Supplementary information of

Unveiling the drivers of tropical Indian Ocean warming through machine learning-assisted surface wind

Weihao Guo[1,2]†, Rongwang Zhang[1,2]†, Xin Wang[1,2,3]*, Chunzai Wang[1,2,3]*, Xiaofeng Li[4,5], Weiqing Han[6], Lei Zhang[1,2,3]

Affiliations

[1] State Key Laboratory of Tropical Oceanography, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China.

[2] Global Ocean and Climate Research Center, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China.

[3] Guangdong Key Laboratory of Ocean Remote Sensing, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China.

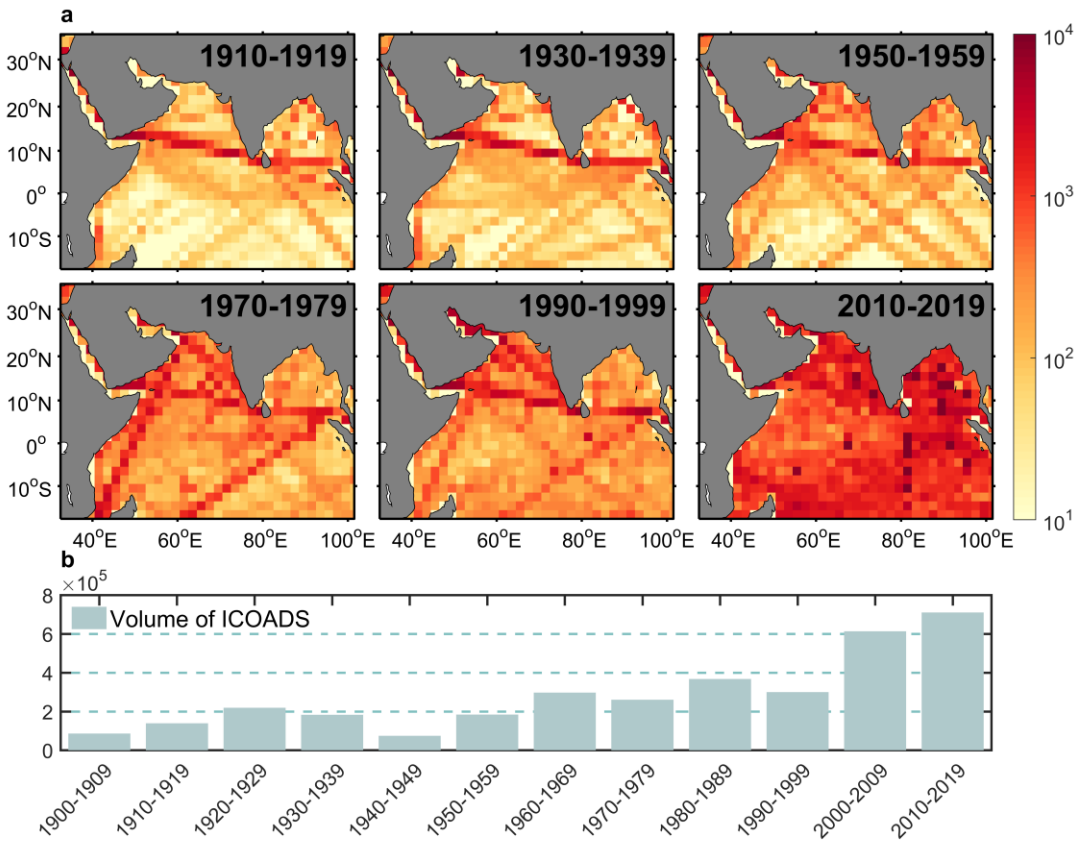
[4] Key Laboratory of Ocean Circulation and Waves, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China.

[5] Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao, China.

