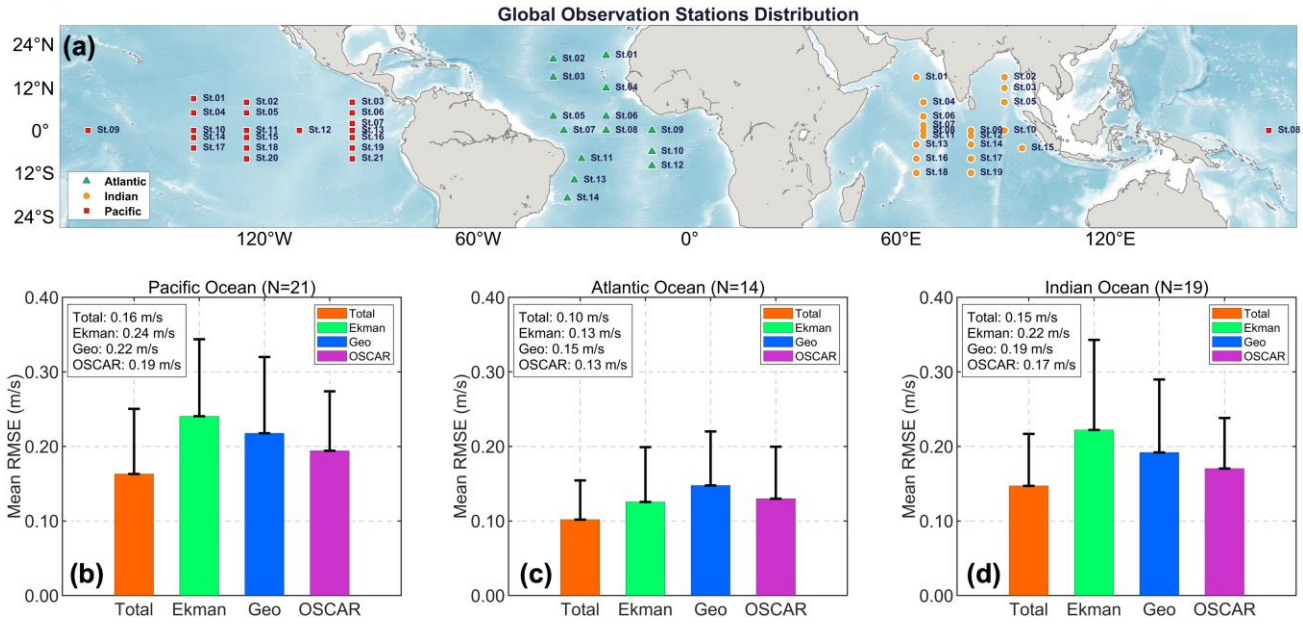



**Figure R1.** (a) Spatial distribution of TAO/PIRATA/RAMA mooring sites across the Pacific, Atlantic, and Indian Oceans. In-situ observation stations in each ocean basin are individually numbered to facilitate subsequent comparative experiments. (b)–(c) Bar charts illustrating the averaged root-mean-square error (RMSE) between TAO/PIRATA/RAMA mooring observations, OSCAR products and the AC-PIDNN-corrected current fields, geostrophic currents, and Ekman corrected currents for the Pacific, Atlantic, and Indian Oceans, respectively.

**2. Error distribution characteristics and systematic differences**

To address your query regarding whether the errors exhibit systematic differences across different ocean regions, we conducted an in-depth latitudinal analysis using the in-situ data, as shown in Figure R2. The results reveal that the dataset's error characteristics do exhibit spatial differences:

- Larger errors are primarily concentrated near the equator (within a latitude band of approximately ±5°).
- Conversely, accuracy significantly improves in mid-to-low latitudes, aligning much more closely with field observations. Specifically, the RMSE gradually decreases from 0.25 m/s near the equator down to 0.05 m/s in these higher-latitude observation areas.

Overall, within the ±25° latitude range, the AC-PIDNN-corrected product consistently remains closer to the field observations than the OSCAR product. Regarding the valuable suggestion to include high-latitude, it is important to note that there is a lack of long-term continuous surface observations. We have explicitly

acknowledged this limitation regarding high-latitude applicability in the **section 4.1.2** of the revised manuscript. We once again thank you for your valuable and constructive comments.

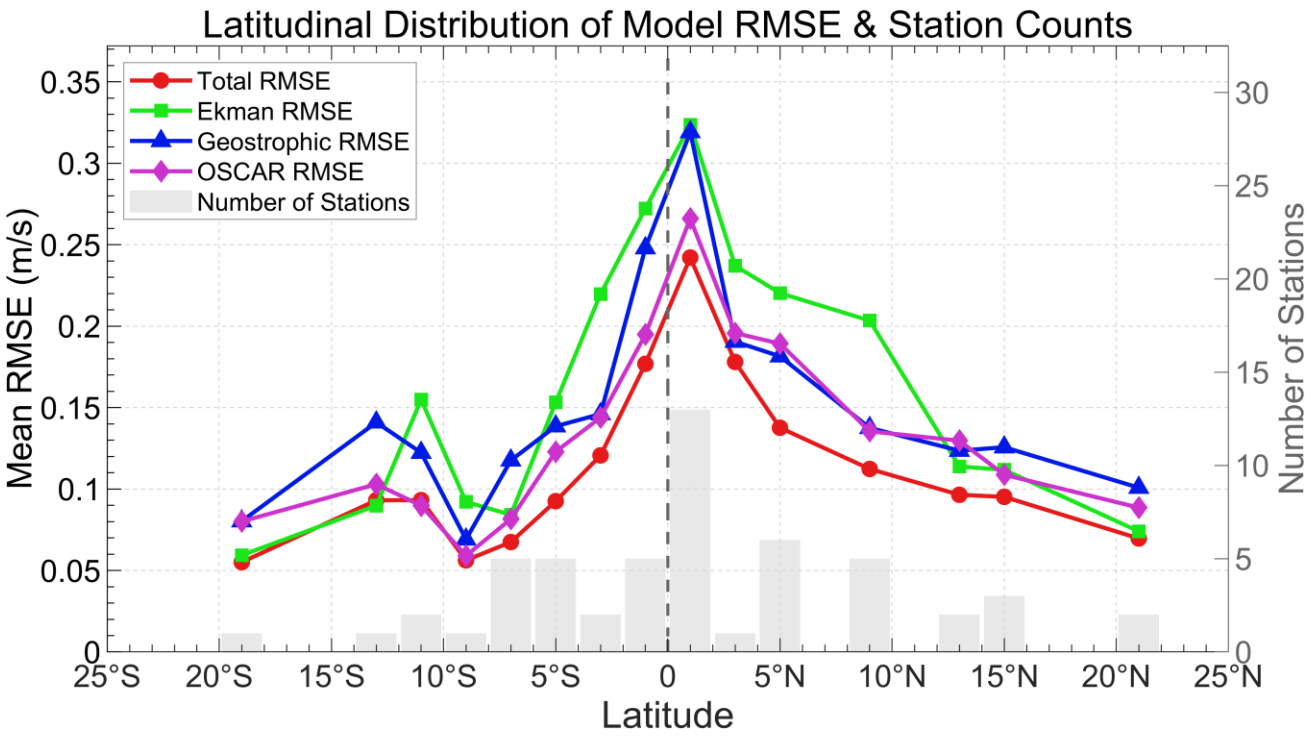



**Figure R2** presents the latitudinal error curves of the in-situ measurements against both the corresponding AC-PIDNN-corrected currents and the OSCAR products across the 54 observation stations.

**Q2: Comparisons with existing mature surface current products, such as OSCAR, or other commonly used products, would strengthen the manuscript. At present, the manuscript mainly compares AC-PIDNN with geostrophic currents, Ekman-corrected currents, an iterative method, and ROMS simulations. However, readers will likely want to know how this new dataset differs from, or improves upon, existing surface current products. Such comparisons would help clarify the actual added value of the proposed dataset.**

**Respond Q2:**

We sincerely thank the reviewer for this constructive comment. Besides the comparison shown in Response 1, we further examined the circulation pattern and divergence with the OSCAR.

We compared the 30-year mean (1993–2022) total surface flow fields derived from OSCAR with those reconstructed by our method. Overall, the two datasets exhibit similar large-scale circulation patterns. However, because the PIDNN reconstruction additionally accounts for the nonlinear current component, its current magnitudes are generally larger than those in OSCAR and show better agreement with the observations (Figures R1–R2). The global mean total velocity for OSCAR is 0.16 m/s, compared to 0.19 m/s for AC-PIDNN. The differences are particularly pronounced in high-kinetic-energy western boundary current regions, such as the Kuroshio and Gulf Stream, equatorial region, and Southern Ocean. In these regions, AC-PIDNN more effectively reconstructs the strong currents and intense nonlinear advection characteristic of western boundary current systems.

