Response to Referee #2

Blue font indicates the reviewer's comments, black font represents responses to these comments, and red font marks the corresponding modifications in the manuscript.

The paper applies an interpretable machine learning model based on the random forest algorithm to generate a 1° × 1° monthly sea surface wind dataset (MLAWind) spanning 1950-2023 and covering the near-global ocean between 60°S and 60°N. I think the authors present a novel and valuable approach for estimating sea surface winds directly from observed ICOADS data, rather than relying on winds derived from atmospheric data assimilation products such as ERA5.

Although the paper presents climatological 10-m winds for the period 1981-2010 (Fig. 7) and devotes substantial discussion to interannual variability and long-term trends (Sections 4.2 and 4.3), these analyses do not fully address the reliability of the dataset in earlier decades and data-sparse regions. The climatological period of 1981-2010 largely overlaps with the model training period (1993-2022), during which observational coverage is relatively dense. In addition, sea surface winds in tropical and low-latitude regions are generally easier to reconstruct because of the stronger relationship between SST and surface winds compared with mid- and high-latitude regions.

My primary concern is the quality and reliability of the dataset during data-sparse periods and in poorly observed regions, particularly prior to 1970 and over the Southern Ocean. To increase confidence among readers and potential users of the dataset, I recommend that the authors provide additional evidence demonstrating the robustness and accuracy of MLAWind under these challenging conditions.

Specifically, I suggest a comparative evaluation against the ERA5 reanalysis for three distinct periods: 1950-1969, 1970-1992, and 1993-2023. ERA5 is a widely used and well-validated reanalysis product and would provide a useful benchmark for assessing the temporal consistency of MLAWind. To facilitate a direct comparison, MLAWind could first be interpolated onto the ERA5 grid.

We sincerely appreciate your constructive comments. Following your suggestions, the quality and reliability of MLAWind during data-sparse periods and in poorly observed regions are evaluated in the revised manuscript. Specifically, a comprehensive comparison between MLAWind and ERA5 is conducted across three distinct periods: 1950–1969, 1970–1992, and 1993–2023. The assessment encompassed climatological means, standard deviations, correlation coefficients, and root-mean-square deviations (Figs. R1-6). Additionally, the time series of MLAWind and ERA5 are compared across different regions during 1950–2023 (Figs. R7-8). The results demonstrate that MLAWind achieves performance comparable to that of ERA5 under the challenging conditions, effectively capturing the spatiotemporal characteristics of sea surface winds during non-satellite periods and over the Southern Ocean.

Our point-by-point replies to the specific suggestions are as follows.

1. Global maps of the climatological mean of the 10-m zonal (10u) and meridional (10v) wind components, together with maps of their differences relative to ERA5.

Following your suggestion, the climatological distributions of MLAWind and ERA5 are compared across different periods (Fig. R1). The spatial distributions of MLAWind are highly consistent with those of ERA5, even during the pre-satellite periods (1950–1969 and 1970–1992). Both datasets capture prominent westerlies over the Kuroshio, Gulf Stream, and mid-latitude Southern Hemisphere. They reproduce the strong easterlies over the equatorial Pacific and Atlantic, together with the weak annual-mean winds over the equatorial Indian Ocean that arise from monsoon-driven seasonal reversals (Schott et al., 2009).

The absolute errors of MLAWind and multiple reanalysis datasets relative to ERA5 are compared. Across all three periods, MLAWind consistently exhibits significantly smaller errors than the other reanalysis datasets. Specifically, the differences between MLAWind and ERA5 are minimal during 1993–2023 (Fig. R1k), which is attributed to the inclusion of CCMP satellite data in the generation of both datasets (Hersbach et al., 2019). Owing to the lack of satellite observations, the discrepancies between MLAWind and ERA5 during the periods 1950–1969 and 1970–

1992 are slightly larger than those during 1993–2023 (Fig. R1c and g). Quantitative evaluation indicates that MLAWind demonstrates the strongest agreement with ERA5 among all evaluated datasets during the non-satellite periods, with global-mean errors remaining below 0.4 m/s (Table R1).