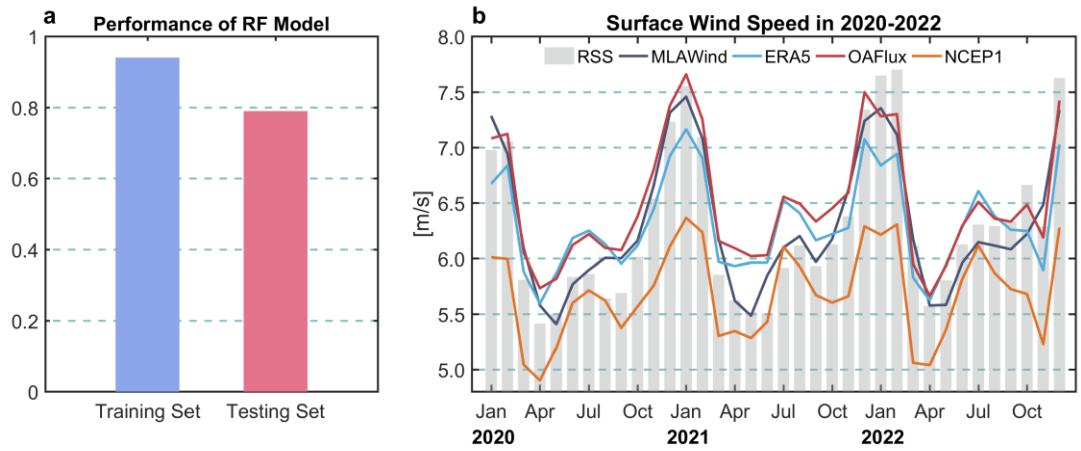
[6] Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado, USA.

* Corresponding authors: Xin Wang (wangxin@scsio.ac.cn); Chunzai Wang (cwang@scsio.ac.cn)

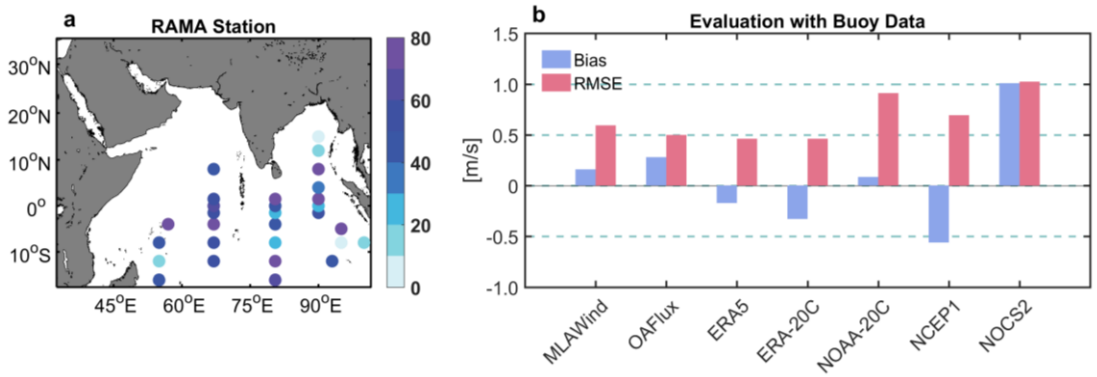
† These authors contributed equally to this work: Weihao Guo, Rongwang Zhang.



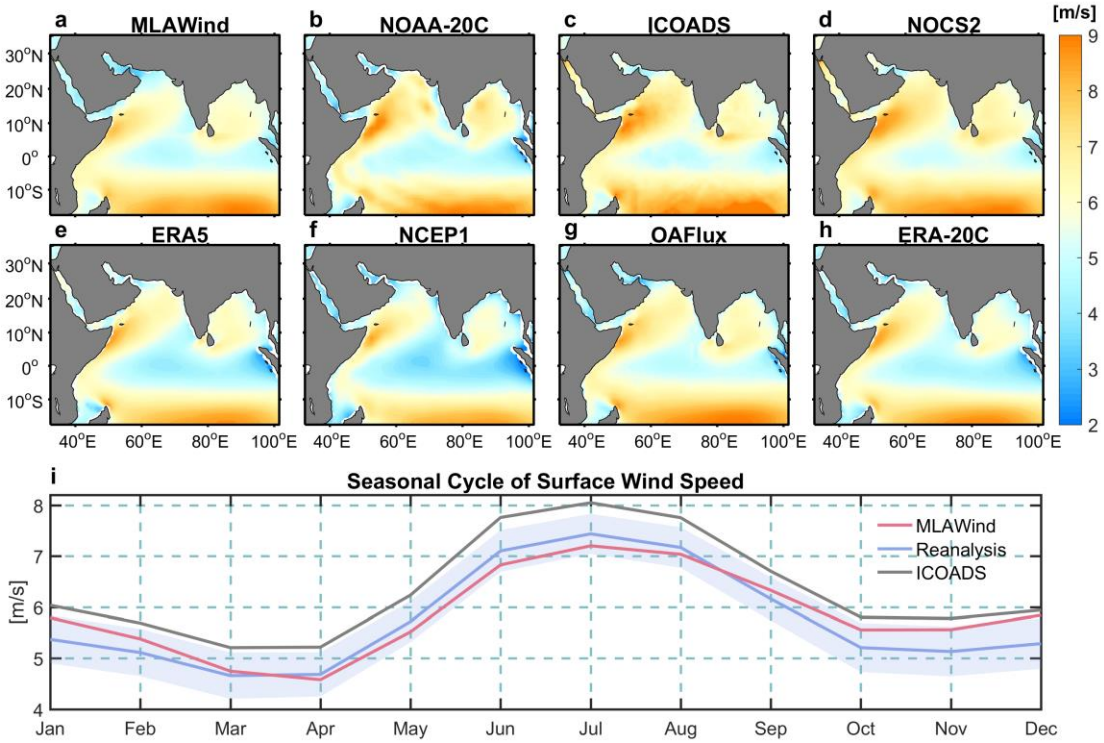
Supplementary Figure 1 | Variations in the volumes of ICOADS data over the Indian Ocean. a Distribution of ICOADS data in different periods. The shadings indicate the volumes of available data. **b** Regional mean volumes of ICOADS data over the Indian Ocean (40°E-100°E, 17.5°S-35°N).