The incorporation of nonlinear effects also leads to stronger surface divergence in AC-PIDNN than in OSCAR, although the two datasets retain broadly similar large-scale spatial patterns. As shown in **Figure R3**, the divergence field reconstructed by AC-PIDNN exhibits greater intensity and sharper spatial gradients, particularly in dynamically energetic regions. These differences indicate that AC-PIDNN provides an improved representation of multiscale surface-current dynamics and associated convergence–divergence structures. We have added these results to the Supplementary Material as Figure S3.

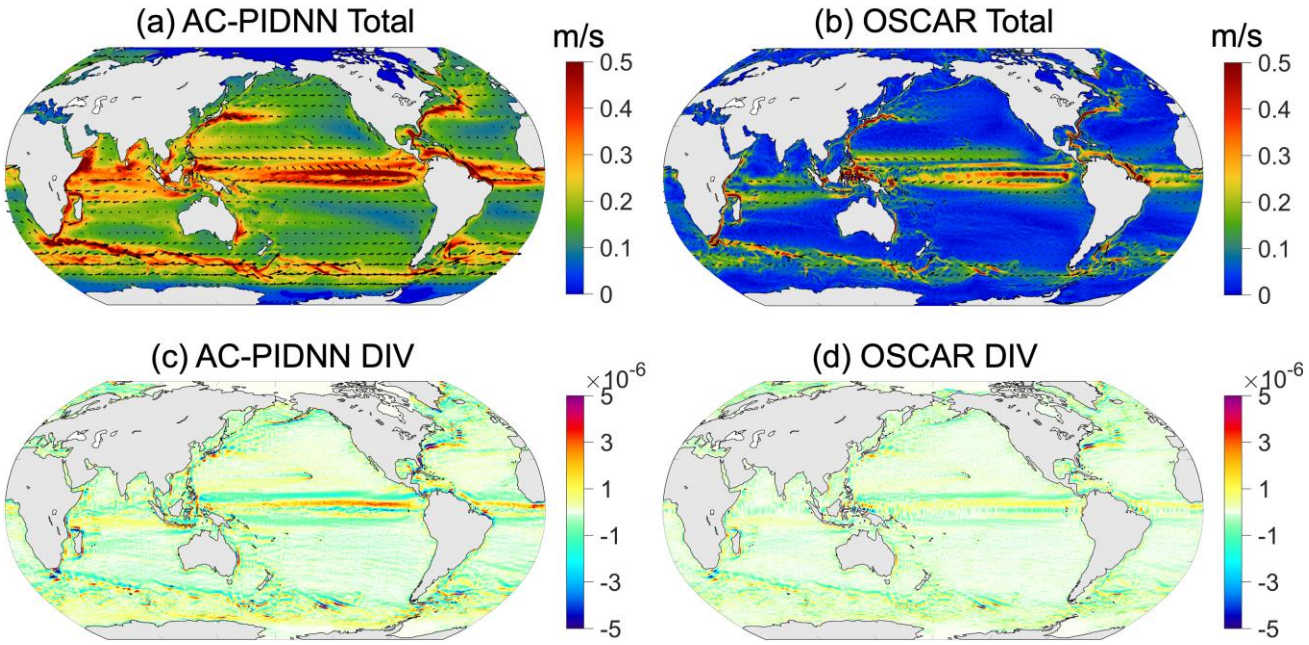


Figure R3. (a) 1993–2022 mean flow field from AC-PIDNN; (b) 1993–2022 mean flow field from the OSCAR product; (c) divergence of the 1993–2022 mean flow field from AC-PIDNN; (d) divergence of the 1993–2022 mean flow field from the OSCAR product.

**Q3: A simple ablation experiment would help show what the physics loss actually contributes. Since the first training stage already uses only Loss_data, the authors could simply save the Stage 1 output and use it as a data-loss-only baseline. Comparing this baseline with the final AC-PIDNN output using station-based observational errors and the Bal balance would make it clearer whether the later physics-informed stages provide a real improvement beyond the initial data fitting.**

**Respond Q3:**

We sincerely thank the reviewer for this insightful suggestion. To systematically address this, using the year 2021 as an example, we executed the suggested ablation experiment to examine the improvement by physics-informed stages. By comparing output from Stage 1, Stage 2 and Stage 3 against both theoretical momentum balance metrics (Bal) and station-based observational errors, we can clearly delineate the value added by the physical constraints beyond pure data-driven mapping.

**1. Impact on Momentum Balance (Bal)**

**Figure R4**(a)–(c) illustrate the annual mean surface current fields generated by the AC-PIDNN model during Stages 1, 2, and 3, respectively. Correspondingly, Figures R4(d)–(f) map the momentum balance residual (Bal) for these three stages. The progression reveals a stark contrast in dynamical consistency:

- **Stage 1 (Data-Loss Only Baseline):** Because this initial stage relies exclusively on $Loss_{data}$ without any embedded physical constraints, the output entirely lacks momentum balance correction, leading to severe dynamical inconsistencies where the momentum residuals exceed **100%** across the vast majority of the global ocean regions, as shown in **Figure R4** (a) and (d).
- **Stage 2 (Hybrid Loss Optimization):** Upon introducing a hybridized loss function that begins blending physical constraints with observational data, we observe a dramatic correction in the flow

field's dynamical structure. The global mean momentum residual drops precipitously to approximately **51%**, as shown in **Figure R4** (b) and (e).

- **Stage 3 (Fully Physics-Constrained Final Output):** In the final stage, the output is rigorously constrained by the complete physics loss. This proper parameterization forces the reconstructed flow field to obey the fundamental equations of motion, vastly improving the momentum balance and reducing the global average momentum residual to a mere **1.2%**, as shown in **Figure R4** (c) and (f).

This progression clearly demonstrates that the later physics-informed stages provide a profound improvement that extends well beyond initial data fitting. By enforcing momentum conservation on the purely data-driven output, the subsequent physics-informed stages produce a flow field that is both physically coherent and statistically accurate, thereby enhancing the hydrodynamic reliability of the reconstructed surface currents.

**2. Validation Against In-Situ Observational Errors and OSCAS products**

To confirm that this improved dynamical balance translates to real-world accuracy, we further evaluated the outputs of all three training stages, as well as the standard OSCAR product, against independent in-situ mooring data as shown in **Figure R5**. As expected, the progressive incorporation of physical laws systematically refines the accuracy of the modeled velocities:

- **Stage 1 Average Error:** 0.17 m/s
- **Stage 2 Average Error:** 0.16 m/s
- **Stage 3 Average Error:** 0.13 m/s
- **OSCAR Product Average Error:** 0.17 m/s

The data clearly shows that the Stage 3 output aligns most closely with the in-situ measurements, significantly outperforming the Stage 1 baseline, the intermediate Stage 2 output, and the OSCAR product.

This direct comparison with station-based observational errors firmly substantiates the superiority and absolute necessity of the physics-informed training stages. We have incorporated the comprehensive results of this ablation experiment, alongside the associated figures into the **Supplementary Material (Figure S1 to Figure S2)**. We once again thank you for your valuable and constructive comments.

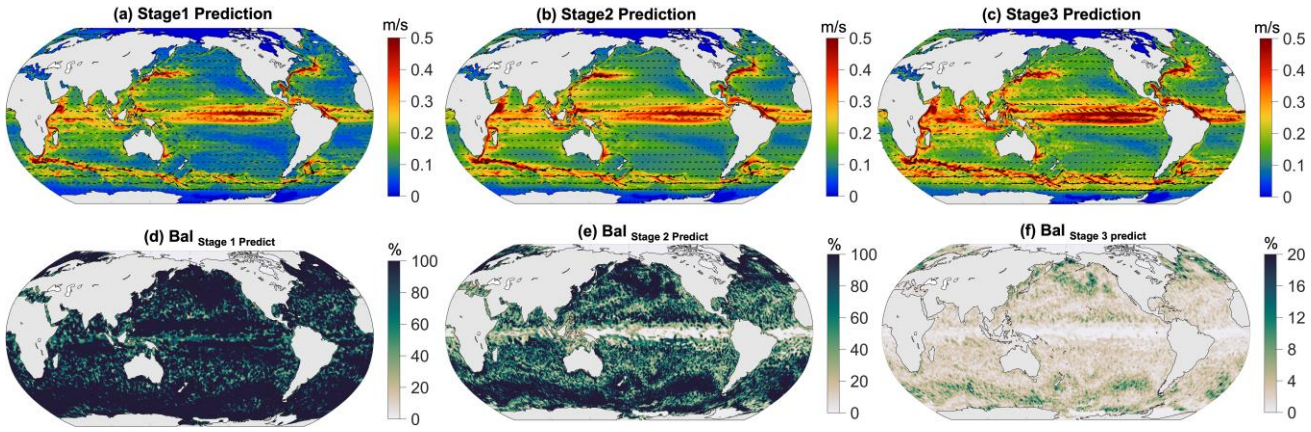



**Figure R4.** (a) Mean output current field of AC-PIDNN Stage 1 in 2021; (b) Mean output current field of AC-PIDNN Stage 2 in 2021; (c) Final mean output current field of AC-PIDNN Stage 3 in 2021; (d) Momentum balance (Bal) results for Stage 1; (e) Momentum balance (Bal) results for Stage 2; (f) Momentum balance (Bal) results for Stage 3.