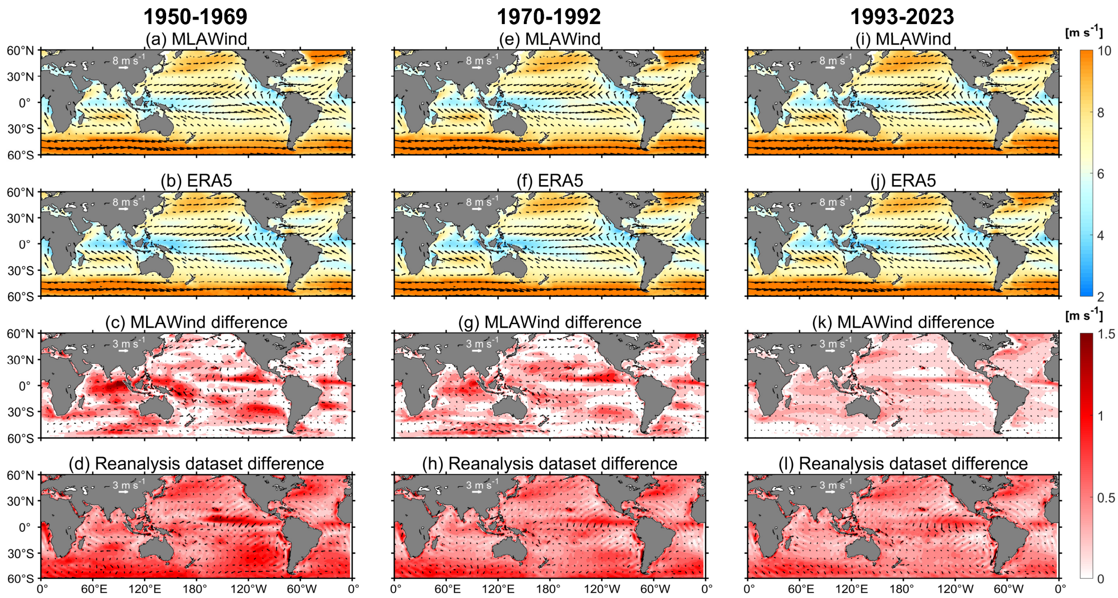

Figure R1. (a-b) Climatology of 10-m wind velocity (vector, m s^{-1}) and sea surface wind speed (shading, m s^{-1}) of MLAWind and ERA5 from 1950–1969. (c) The absolute error of MLAWind relative to ERA5. (d) The mean absolute error of multiple reanalysis datasets relative to ERA5. (e-h) and (i-l) Same as (a-d) but for 1970–1992 and 1993–2023, respectively. All datasets are interpolated onto a 1°×1° grid.

Table R1. Global-mean climatological sea surface wind speed errors (m s^{-1}) of multiple datasets relative to ERA5.

	MLAWind	JRA-55	NCEP1	NOAA-20C
1950–1969	**0.39**	0.44	0.72	0.83
1970–1992	**0.35**	0.37	0.74	0.65
1993–2023	**0.27**	0.30	0.73	0.59

A table is added to the revised manuscript to compare the climatological characteristics between the MLAWind datast and the reanalysis datasets across different periods. The corresponding description is provided in L411-L427:

"The climatological distributions between the MLAWind dataset and the reanalysis datasets are further compared across three distinct periods (i.e. 1950–1969, 1970–1992, and 1993–2022). Similar to Fig. 8, the spatial distributions of MLAWind are highly consistent with those of the reanalysis datasets, even during the pre-satellite periods (1950–1969 and 1970–1992). Both datasets capture prominent westerlies over the Kuroshio, Gulf Stream, and mid-latitude Southern Hemisphere. They reproduce the strong easterlies over the equatorial Pacific and Atlantic, together with the weak annual-mean winds over the equatorial Indian Ocean that arise from monsoon-driven seasonal reversals (Schott et al., 2009). The quantitative evaluation indicates that global-mean climatological sea surface wind speed of the MLAWind dataset exhibits low bias and RMSE relative to the reanalysis datasets, accompanied by a high correlation above 0.97 over the three examined periods (Table 3)."

Table 3. Comparison of global-mean climatological sea surface wind speed between the MLAWind dataset and the reanalysis datasets across different periods.

	1950–1969	1970–1992	1993–2022
Bias (m s^{-1})	0.37	0.37	0.28
RMSE (m s^{-1})	0.38	0.32	0.23
CORR	0.97	0.98	0.99

2. Global maps of the standard deviation of 10u and 10v anomalies (removed the seasonal cycle), together with maps of their differences relative to ERA5.

Following your suggestion, the standard deviations of MLAWind and ERA5 are compared across different periods (Figs. R2-3). The standard deviation of the 10-m zonal wind anomalies derived from MLAWind exhibits strong spatial consistency with that of ERA5 across all three evaluated periods, despite differences in magnitude. Both datasets reveal that high-variability regions are concentrated along the Kuroshio Current, the Gulf Stream, the tropical western Pacific Ocean, the tropical Indian Ocean, and the Southern Ocean, whereas low-variability regions are observed over the tropical eastern Pacific and Atlantic Ocean (Fig. R2).

MLAWind and reanalysis datasets exhibit comparable error distributions and magnitudes across different periods. During the non-satellite periods, they maintain small biases relative to ERA5 over most regions, with global mean errors of approximately 0.1 m/s. The slightly larger bias is both confined to the tropical western Indian Ocean during 1950–1969, while it is located in the tropical western Pacific Ocean during 1970–1992. These two regions are characterized by active air–sea interactions and strong wind field variability (e.g. Bjerknes, 1969; Kug et al., 2009; Wang, 2019). Previous studies have revealed considerable uncertainties among different datasets on decadal-to-interdecadal timescales (Guo et al., 2025a).

Similarly, the standard deviation of the 10-m meridional wind anomalies derived from MLAWind is highly correlated with that of ERA5 across all three evaluation periods, with error magnitudes comparable to those of mainstream reanalysis products (Fig. R3).