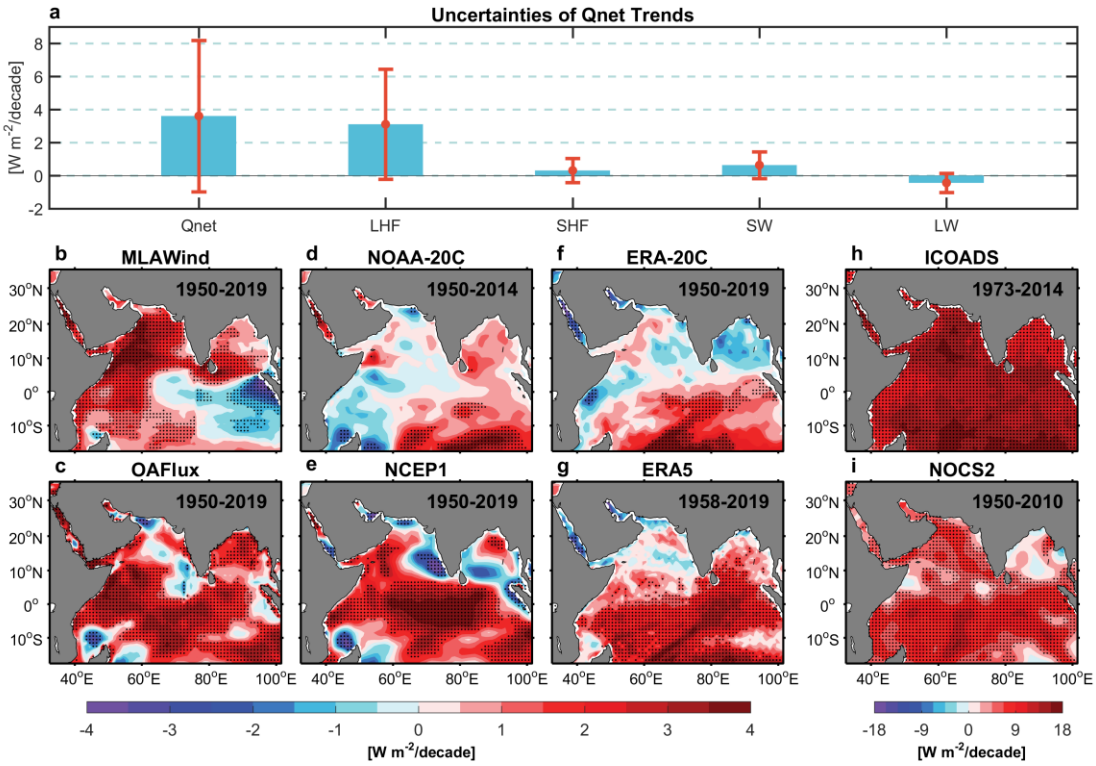
Supplementary Figure 2 | Evaluation of the random forest model. a Performance of the random forest model in the training set and testing set. R-squared is used as the metric of evaluation. **b** Time series of surface wind speed in the Indian Ocean (40°E-100°E, 17.5°S-35°N) for multiple datasets from 2020–2022. The grey bar indicates the labeled surface wind speed (RSS) in the random forest model. ERA5, OAFlux and NCEP1 are integrated in the current evaluation because of their available surface wind speed data for 2020–2022.