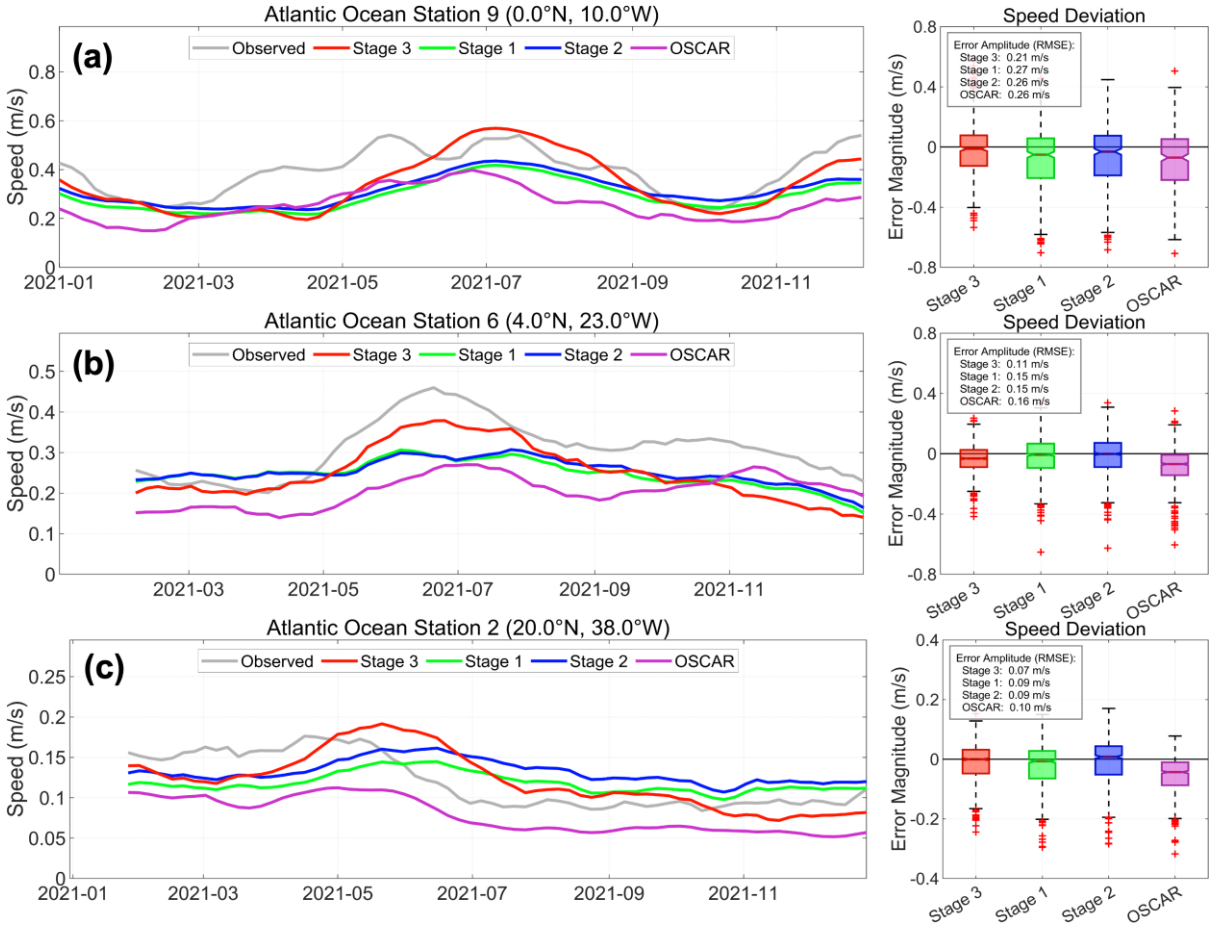



**Figure R5.** Comparison of the output current fields from the three training stages of AC-PIDNN and OSCAR product velocities against in-situ observational data.

**Q4: The data upload method on the Zenodo website may not meet the requirements of ESSD. Data must be submitted in a more appropriate manner and additional information must be provided regarding the format of the data set, file naming conventions, variable units, missing values, metadata, and whether the NetCDF files comply with the CF standard.**

**Respond:**

We sincerely thank the reviewer for highlighting these critical omissions. We fully agree that comprehensive metadata and strict adherence to data standards are essential for the usability and longevity of the dataset. We have thoroughly revised our Zenodo repository and the dataset's internal documentation to align with ESSD requirements. The updated repository is now available at: DOI: 10.5281/zenodo.19966716 Specifically, we have addressed the reviewer's concerns as follows:

**1. Data Upload Method and Accessibility Data Upload Method and Accessibility**

Due to the comprehensive 31-year span and high spatial resolution (0.25°) of our dataset, the total data volume exceeds 500 GB. Because this vastly surpasses Zenodo's standard per-upload storage limits, hosting the entire dataset directly on Zenodo in a single upload is technically unfeasible. To provide a practical and compliant solution, we have adopted a hybrid access approach:

- We have uploaded a **r**epresentative sample .nc file directly to the Zenodo repository so that users can immediately inspect the data structure, metadata, and variables.
- We have included a comprehensive README.txt file on Zenodo. This document details the exact directory structure and provides direct, high-speed download links to the complete 500 GB dataset hosted on a dedicated S3 server (https://s3.jn1.is.shanhe.com/umdr2026/).

**2. CF Standard Compliance and Metadata Enhancements CF Standard Compliance and Metadata Enhancements**

We have explicitly updated the Zenodo description and ensured our NetCDF files comply with the Climate and Forecast (CF) Metadata Conventions (CF-1.8). The repository page now clearly documents the following details:

- Data Format: NetCDF-4 Classic (netcdf4_classic).
- File Naming Conventions: Detailed the standardized directory and file naming structure ([YYYY]/[MM]/[YYYY]_[MM]_[DD].nc).
- Variable Details & Units: Provided a comprehensive table defining every variable (e.g., ugeo, uekman, uadv, utotal), including their standard long names, dimensions, and strict inclusion of scientific units (e.g., m/s, m/s2, m).
- Missing Values: Clarified the handling of missing data (e.g., over landmasses or un-computable grid points) by explicitly defining the use of the standard _FillValue = NaN indicator.
- Global Metadata: Outlined the inclusion of global attributes to track provenance, including creation_date, author, and Conventions.

We believe these updates provide the clarity, transparency, and standardization required by ESSD, ensuring the data is highly accessible and easily interpretable for future researchers.

## Specific Comments

**Q5: The method description is not yet sufficiently clear, especially regarding the training labels. Based on the code provided by the authors, the input variables described in the manuscript do not seem to fully match the inputs used in the code. Also, the training labels in the code appear to be a pseudo-target obtained by adding the geostrophic current and the Ekman current, rather than an independently observed total current. This approach may be reasonable, but it should be stated clearly in the method section. Otherwise, it is hard to understand how the model moves from a geostrophic + Ekman baseline toward a nonlinear ageostrophic correction.**

**Respond:**

We sincerely thank the reviewer for this meticulous and insightful observation. We fully agree that the original description of the model's inputs and training labels lacked the necessary clarity and transparency. We have carefully revised the methodology section to ensure strict alignment between the manuscript and the provided codebase, and to explicitly explain the rationale behind our training strategy.

Specifically, we have made the following modifications in the revised manuscript:

**(1) Alignment of Input Variables**

As the reviewer correctly pointed out, there was a discrepancy between the text and the actual code. We have thoroughly reviewed and rewritten the relevant paragraphs in the Methodology section (see **Section 3 in the revised manuscript**) to comprehensively list and define all input variables. The updated text now strictly mirrors the input features explicitly used in our deep learning code, eliminating any ambiguity.

**(2) Clarification of the "Pseudo-Target" Training Labels**

We highly appreciate the reviewer's acute observation regarding the training labels. We have now explicitly stated in the revised manuscript (see **Section 3.2.3 of the revise manuscript**) that the training labels are indeed "pseudo-targets" derived from the superposition of the geostrophic current and the Ekman corrected current, rather than dense, independent real-world observations (such as drifter buoy data).

**(3) Explanation of the Learning Mechanism for the Ageostrophic Correction**

To address the reviewer's valid concern about how the model learns the nonlinear ageostrophic correction from this baseline, we have added a dedicated explanation of our physical and algorithmic logic:

- We clarify that because high-resolution, global, and independent observations of total sea surface currents are too sparse to support AI model training, we employ the physics-derived baseline (geostrophic + Ekman) as the primary training target.
- We further elaborate on how the model utilizes these pseudo-targets in conjunction with the designed loss functions and physical constraints (e.g., incorporating the nonlinear advective terms and momentum residuals) to move beyond the linear baseline. This approach forces the model to capture

and output the non-linear ageostrophic corrections necessary to bridge the gap between the simplified linear dynamics and the actual complex flow field.