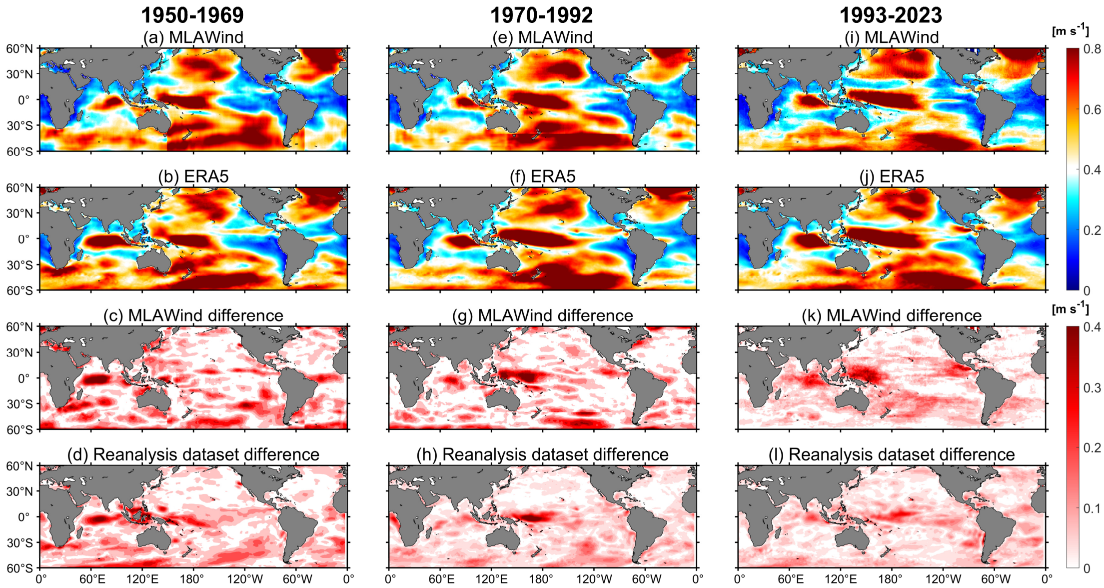

Figure R2. (a-b) Standard deviation of 10-m zonal wind anomalies (m s^{-1}) of MLAWind and ERA5 from 1950–1969. (c) The absolute error of MLAWind relative to ERA5. (d) The mean absolute error of multiple reanalysis datasets relative to ERA5. (e-h) and (i-l) Same as (a-d) but for 1970–1992 and 1993–2023, respectively. All datasets are interpolated onto a 1°×1° grid.

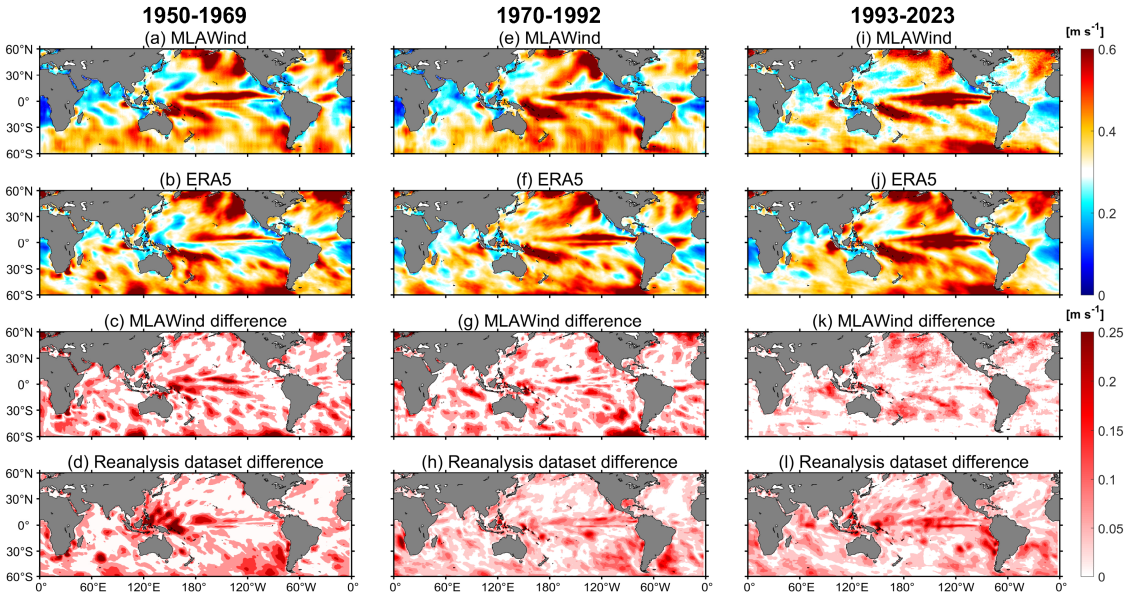

Figure R3. Same as Fig. R2, but for 10-m meridional wind anomalies (m s^{-1}).

A comparison of the standard deviations of the MLAWind dataset and the reanalysis datasets is added to the revised manuscript in L428–L436:

"The standard deviations of the 10-m zonal and meridional wind anomalies derived from the MLAWind dataset and the reanalysis datasets are assessed across different periods. The MLAWind dataset exhibits strong spatial consistency with the reanalysis datasets, despite differences in magnitude. Both datasets consistently indicate that high-variability regions are confined to the westerly wind belt and the Indo-Pacific Warm Pool region. These results demonstrate that the MLAWind dataset effectively captures the variability of sea surface winds, even during periods with sparse satellite data and in regions with limited observations."

3. Global maps of the correlation coefficient and root-mean-square deviation (RMSD) of 10u and 10v relative to ERA5.

Following your suggestions, the consistency between MLAWind and ERA5 is assessed using two statistical metrics: the correlation coefficient and the root-mean-square deviation (RMSD). During 1993–2023, the 10-m zonal and meridional winds from MLAWind exhibit the strongest positive correlations with those from ERA5, which can be attributed to the incorporation of CCMP satellite data in the generation of both datasets (Hersbach et al., 2019). During 1950–1992, MLAWind maintains significant positive correlations with ERA5 over most of the globe, indicating that MLAWind effectively captures the spatiotemporal characteristics of sea surface winds, even in the absence of satellite data (Fig. R4).

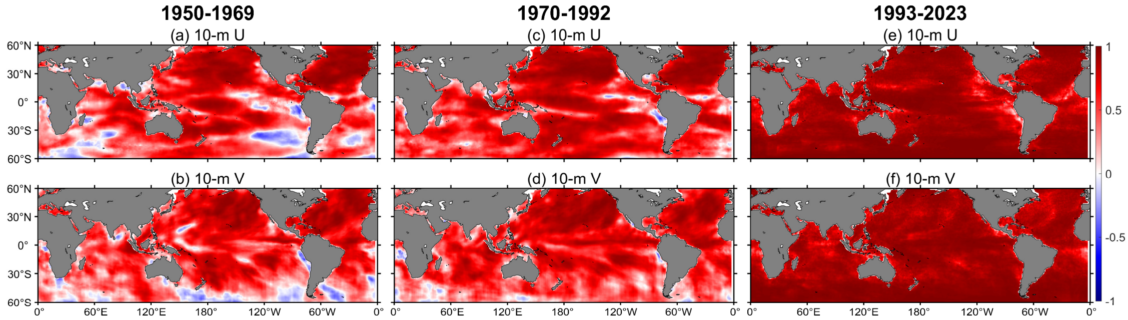

Figure R4. (a-b) Correlation coefficient between MLAWind and ERA5 for 10-m zonal and meridional winds during 1950–1969. (c-d) and (e-f) Same as (a-c) but for 1970–1992 and 1993–2023, respectively.