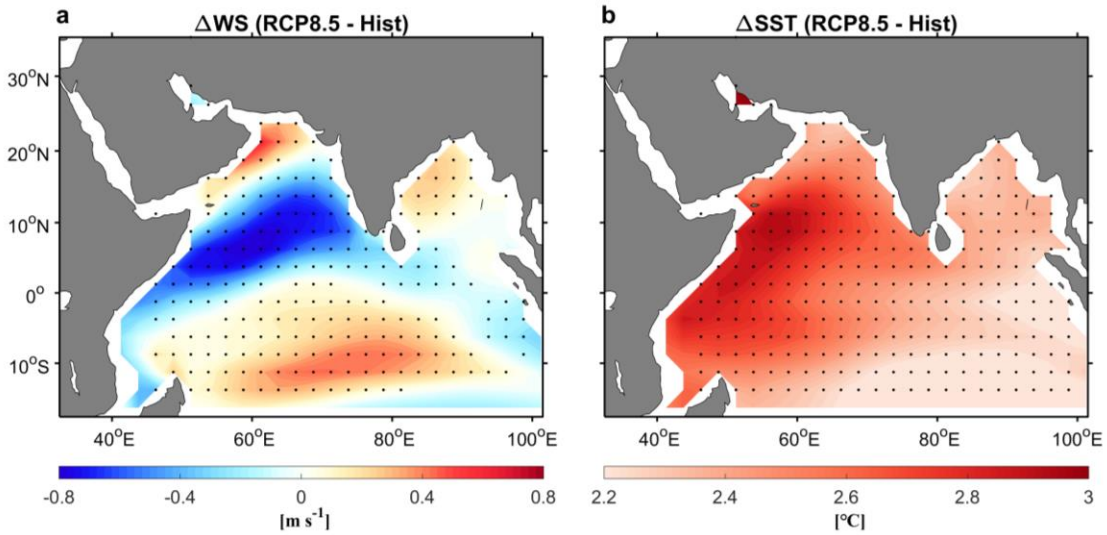
Supplementary Figure 3 | Evaluation of MLAWind via buoy data. a Distribution of RAMA in the Indian Ocean. Monthly mean data are used in this study, and the shadings at each station indicate the volumes of available data from 1990–2019. **b** Bias and RMSE values of multiple datasets evaluated with RAMA from 1990–2019.



Supplementary Figure 4 | Spatial distribution and seasonal cycle characteristics of surface wind speed in multiple datasets. a-h Climatology of surface wind speed (m/s) of MLAWind, NOAA-20C, ICOADS, NOCS2, ERA5, NCEP1, OAFlux and ERA-20C from 1981–2010. **i** Regional mean surface wind speed (m/s) over the Indian Ocean (40°E-100°E, 17.5°S-35°N) in different months. The blue shading indicates one standard deviation of multiple reanalysis datasets.



Supplementary Figure 5 | Annual mean heat flux trends in multiple datasets. a Ensemble mean surface net heat flux (Qnet), latent heat flux (LHF), sensible heat flux (SHF), shortwave radiation (SW) and longwave radiation (LW) trends (W/m^2/decade) obtained from multiple datasets over the tropical Indian Ocean (40°E-100°E, 10°S-10°N). The red error bars denote one standard deviation of multiple datasets. **b-g** Annual mean LHF trends (W/m^2/decade) over the Indian Ocean of MLAWind, OAFlux, NOAA-20C, NCEP1, ERA-20C and ERA5. The bold texts indicate the periods used to calculate the trends. The black dots indicate the linear trends exceeding the 99% significance level based on the *F*-test. **h-i** Same as **b-g** but for ICOADS and NOCS2.



Supplementary Figure 6 | Future variations in surface wind speed (WS) and sea surface temperature (SST) over the Indian Ocean. a Composites of WS difference (m/s) between RCP8.5 scenarios (2045–2100) and historical simulations (1950–2005) in 35 CESM-LENS models. The black dots indicate the composite exceeding the 99% significance level based on Student's *t*-test. **b** Same as **a** but for SST difference (°C).