We believe these revisions directly address your concerns, significantly improving the methodological transparency, reproducibility, and physical interpretability of our model (**please refer to Section 3.2.3 in the revised manuscript and Figure S1 in the supplementary material**). Thank you again for pointing out this critical detail.

**Q6: The manuscript mentions an activation function, but it does not specify which activation function was actually used. The final model settings should be included in the manuscript.**

**Respond:**

We thank the reviewer for pointing out this omission. We agree that specifying the exact activation function and hyperparameters is crucial for the reproducibility of our deep learning model. In the Supplementary Material, we have added Table R1 to explicitly detail the activation functions and provide a comprehensive summary of the final model configurations. Specifically, the following details have been added:

- **Activation Function:** We have explicitly stated that the **tanh** activation function was utilized in the hidden layers of our network because of its efficiency in mitigating vanishing gradients and accelerating convergence during training.
- **Final Model Settings:** To present the model configurations transparently, we have compiled all hyperparameters into a table, as shown in Table R1 (**Table S1** in the Supplementary Material), We believe these additions fully resolve the ambiguity and ensure that our model architecture is fully transparent and reproducible.

We have detailed these revisions in Lines 210 of the revised manuscript and the accompanying **Supplementary Material Table S1**.

**Table R1. Hyperparameters and training configurations of the AC-PIDNN model.**

| Category | Hyperparameter / Setting | Value / Description |
|---|---|---|

| | | |
|---|---|---|
| Network Configuration | Optimizer | Adam |
| | Weight Initialization | Xavier |
| | Activation Function | Tanh |
| Regularization (to prevent overfitting) | Dropout Rate | 0.3 |
| | L2 Regularization Factor (λ) | 0.01 |
| Stage 1: Initialization (Quickly fit geostrophic current) | Loss Function | $L_{data}$ |
| | Learning Rate | 0.001 |
| | Epochs | 1000 |
| Stage 2: Transition (Gradually introduce physical constraints) | Loss Function | $L_{hybrid}$ |
| | Hybrid Loss Proportion | 1:1 |
| | Learning Rate | 0.0001 |
| | Epochs | 1000 |
| Stage 3: Physics-Informed (Minimize momentum equation residuals) | Loss Function | $L_{physics}$ |
| | Learning Rate | 0.00001 |
| | Epochs | 10000 |

**Q7: In Figure 4(a-c), the nonlinear advective component is relatively weak, and under the current color scale its main spatial patterns are difficult to see by eye.**

**Respond:**

We gratefully acknowledge the reviewer for pointing out this visual clarity issue. To address this, we have regenerated **Figure 4 as shown in Figure R9**. Specifically, we have applied an independent, narrower color scale (adjusted colorbar limits) for the nonlinear advective component in panels (a-c). This adjustment significantly enhances the visual contrast, making the spatial patterns and localized variations of the advective term particularly its intensified signals around dynamically active regions such as western boundary currents and mesoscale eddies clearly visible to the reader. The updated Figure 4, along with a revised caption explaining the distinct color scales, has been incorporated into the revised manuscript.

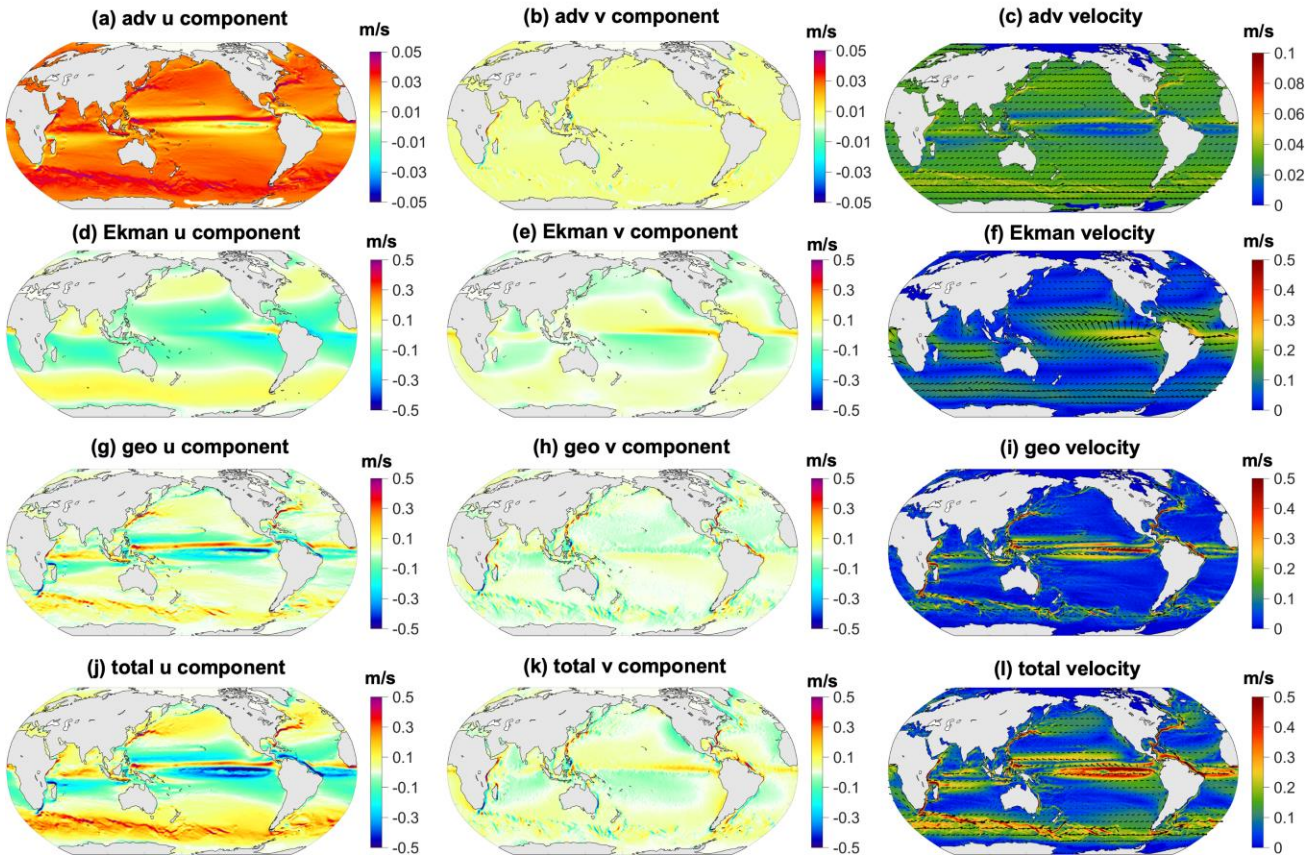



**Figure R9** Results of Sea Surface Circulation Correction from 1993 to 2023, using the AC-PIDNN model, incorporating geostrophic, nonlinear advective and Ekman corrected components. Panels (a)–(c) show the u-component, v-component, and flow field of the nonlinear advective component; Panels (d)–(f) show the u-component, v-component, and flow field of the Ekman corrected component; Panels (g)–(i) show the u-component, v-component, and flow field of the geostrophic component; Panels (J)–(L) show the u-component, v-component, and flow field of the total circulation.

**Q8: Line 157 contains "Error! Reference source not found." This should be fixed.**

**Respond Q8:**

We appreciate the reviewer pointing out this oversight. The broken cross-reference has been corrected to refer to **Line 155** in the revised manuscript. We have also verified all other citations and cross-references throughout the document to ensure their accuracy.

**Q9: The Figure 3 caption contains the phrase "the second column and column," which appears to be incorrect.**

**Respond Q9:**

We thank the reviewer for their careful reading and for catching this typographical error. We have corrected the phrase in the caption of **Figure 3 (Figure 5 in revised manuscript)**. The text has been revised from "the second column and column" to "the second column" to accurately describe the corresponding panels in the figure. We have also carefully proofread all other figure captions in the manuscript to ensure accuracy and readability.

**Q10: Around line 43 0, "Guanglinag Liu" appears to be a spelling error.**

**Respond Q10:**

We sincerely thank the reviewer for their extremely careful reading of our manuscript and for catching this typographical error. We have corrected the spelling from "Guanglinag Liu" to "**Guangliang Liu**" in the revised manuscript. Additionally, we have conducted a thorough proofreading of the entire text and the reference list to ensure that all names and terminologies are spelled correctly.