Figures R5 and R6 present the RMSD of MLAWind and NCEP1 relative to ERA5, respectively. The 10-m zonal wind of MLAWind exhibits minimal RMSD values during 1993–2023. During the non-satellite period, its RMSD remains modest over most of the global ocean, with relatively large values concentrated in the Southern Ocean and the eastern tropical Indian Ocean, particularly during 1950–1969 (Fig. R5a, c and e). It is noteworthy that the other reanalysis dataset exhibits comparable RMSD distributions (Fig. R6a, c and e). The pronounced discrepancies over the Southern Ocean can be attributed to the sparse observational coverage within the 30°S–60°S band, which introduces substantial uncertainties across all products. In addition, our previous study has revealed considerable uncertainties among different datasets in the tropical Indian Ocean, where MLAWind are found to be more reliable than the existing reanalysis datasets (Guo et al., 2025a).

The RMSD of the MLAWind 10-m meridional wind is comparable to that of the other reanalysis dataset, indicating strong spatiotemporal consistency between MLAWind and ERA5 (Figs. R5-6).

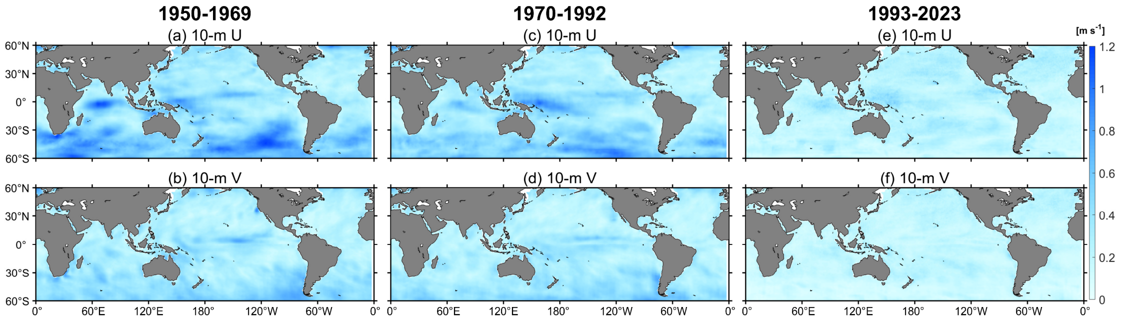

Figure R5. (a-b) Root-mean-square deviation (RMSD) between MLAWind and ERA5 for 10-m zonal and meridional winds during 1950–1969. (c-d) and (e-f) Same as (a-c) but for 1970–1992 and 1993–2023, respectively.

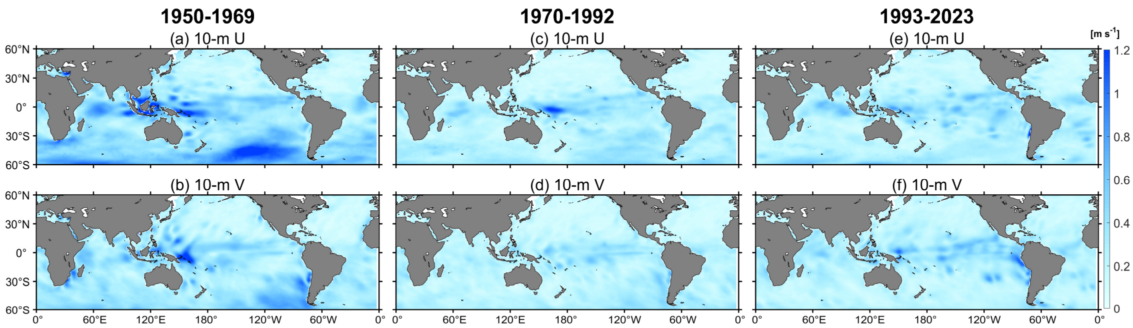

Figure R6. Same as Fig. R5, but for the RMSD between NCEP1 and ERA5 (m s^{-1}).

The analysis of the correlation coefficient and RMSD between MLAWind and ERA5 is added to the revised manuscript in L518–L539:

"In addition to the tropical Pacific Ocean, the inter-annual variability of the MLAWind dataset is systematically assessed on a global scale. Figure 12 presents the CORR and RMSE between MLAWind and ERA5. During 1993–2023, the 10-m zonal winds from MLAWind exhibit the strongest positive correlations with those from ERA5, which can be attributed to the incorporation of CCMP satellite data in the generation of both datasets (Hersbach et al., 2019). During 1950–1992, MLAWind maintains significant positive correlations with ERA5 over most of the globe. It indicates that MLAWind can effectively reproduce the inter-annual variability characteristics of reanalysis datasets, even in the absence of satellite data (Fig. 12).

The 10-m zonal wind of MLAWind exhibits minimal RMSE values during the satellite period. During 1950–1992, its RMSE remains modest over most of the global ocean, with relatively large values concentrated in the Southern Ocean and the eastern tropical Indian Ocean (Fig. 12). The pronounced discrepancies over the Southern Ocean can be attributed to the sparse observational coverage within the 30°S–60°S band, which introduces substantial uncertainties across all products. In addition, previous study has revealed considerable uncertainties among different datasets in the tropical Indian Ocean, where MLAWind are found to be more reliable than the existing reanalysis datasets (Guo et al., 2025a)."