**Reviewer 2**

**General assessment**

**This manuscript presents a global daily sea-surface ageostrophic current dataset for 1993–2023 at 0.25° resolution using a physics-informed dense neural network, AC-PIDNN. The topic is timely and relevant for ESSD. A long-term global dataset that goes beyond purely geostrophic currents could be highly useful for studies of surface transport, eddy energetics, air-sea exchange, and model evaluation.**

**However, I recommend major revision before publication. The dataset is potentially significant, but the manuscript does not yet provide sufficient evidence that the reconstructed currents represent physically realistic ageostrophic circulation rather than a velocity field constrained to minimize the prescribed momentum residual. The most important issues are the limited independent validation, insufficient discussion of physical assumptions, and incomplete documentation of the dataset and machine-learning workflow.**

**Q1: Dataset validation is not sufficient for a global ESSD data product**

**The main validation is based on one TAO mooring location and one ROMS regional simulation. This is not sufficient for a 30-year global dataset. The authors should provide broader independent validation using multiple observational references, such as surface drifters, additional TAO/PIRATA/RAMA moorings, OSCAR or other current products, and regional comparisons in western boundary currents, the Southern Ocean, the tropics, and eastern boundary current systems.**

**The current validation demonstrates some skill in selected cases, but it does not establish global reliability.**

**Respond Q1:**

We sincerely thank the reviewer for this highly constructive comment. Following the suggestion, we have systematically expanded our validation against 54 in-situ observations from the TAO, PIRATA, and RAMA arrays (https://climatedataguide.ucar.edu/climate-data/tropical-moored-buoy-system-tao-triton-pirata-rama-toga). Besides, the widely used OSCAR dataset was also included in the comparison.

**1. Expanded In-situ Validation**

As shown in Figure R1(a), these 54 stations are widely distributed across the Pacific (21 stations), Atlantic (14 stations), and Indian Oceans (19 stations). Figure R1(b)–(d) display the Root Mean Square Error (RMSE) of the AC-PIDNN-corrected total currents, OSCAR currents, geostrophic currents (Geo), and Ekman corrected currents (Ekman) against the in-situ measurements. The statistical results are as follows:

- Pacific Ocean: The RMSE of the AC-PIDNN-corrected total current is 0.16 m/s, which outperforms both the Ekman corrected (0.24 m/s) and Geo (0.22 m/s) currents. Furthermore, compared to the standard OSCAR product's error of 0.19 m/s against in-situ data, the AC-PIDNN product reduces the error by approximately 15.8%.
- Indian Ocean: The RMSE of our corrected current is 0.10 m/s, outperforming Ekman corrected (0.13 m/s) and Geo (0.15 m/s) currents. Compared to the OSCAR product's error of 0.13 m/s, the AC-PIDNN product reduces the error by approximately 23.1%.
- Atlantic Ocean: The RMSE of our corrected current is 0.15 m/s, outperforming Ekman corrected (0.22 m/s) and Geo (0.17 m/s) currents. Compared to the OSCAR product's error of 0.17 m/s, the AC-PIDNN product reduces the error by approximately 11.8%.

Meridionally, larger errors are primarily concentrated near the equator (within a latitude band of approximately ±5°). Conversely, accuracy significantly improves in mid-to-low latitudes further away from the equator, aligning much more closely with field observations. Specifically, the RMSE gradually decreases from 0.25 m/s near the equator down to 0.05 m/s in these higher-latitude observation areas.

Overall, across all three major ocean basins, the corrected surface circulations are closer to the in-situ observations compared to the traditional geostrophic and Ekman corrected components, and they consistently demonstrate superior accuracy when compared directly to the OSCAR products.

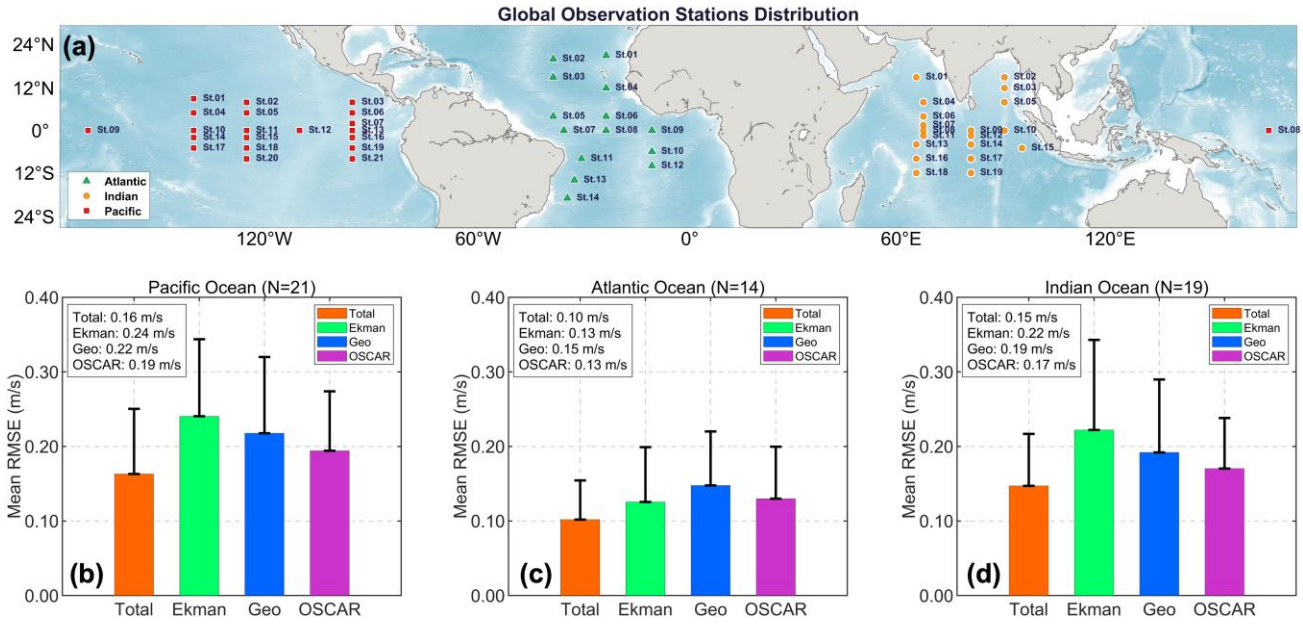



**Figure R1.** (a) Spatial distribution of TAO/PIRATA/RAMA mooring sites across the Pacific, Atlantic, and Indian Oceans. In-situ observation stations in each ocean basin are individually numbered to facilitate subsequent comparative experiments. (b)–(c) Bar charts illustrating the averaged root-mean-square error (RMSE) between TAO/PIRATA/RAMA mooring observations, OSCAR products and the AC-PIDNN-corrected current fields, geostrophic currents, and Ekman corrected currents for the Pacific, Atlantic, and Indian Oceans, respectively.

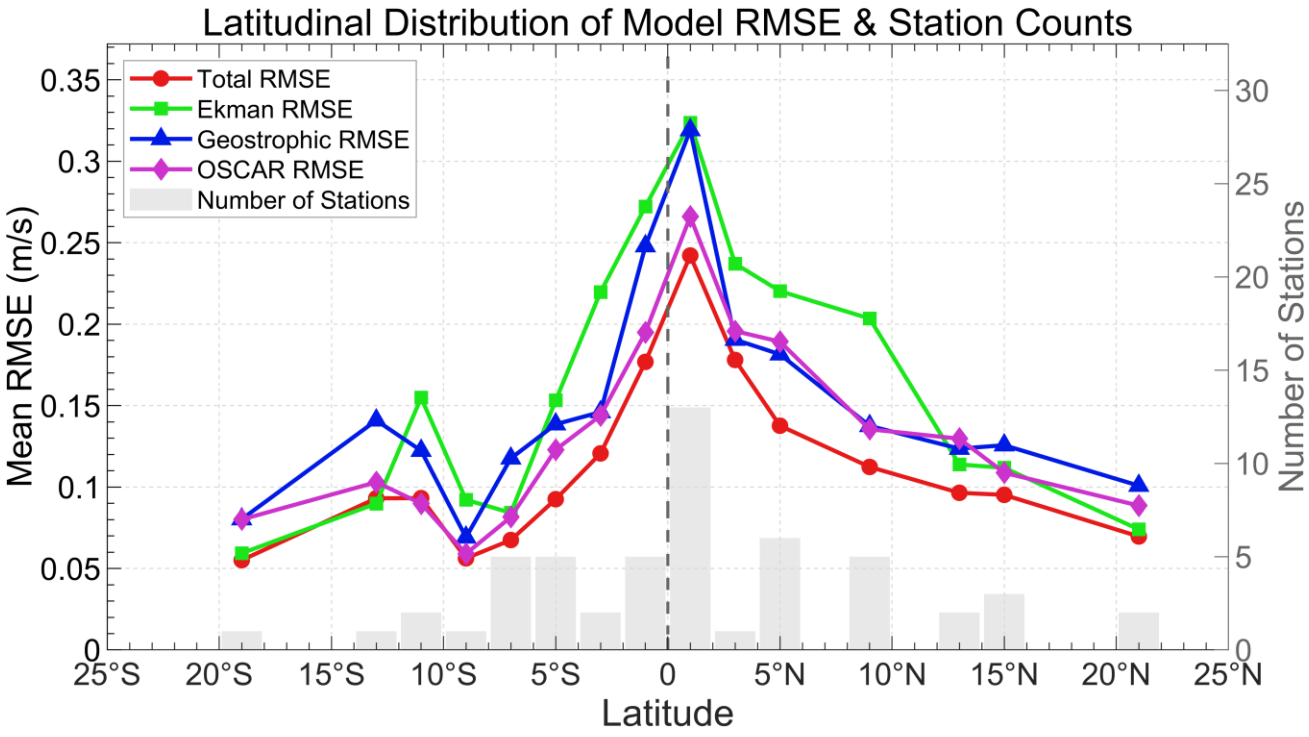


**Figure R2** presents the latitudinal error curves of the in-situ measurements against both the corresponding AC-PIDNN-corrected currents and the OSCAR products across the 54 observation stations.

**2. Global and Regional Spatial Comparisons (1993–2022 Means)**

Besides the comparison with the in-situ mooring observations, we further examined the circulation pattern with the OSCAR. The 30-year mean (1993–2022) total surface flow fields derived from OSCAR with those reconstructed by our method. Overall, the two datasets exhibit similar large-scale circulation patterns. However, because the PIDNN reconstruction additionally accounts for the nonlinear current component, its current magnitudes are generally larger than those in OSCAR and show better agreement with the observations (**Figures R1–R2**). The global mean total velocity for OSCAR is 0.16 m/s, compared to 0.19 m/s for AC-PIDNN.

The differences are particularly pronounced in high-kinetic-energy western boundary current regions, such as the Kuroshio and Gulf Stream, equatorial regions, and Southern Ocean. In these regions, AC-PIDNN more effectively reconstructs the strong currents and intense nonlinear advection characteristics. The incorporation of nonlinear effects also leads to stronger surface divergence in AC-PIDNN than in OSCAR in those regions **as shown in Figure R3**, although the two datasets retain broadly similar large-scale spatial patterns of divergence. The divergence field reconstructed by AC-PIDNN exhibits greater intensity and sharper spatial gradients, particularly in dynamically energetic regions of western boundary current (typical represent by the Kuroshio and Gulf Stream). These differences indicate that AC-PIDNN provides an improved representation of multiscale surface-current dynamics and associated convergence–divergence structures.

All the detailed comparative maps, statistical charts, and discussions mentioned above have been incorporated into **Section 4.1.2** of the revised manuscript and **Supplementary Figure S3**. We thank the reviewer again for this critical suggestion, which has vastly improved the robustness of our validation.

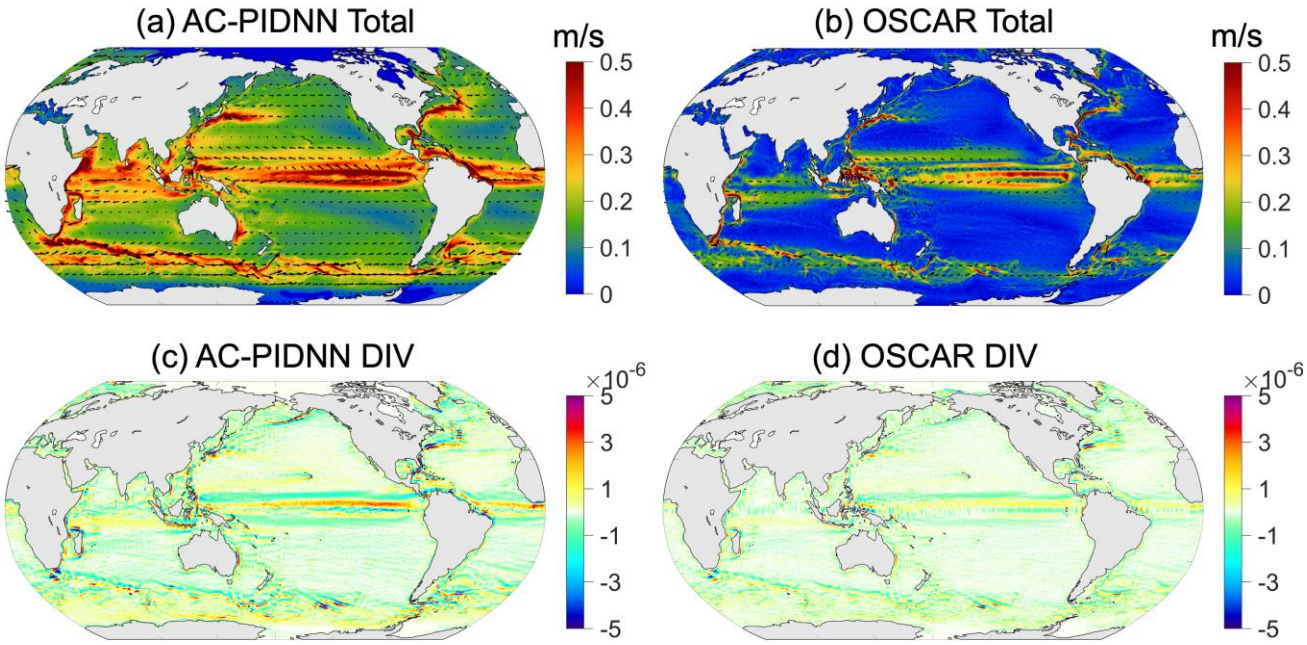



**Figure R3.** (a) 1993–2022 mean flow field from AC-PIDNN; (b) 1993–2022 mean flow field from the OSCAR product; (c) divergence of the 1993–2022 mean flow field from AC-PIDNN; (d) divergence of the 1993–2022 mean flow field from the OSCAR product

**Q2: Momentum residual is not an independent validation metric**

**The manuscript emphasizes that the reconstructed currents satisfy the imposed momentum balance. However, the momentum residual is part of the training objective. Therefore, a low residual is expected and should not be interpreted as independent evidence that the currents are accurate.**

**The authors should clearly distinguish between internal consistency with the imposed equations and external validation against independent observations or simulations.**

**Respond:**

We appreciate the reviewer's professional insight regarding the momentum residual. We agree that the momentum residual is not an "independent external validation" metric, but rather as a rigorous diagnostic criterion for the model's "internal physical consistency." The core rationale of our study is to utilize physical equations to correct the total flow field and isolate the non-linear advection terms. Within this physics-informed inversion framework, maintaining a sufficiently small momentum-equation residual is a prerequisite for the

reliable extraction of the nonlinear advection terms. Therefore, while a minimized residual is indeed the result of the training objective converging, it serves as proof that the network has successfully resolved the complex non-linear dynamics.

Following reviewer's concern, in the revised manuscript, we have rigorously distinguished between the concepts of "internal physical consistency" (demonstrated by the momentum residual) and "independent external validation" (demonstrated by comparisons with in-situ observations, as detailed in our previous responses). Furthermore, we have expanded the **Section 3.2.3 of the revised manuscript** to note that incorporating more comprehensive physical mechanisms remains a vital future direction to further bridge the gap between theoretical equation-based reconstructions and real-world measurements.

**Q3: The physical problem is underdetermined**

**The reconstruction attempts to infer total surface currents and ageostrophic components from SSH-derived geostrophic currents, Ekman currents, coordinates, and physical constraints. However, the steady momentum equations do not uniquely determine the ageostrophic current field. Multiple velocity fields may satisfy similar residual constraints.**

**The authors should discuss uniqueness, regularization, and sensitivity to initialization and loss weighting. This is essential because the dataset is being presented as a physical product, not only as a machine-learning output.**

**Respond:**

We sincerely appreciate this insightful and expert comment. We agree that inverting the total surface current field solely from steady-state momentum equations is fundamentally an underdetermined problem. To ensure that this dataset serves as a physical product rather than merely an output of a deep learning algorithm, we have implemented a multi-layered framework to constrain the solution space, ensure uniqueness, and maintain dynamical consistency.

**1. On Uniqueness and the Strong Data Prior**

As you correctly noted, underdetermined problems are sensitive to initialization. A random initialization (e.g., Xavier) combined with immediate physics-informed training would likely lead to divergence and cannot guarantee the uniqueness of solutions. This is the primary motivation for our "Three-Stage Training Strategy."

- Stage 1 (Pure Data-Driven Phase): This acts as an explicit "Physical Background Initialization." By forcing the neural network to first anchor to the ocean circulation, we provide a physically consistent starting point.
- Stage 2 **(**Hybrid Loss Optimization**):** Upon introducing a hybridized loss function that begins blending physical constraints with observational data, we observe a dramatic correction in the flow field's dynamical structure.
- Stage 3 **(**Fully Physics-Constrained Final Output**):** In the final stage, the output is rigorously constrained by the complete physics loss. This proper parameterization forces the reconstructed flow field to obey the fundamental equations of motion.

As demonstrated by the comparison with the observational data and the OSCAR product in the previous response, the solutions align more closely with the actual measurements. The momentum equations do not act as the sole source of information but constrain the model within the subspace defined by observational reality, effectively filtering out non-physical results.

To further demonstrate this, we compare our method with a model using Xavier initialization and direct physics-loss optimization. The comparison clearly shows that without Stage 1, as shown in **Figure R4**, the resulting velocity field is excessively smoothed and deviates significantly from physical reality.