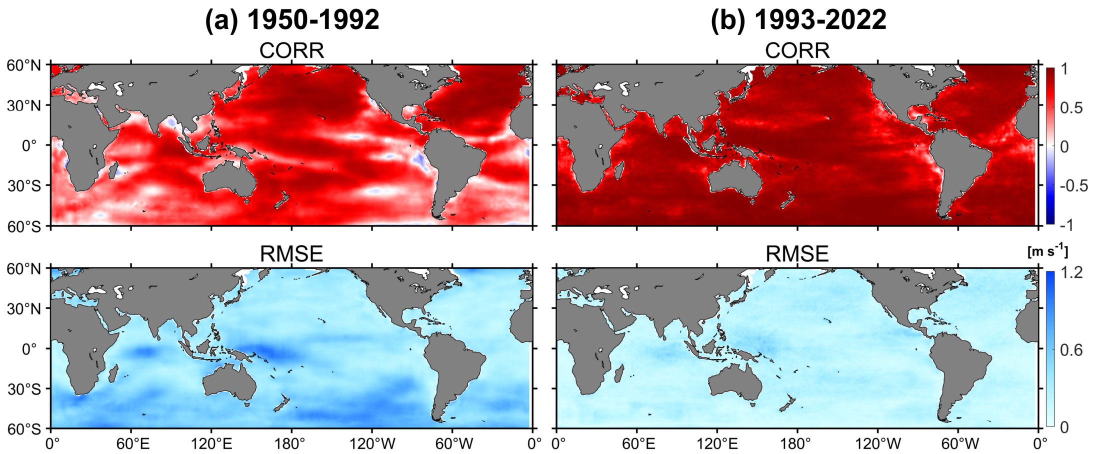

Figure 12. (a) Correlation coefficient (CORR) and Root Mean Square Error (RMSE) between MLAWind and ERA5 for 10-m zonal winds during 1950–1992. (b) Same as (a) but for 1993–2022.

4. In addition, for three broad regions (20°S-60°S, 20°S-20°N, and 20°N-60°N), I recommend presenting time series from January 1950 to December 2023 showing: The differences for regional mean 10u and 10v between MLAWind and ERA5.

Thanks for your valuable comments. Following your suggestions, the time series of 10-m zonal and meridional wind anomalies from multiple datasets across different latitude bands have been added to the revised manuscript (Figs. R7-8). Three slightly adjusted broad regions are used: 30°S–60°S, 30°S–30°N, and 30°N–60°N. The 30°S–60°S band is examined separately due to the relatively large inter-dataset uncertainty within this region (Figs. R5-6).

The 10-m zonal wind anomalies from MLAWind exhibit strong temporal consistency with those from ERA5 across different regions, even during periods without satellite observations (Fig. R7). In the 30°S–30°N and 30°N–60°N regions, the inter-annual variations of MLAWind are comparable to those of the reanalysis datasets, with small deviations from ERA5. In the 30°S–60°S westerly wind belt, the reanalysis datasets exhibit substantial uncertainty during 1950–1959 (blue shading in Fig. R7c). MLAWind exhibits closer agreement with ERA5 than the reanalysis ensemble mean, thereby demonstrating enhanced reliability (Fig. R7c). Similarly, MLAWind is effective in capturing the inter-annual variability of the 10-m meridional wind of ERA5 across different latitude bands, albeit with small differences in magnitude (Fig. R8). The result demonstrates that MLAWind maintain robust applicability during non-satellite periods and over the Southern Ocean. It can serve as a reliable data source for global wind field change studies.