Therefore, our progressive strategy explicitly overcomes the severe sensitivity to initialization inherent in underdetermined problems, ensuring the model converges to a unique oceanographic reality rather than being trapped in a mathematical artifact.

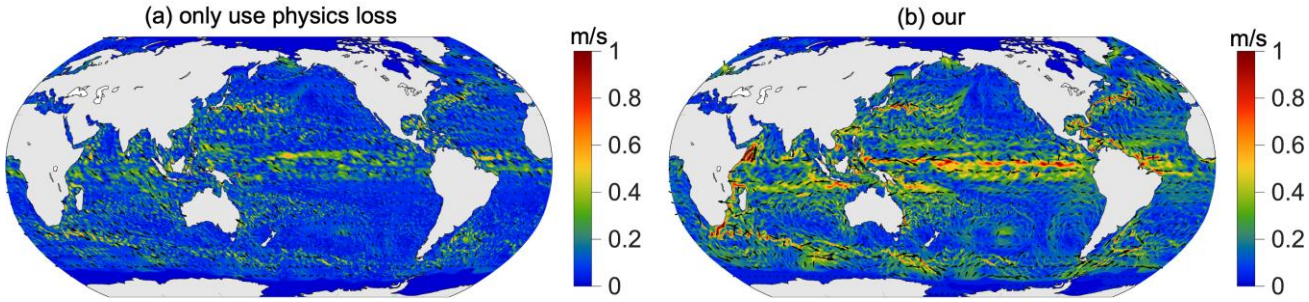



**Figure R4** (a) Results of training using only physical loss, (b) Final output of the method in this study.

**2. Sensitivity to Loss Weighting and Adaptive Loss Scaling**

We recognize that AC-PIDNN models are highly sensitive to the balance between physical and data losses. Static weighting often leads to "gradient pathologies" or overfitting to the steady-state equations. To address this and eliminate the need for manual hyperparameter tuning (e.g., grid search), we implemented an **Adaptive Loss Scaling** mechanism during the Stage 2 hybrid training:

$$\mathcal{L}_{hybrid} = \mathcal{L}_{data} + \lambda \times \mathcal{L}_{physics}$$

where the adaptive weight $\lambda$ is calculated as $\lambda = \text{sg}\left(\frac{L_{data}}{L_{physics}}\right)$. Here, $\text{sg}(\cdot)$ denotes the stop-gradient operation, ensuring that the weight acts only as a scaling factor and does not interfere with the backpropagation process. This mechanism maintains a 1:1 balance between data and momentum constraints throughout the training, ensuring robustness and reproducibility.

Ultimately, this adaptive scaling mechanism directly addresses your concern regarding loss weighting sensitivity, dynamically harmonizing physical constraints with observational data without relying on arbitrary, trial-and-error hyperparameter tuning.

**3. Physical Mapping of Regularization**

While we employ $L_2$ regularization and Dropout to mitigate overfitting, they serve distinct physical roles.

- $L_2$ Regularization: This is physically analogous to sub-grid scale parameterization and eddy viscosity (*Frederiksen,* 2006). In the context of inverse problems, it suppresses non-physical, high-frequency oscillations in regions of high shear, ensuring the spatial continuity of the divergence and velocity fields (*Smagorinsky*, 1963).

- Dropout: This functions purely as a statistical parameter constraint and lacks a direct fluid-dynamic counterpart, which we have clarified in the revised text.

Consequently, the regularization techniques applied here are not mere mathematical tricks; the $L_2$ penalty serves an explicit physical purpose by parameterizing sub-grid processes, thereby ensuring the fluid-dynamical stability of the reconstructed currents.

We have integrated these discussions on solution uniqueness, physical initialization, adaptive loss balancing, and the dynamical implications of regularization into the revised **Section 3.2.3 and Supplementary Table S1**. Thank you once again for your professional guidance in strengthening the physical validity of our work.

Reference

Frederiksen, J. S., & Kepert, S. M. (2006). Dynamical subgrid-scale parameterizations from direct numerical simulations. *Journal of the Atmospheric Sciences*, *63*(11), 3006–3019. https://doi.org/10.1175/jas3795.1

Smagorinsky, J. (1963). General circulation experiments with the primitive equations. *Monthly Weather Review*, *91*(3), 99–164. https://doi.org/10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2

**Q4: Loss function design and training strategy**

**The novelty of this work lies primarily in the loss function design and the three-stage training strategy rather than in the neural network architecture itself. While the approach is interesting, its implications deserve further discussion.**

**Stage 1 trains the network using only the data loss, effectively reproducing the geostrophic current field. Stage 2 combines the data and physics losses, allowing the solution to deviate from geostrophy while satisfying the momentum balance. However, Stage 3 removes the data loss entirely and optimizes only the physics loss. Consequently, the final solution is no longer explicitly constrained by the observational prior, but instead by the momentum equations.**

**This raises two important questions. First, does the final training stage improve agreement with independent observations, or does it primarily reduce the momentum residual? Second, does minimizing the physics loss alone lead to a unique and physically meaningful reconstruction?**

**A simple ablation experiment (as suggested by Reviewer #1) comparing the outputs after Stages 1, 2, and 3 against independent observations would help quantify the contribution of each training stage and clarify the balance between observational fidelity and dynamical consistency.**

**Respond Q4:**

We sincerely thank the reviewer for this insightful comment. To address two critical questions, we have detailed the rationale behind our approach and conducted the suggested ablation experiment.

**1. Rationale of the Three-Stage Training Strategy**

Since training deep learning models solely on physical equations has the risk of converging to non-unique, trivial solutions or overly smoothed fields, we mitigate this by gradually shifting the optimization landscape from a data-driven initialization to a strictly physics-constrained refinement. The loss functions are defined as follows:

$$\mathcal{L}_{data} = \frac{1}{N}\sum_{i=1}^{N}(\|u_{Pseudo\ labels} - u_i\|^2 + \|v_{Pseudo\ labels} - v_i\|^2)$$

$$\mathcal{L}_{physics} = \frac{1}{N}\sum_{i=1}^{N}\left(\|ADV_x + COR_x + PGF_x + WIND_x\|^2 + \left\|ADV_y + COR_y + PGF_y + WIND_y\right\|^2\right)$$

$$\mathcal{L}_{hybrid} = \mathcal{L}_{data} + sg\left(\frac{\mathcal{L}_{data}}{\mathcal{L}_{physics}}\right) \times \mathcal{L}_{physics}$$

Where the operator sg(·) denotes a stop-gradient operation ensuring that $sg\left(\frac{\mathcal{L}_{data}}{\mathcal{L}_{physics}}\right)$ acts strictly as a numerical scaling factor rather than a differentiable variable, $(u_{Pseudo\ labels}, v_{Pseudo\ labels})$ represents the pseudo-labels for model initialization, which are constructed using the sum of the corresponding geostrophic and Ekman components. The training is systematically executed in three stages:

- **Stage 1 (Data-Driven Initialization):** Using only $\mathcal{L}_{data}$ (Learning rate = 0.001, Epochs = 1000). This allows the model to rapidly capture the features of the geostrophic and Ekman corrected flows. It establishes a physically reasonable starting point, preventing the non-uniqueness problem that occurs when starting from random initialization, as shown in **Figure R4 (Q3)**.
- **Stage 2 (Transitional Hybrid Optimization):** Using $\mathcal{L}_{hybrid}$ (Learning rate = 0.0001, Epochs = 1000). Enforcing strong physical constraints too abruptly can cause gradient instability. This hybrid stage acts as a buffer, gently introducing momentum balance constraints to guide the model away from purely data-fitted local minima.
- **Stage 3 (Pure Physics Refinement):** Using only $\mathcal{L}_{physics}$ (Learning rate = 0.00001, Epochs = 10000). By removing the data constraint, we allow the network to minimize the momentum residuals strictly. Because the model is already anchored in a realistic state by Stages 1 and 2, minimizing $\mathcal{L}_{physics}$ does not lead to an arbitrary solution; instead, it accurately resolves the non-linear dynamical components that cannot be captured by data fitting alone.

Ultimately, this progressive training design addressed your concern: by systematically securing observational fidelity in the early stages, the optimization of the physics loss alone in Stage 3 acts as a final dynamical polishing. As proved by the comparison with only trained on the physics loss in our previous response, this guarantees a unique, physically meaningful reconstruction.

**2. Ablation Experiment: Observational Fidelity (Figure R6) vs. Dynamical Consistency (Figure R5)**

To answer your question regarding whether Stage 3 improves agreement with independent observations or merely reduces momentum residuals, we conducted a quantitative ablation experiment using 2021 flow field data. We saved the output of Stage 1, Stage 2 and Stage 3 to compare against the in-suit observation.

**(1) Dynamical Consistency**

Figures R9(a)–(c) illustrate the annual mean surface current fields generated by the AC-PIDNN model during Stages 1, 2, and 3, respectively. Correspondingly, Figures R9(d)–(f) map the momentum balance residual (Bal) for these three stages. The progression reveals a stark contrast in dynamical consistency.

- **Stage 1 (Data-Loss Only Baseline):** Because this initial stage relies exclusively on $Loss_{data}$ without any embedded physical constraints, the neural network acts as a standard data-fitting algorithm. Consequently, the output entirely lacks momentum balance correction, leading to severe dynamical inconsistencies where the momentum residuals exceed **100%** across the vast majority of the global ocean regions, as shown in Figure R9 (a) and (d).
- **Stage 2 (Hybrid Loss Optimization):** Upon introducing a hybridized loss function that begins blending physical constraints with observational data, we observe a dramatic correction in the flow field's dynamical structure. The global mean momentum residual drops precipitously to approximately **51%**, as shown in Figure R9 (b and (e).
- **Stage 3 (Fully Physics-Constrained Final Output):** In the final stage, the output is rigorously constrained by the complete physics loss. This proper parameterization forces the reconstructed flow field to obey the fundamental equations of motion, vastly improving the momentum balance and reducing the global average momentum residual to a mere **1.2%**, as shown in Figure R9 (c) and (f).

Therefore, this progression proves that Stage 3 does not merely reduce numerical errors, but fundamentally ensures the physical realism of the reconstructed currents by driving momentum residuals to negligible levels.

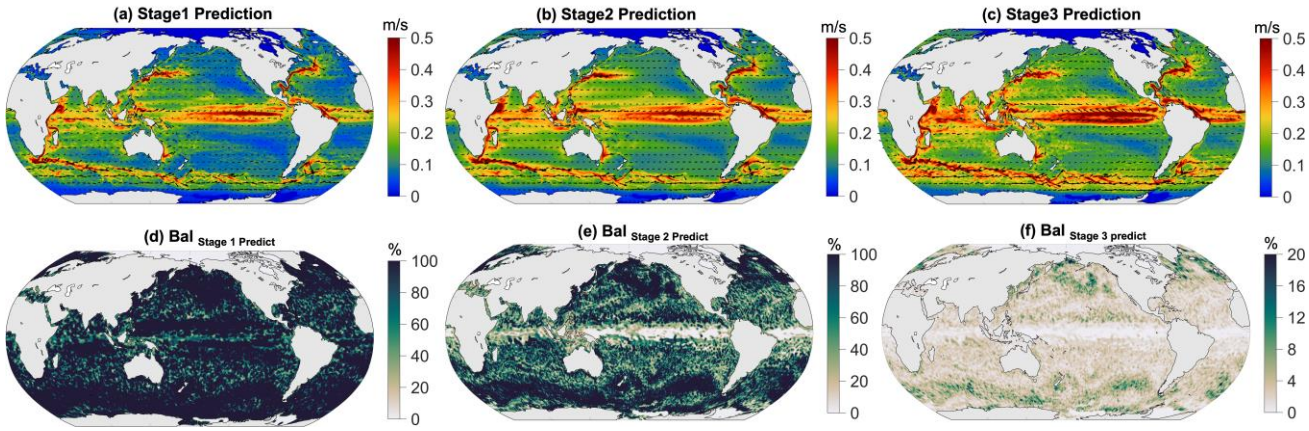



**Figure R5.** (a) Mean output current field of AC-PIDNN Stage 1 in 2021; (b) Mean output current field of AC-PIDNN Stage 2 in 2021; (c) Final mean output current field of AC-PIDNN Stage 3 in 2021; (d) Momentum

balance (Bal) results for Stage 1; (e) Momentum balance (Bal) results for Stage 2; (f) Momentum balance (Bal) results for Stage 3.

**(2) Observational Fidelity**

To confirm that this improved dynamical balance translates to real-world accuracy, we further evaluated the outputs of all three training stages, as well as the standard OSCAR product, against independent in-situ mooring data. Due to the incomplete coverage of the 2022 OSCAR dataset and computational time constraints during training, our comparative analysis focused on the velocity time series from successfully matched observation stations in 2021, as illustrated in Figure R8. As expected, the progressive incorporation of physical laws systematically refines the accuracy of the modeled velocities:

- **Stage 1 Average Error:** 0.17 m/s
- **Stage 2 Average Error:** 0.16 m/s
- **Stage 3 Average Error:** 0.13 m/s
- **OSCAR Product Average Error:** 0.17 m/s

The data clearly shows that the Stage 3 output aligns most closely with the in-situ measurements, significantly outperforming the Stage 1 baseline, the intermediate Stage 2 output, and the widely used OSCAR product. Purely data-driven approaches (Stage 1) are highly susceptible to the inherent noise and spatial gaps present in multi-source satellite altimetry and scatterometry data. By enforcing momentum conservation, the subsequent physics-informed stages effectively filter out dynamically impossible artifacts, preventing overfitting to observational noise and yielding a flow field that is both physically coherent and statistically more accurate.

This comparison with station-based observational errors firmly substantiates the superiority and absolute necessity of the physics-informed training stages. In conclusion, these results directly address your concern: the physical constraints applied in Stage 3 do not merely perform a mathematical reduction of momentum residuals, but genuinely enhance the model's predictive skill against independent, real-world observations.

We have incorporated the comprehensive results of this ablation experiment, alongside the associated figures and an expanded dynamical discussion, into the **Supplementary Material (Figure S1 to Figure S2)** of the revised manuscript. We once again thank you for your valuable and constructive comments.

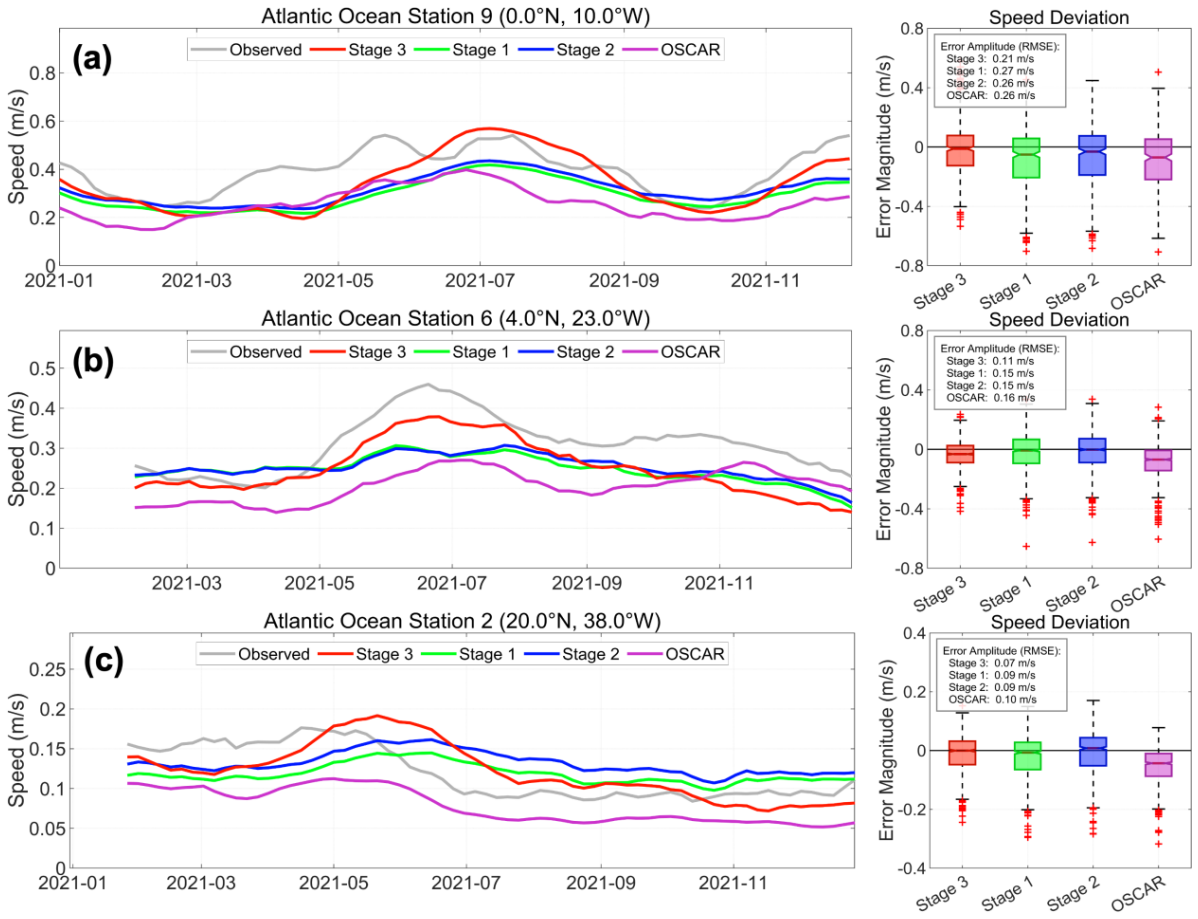



**Figure R6.** Comparison of the output current fields from the three training stages of AC-PIDNN and OSCAR product velocities against in-situ observational data.