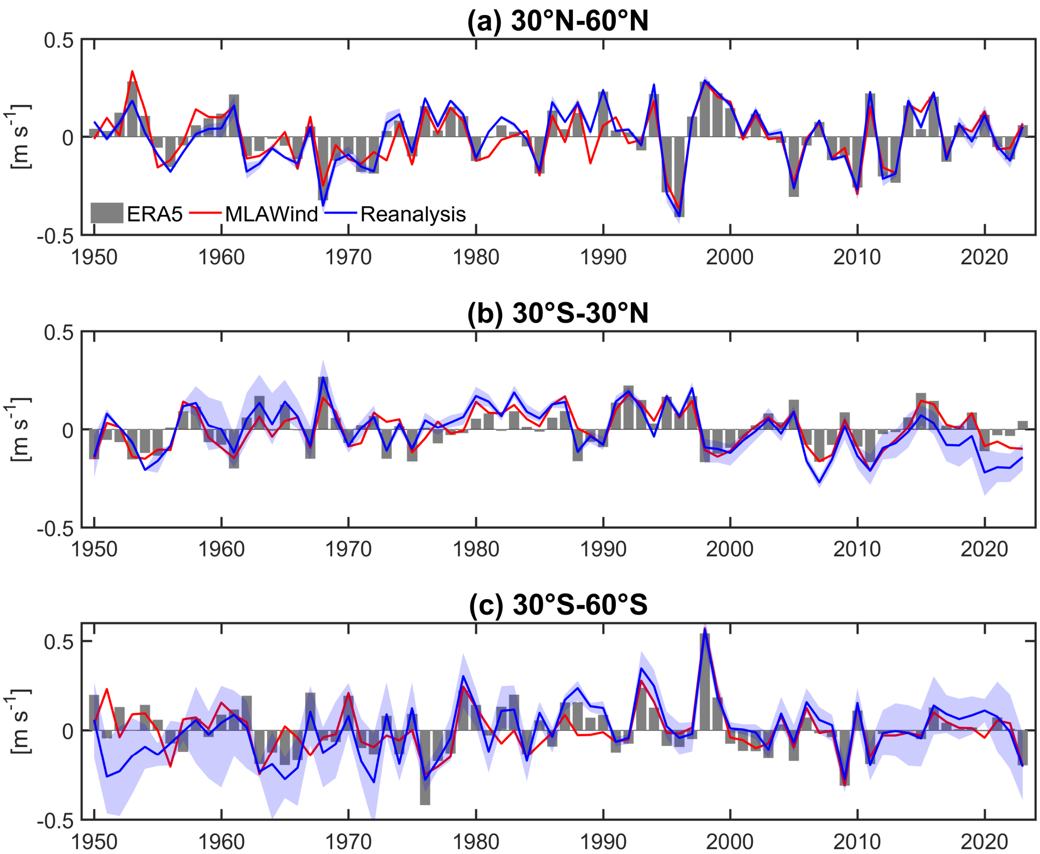

Figure R7. Inter-annual variations of the 10-m zonal wind anomalies (m s^{-1}) across different latitude bands. Long-term linear trends from 1950–2023 are removed. The blue line indicates the ensemble-mean time series of multiple reanalysis datasets (i.e.

NCEP1, JRA-55 and NOAA-20C). The blue shading indicates the spread in 10-m zonal wind anomalies among these reanalysis datasets.

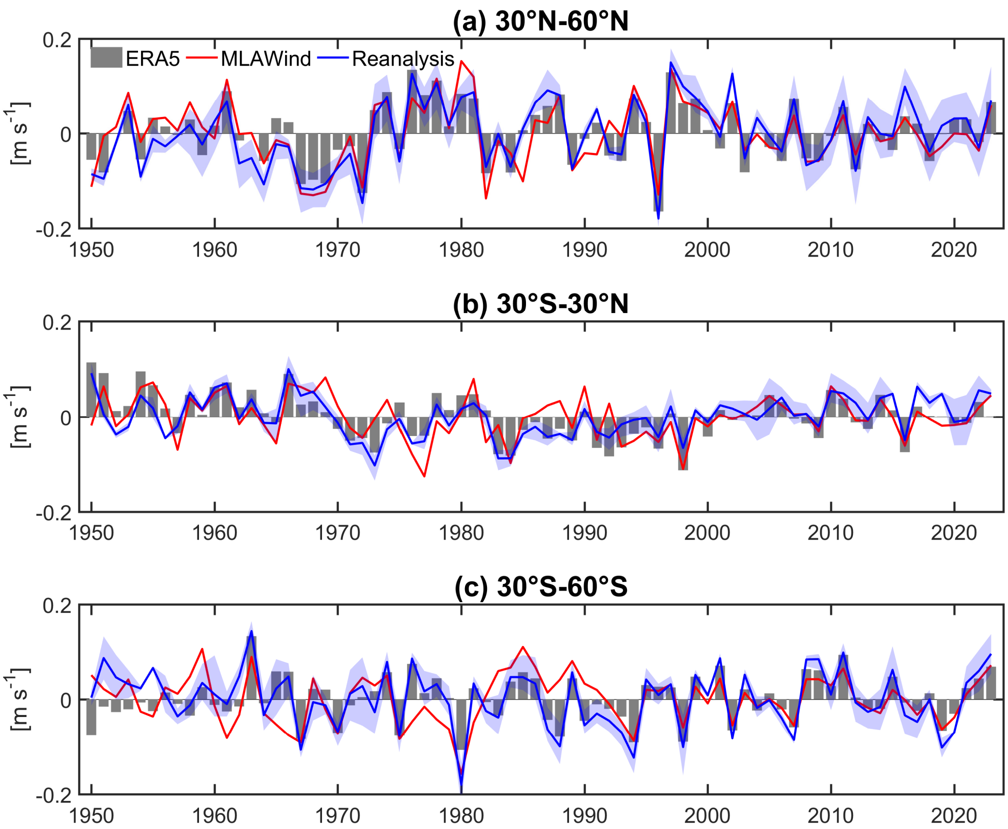

Figure R8. Same as Fig. R7, but for 10-m meridional wind anomalies (m s^{-1}).

The analysis of the time series of the MLAWind dataset and the reanalysis datasets across different latitude bands is added to the revised manuscript in L540–L554:

"To further compare the performance of the MLAWind dataset and the reanalysis datasets in different regions, their inter-annual variations are assessed across three broad latitude bands: 30°S–60°S, 30°S–30°N, and 30°N–60°N (Fig. 13). The 10-m zonal wind anomalies from MLAWind exhibit strong temporal consistency with those from reanalysis datasets across different regions, even during periods without satellite observations. Despite the relatively large RMSE between MLAWind and ERA5 over the westerly wind belt in the Southern Hemisphere (Fig. 12), MLAWind is effective in capturing the inter-annual variability of the 10-m zonal wind anomalies within the 30°S–60°S region (Fig. 13c). The result demonstrates that MLAWind maintain robust applicability during non-satellite periods and over the Southern Ocean. It can serve as a reliable data source for global wind field change studies."

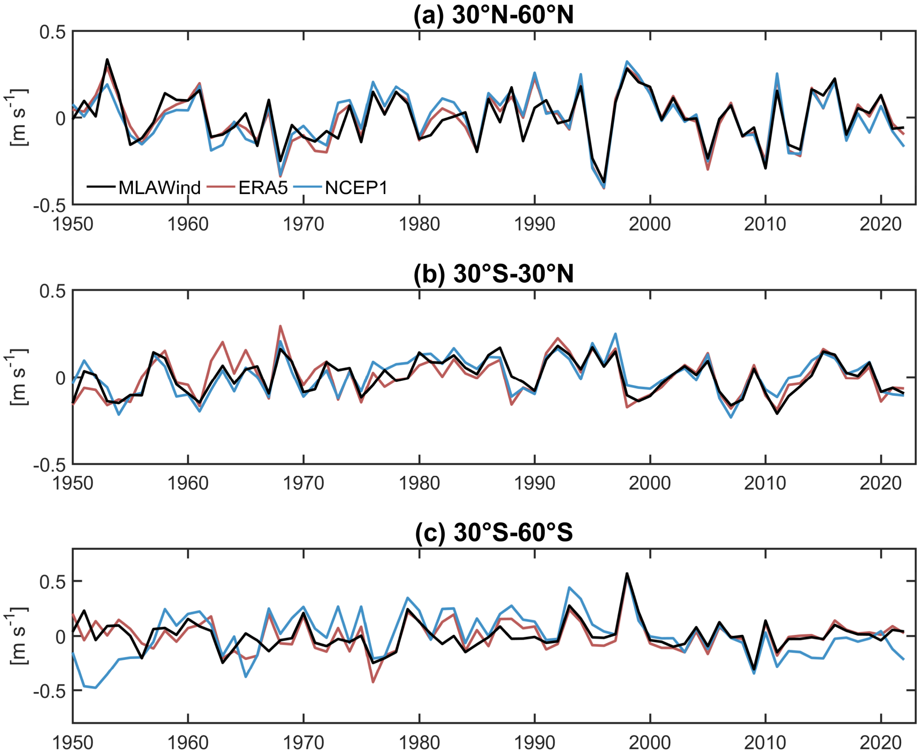

Figure 13. Inter-annual variations of the 10-m zonal wind anomalies (m s^{-1}) across different latitude bands. Long-term linear trends from 1950–2022 are removed.

Minor comments:

Lines 114–115: Please provide appropriate references for SHAP.

The relevant references for SHAP are added to the revised manuscript in L114-L117, L762-L765 and L783-L785:

"SHapley Additive exPlanations (SHAP) is a reliable interpretable module based on cooperative game theory, which demonstrates excellent compatibility across diverse machine learning models (Lundberg and Lee, 2017; Kim et al., 2025)."

"Kim, S., Alizamir, M., Heddam, S., Chang, S. W., Chung, I. M., Kisi, O., and Kulls, C.: Development of the machine learning and deep learning models with SHAP strategy for predicting groundwater levels in South Korea, Sci. Rep., 15, 35523, https://doi.org/ 10.1038/s41598-025-19545-y, 2025."

"Lundberg, S. and Lee, S.-I.: A unified approach to interpreting model predictions, Adv. Neural Inf. Process. Syst., 30, 1–10, https://doi.org/10.48550/arXiv.1705.07874, 2017."

Figure 2: The details in the right section of the "Pre-trained model" box lack clarity.

We sincerely apologize for the lack of clarity in the original Figure 2. In the revised manuscript, the resolution of Figure 2 has been increased to make the details located in the right section of the "Pre-trained model" box much clearer.

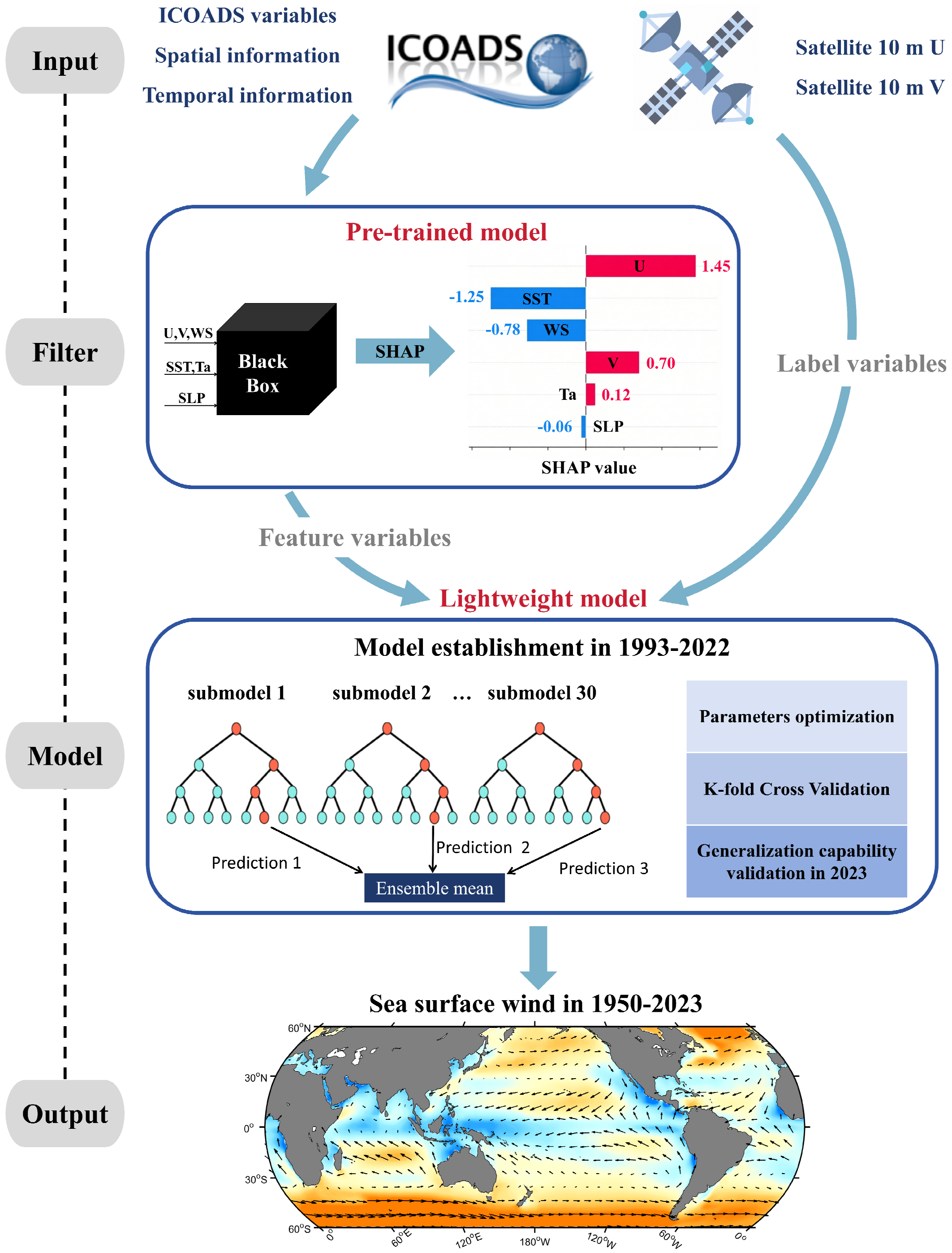

Figure 2. Schematic diagram of the construction process of machine learning-assisted sea surface wind dataset (MLAWind).

Line 221: Were the Ta and SLA variables in ICOADS excluded from the training data?

Yes. In this study, SHAP is used to quantify the contribution of model inputs and works as a filter to eliminate irrelevant features whose relative contribution falls below the 10% threshold. Based on SHAP analysis, the dominant features for reconstructing the 10-m zonal wind are identified as ICOADS 10-m zonal wind, SST, WS, and latitude. In the reconstruction of the 10-m meridional wind, the influential inputs include the ICOADS 10-m meridional wind, latitude, longitude, and month (Fig. R9). These variables are used as the training data, while those with little contribution (i.e. Ta, SLP and year) are excluded.

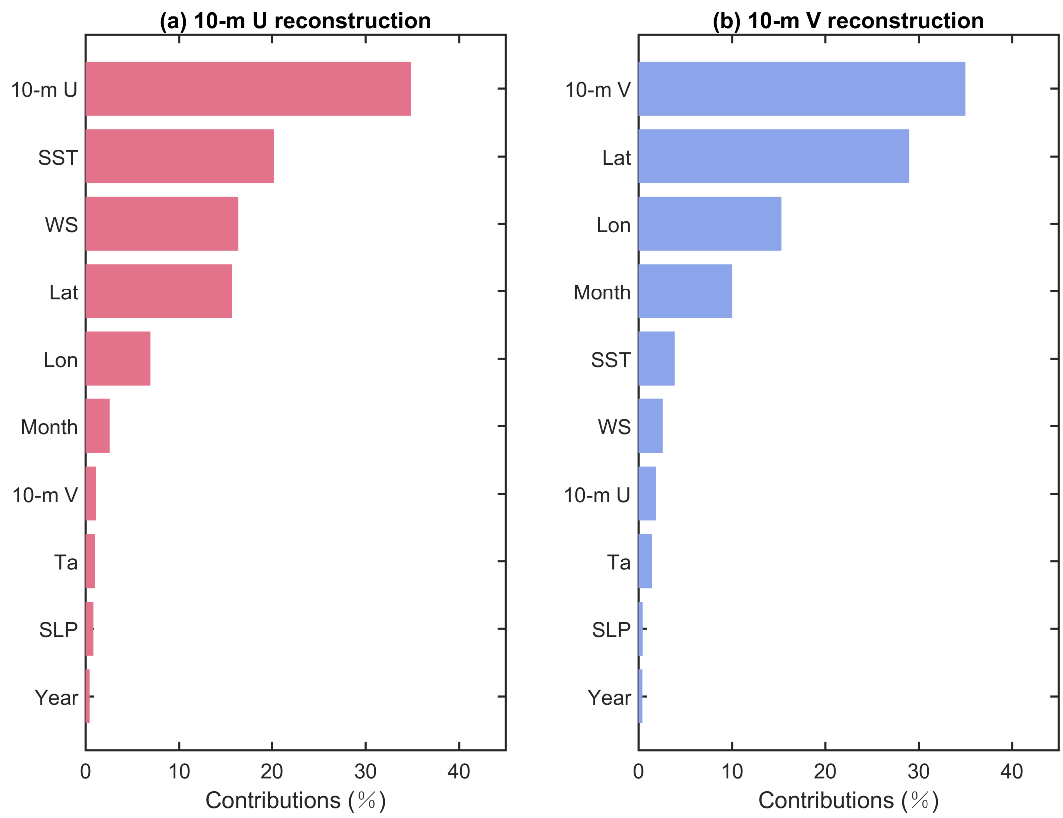

Figure. R9. The relative contributions (%) of different variables in the reconstruction of (a) the 10-m zonal and (b) the 10-m meridional wind,

In the revised manuscript, we explicitly state that the Ta and SLP variables in ICOADS are excluded from the training data in L254-L259:

"Based on the SHAP analysis, the ICOADS 10-m zonal wind, meridional wind, SST, WS, longitude, latitude, and month are identified as the seven key feature variables for sea surface wind reconstruction. They are utilized to establish a lightweight model, while other variables with little contribution (i.e. Ta, SLP and year) are excluded."

Lines 273–274: The formulas (4) and (5) for skewness and are incorrect; the index should start at i=1 rather than i-1.

Thank you for pointing out the typo. We have corrected the formulas as follows:

$$Skewness = \frac{\sqrt{n(n-1)}}{n-2}\frac{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^3}{(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2})^3} \tag{4}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{5}$$

References

Kim, S., Alizamir, M., Heddam, S., Chang, S. W., Chung, I. M., Kisi, O., and Kulls, C.: Development of the machine learning and deep learning models with SHAP strategy for predicting groundwater levels in South Korea, Sci. Rep., 15, 35523, https://doi.org/ 10.1038/s41598-025-19545-y, 2025.

Schott, F. A., Xie, S.-P., and McCreary, J. P.: Indian ocean circulation and climate variability, Rev. Geophys., 47, https://doi.org/10.1029/2007RG000245, 2009.