

1 **Bowen ratio-constrained global dataset of bulk air-sea turbulent**
2 **heat fluxes from 1993 to 2017**

3 Yizhe Wang^{a, b}, Ronglin Tang^{a, b, *}, Meng Liu^c, Lingxiao Huang^{a, b}, Zhao-Liang Li^{a, b, c}

4 ^a State Key Laboratory of Resources and Environment Information System, Institute of
5 Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences,
6 Beijing 100101, China

7 ^b University of Chinese Academy of Sciences, Beijing 100049, China

8 ^c State Key Laboratory of Efficient Utilization of Arable Land in China, Institute of
9 Agricultural Resources and Regional Planning, Chinese Academy of Agricultural
10 Sciences, Beijing 100081, China

11 * Authors to whom correspondence should be addressed: tangrl@lreis.ac.cn

12

13

14

15 **Abstract**

16 Air-sea turbulent heat fluxes, including the sensible heat flux (SHF) and latent heat
17 flux (LHF), along with the Bowen ratio (β , ratio of SHF to LHF), are crucial for
18 understanding air-sea interaction and global energy and water budgets. However, the
19 existing products, primarily developed using the semi-empirical bulk aerodynamic
20 methods and data-driven machine learning approaches, are often weak in accuracy and
21 physical rationality, due to the uncertainties in the environmental forcings and
22 inappropriate parameterizations. In this study, we generated a global daily 0.25° product
23 of bulk air-sea turbulent heat fluxes using the Bowen ratio-constrained Neural Network
24 (NN) model (referred to as the BrTHF model) that could coordinately estimate the SHF
25 and LHF, along with the observations from 197 globally distributed buoys and multi-
26 source remote sensing and reanalysis inputs. The spatial ten-fold cross-validation
27 results showed that the BrTHF model achieved root mean square errors of 6.05 W/m^2 ,
28 23.67 W/m^2 and 0.22 and correlation coefficients of 0.93 , 0.91 and 0.25 for the SHF,
29 LHF and β , respectively. Compared with the physics-agnostic NN model and seven
30 widely used air-sea turbulent heat flux products (including JOFURO3, IFREMER,
31 SeaFlux, ERA5, MERRA2, OAFflux, and OHF), the BrTHF model shows better
32 agreement with observations, primarily attributable to its improved representation of β .
33 Furthermore, the inter-comparison of the spatial distribution of multi-year means, as
34 well as intra-annual and inter-annual change patterns showed that the BrTHF product
35 reliably simulated global SHF, LHF and β , in contrast to the machine learning-based
36 OHF product that failed to replicate these patterns. The main advantage of the BrTHF
37 model lies in its improved rationality of β estimates, successfully eliminating the
38 outliers observed in the physics-agnostic NN model and the seven typical products. The
39 improved SHF, LHF, and β estimates can allow for more accurate quantification of the
40 global air-sea energy and water budgets, enhance our understanding of air-sea
41 interaction, and improve projections of climate change under global warming. The 0.25°
42 daily global product from 1993 to 2017 can be freely accessed from the National

43 Tibetan Plateau Data Center (TPDC) [<https://doi.org/10.11888/Atmos.tpdc.302578>,
44 Tang and Wang (2025)].

45 **Keywords:** Air-sea turbulent heat fluxes; Sensible heat flux; Latent heat flux; Bowen
46 ratio

47 **1. Introduction**

48 Air-sea turbulent heat fluxes, comprising the latent heat flux (LHF) and sensible
49 heat flux (SHF), play vital roles in the Earth's climate system by characterizing the
50 exchange of energy and water between the ocean and atmosphere (Wild et al., 2014;
51 Loeb et al., 2021; Fasullo et al., 2014). Accurate estimation of SHF, LHF and their
52 ratio—the Bowen ratio ($\beta = \text{SHF}/\text{LHF}$) is an essential prerequisite for advancing our
53 understanding of atmosphere-sea interaction (Gentemann et al., 2020), improving the
54 quantification of global water and energy budget (Zhang, 2023), and enhancing the
55 predictability of extreme weather events (Yu, 2019).

56 To estimate global air-sea turbulent heat fluxes, the semi-empirical bulk
57 aerodynamic method was developed based on the Monin-Obukhov similarity theory
58 (Monin and Obukhov, 1954). It establishes scaling relationships between fluxes and
59 near-surface meteorological variables such as wind speed, humidity, and temperature
60 (Yu, 2019). This method, for its ease of application, has been applied to generate tens
61 of widely used products in the past few decades (Shie et al., 2009; Liman et al., 2018;
62 Yu and Weller, 2007; Berry and Kent, 2011; Tomita et al., 2018; Crespo et al., 2019).
63 However, there were huge discrepancies in the global and regional magnitude and
64 patterns of SHF and LHF among these products, which seriously impeded our
65 understanding of the air-sea interaction and the global budget of water and energy
66 (Bentamy et al., 2017; Tang et al., 2024; Yu, 2019). The discrepancies could be partly
67 attributed to the substantial uncertainties in the environmental forcings used to develop
68 these products (Robertson et al., 2020) and the inappropriate parameterizations
69 (Brodeau et al., 2017; Jiang et al., 2024a; Jiang et al., 2024b; Yang et al., 2024). More
70 explicitly, existing parameterizations often rely on simplified assumptions about

71 atmospheric stability and boundary layer dynamics, which may not hold under diverse
72 environmental conditions. For instance, most bulk algorithms are optimized for
73 moderate wind regimes, resulting in degraded performance and increased uncertainty
74 when applied under weak wind regimes (Jiang et al., 2024a; Brunke, 2002). At very
75 high wind speeds, however, observations show that the drag coefficient can decrease
76 due to sea spray and whitecap formation, reducing effective surface roughness and
77 potentially biasing flux estimates (Cai et al., 2025). In addition, simplifications in the
78 treatment of sea surface skin temperature, saturation humidity, and air density in the
79 parameterizations can also introduce substantial uncertainty (Brodeau et al., 2017).
80 Together, these limitations can contribute a lot to the biases in the SHF and LHF
81 estimates which can even lead to the unphysical estimations of β , as Wang et al. (2024)
82 reported. To better describe and comprehend the air-sea interaction and the energy and
83 water budgets, the existing mode to produce global air-sea turbulent heat fluxes needs
84 improvement urgently.

85 Machine learning techniques have been extensively applied to upscale point-scale
86 in-situ measurements of a single variable (such as soil moisture, roughness, or
87 temperature) into grid-scale global datasets (Wang et al., 2023; Peng et al., 2022; O and
88 Orth, 2021; Nelson et al., 2024; Fu et al., 2023). These efforts highlight the great
89 potential of machine learning for more accurate and consistent multivariate coordinated
90 mapping (Karniadakis et al., 2021; Kashinath et al., 2021; Van Der Westhuizen et al.,
91 2023; Wang et al., 2024). However, the application of machine learning in global
92 mapping of air-sea turbulent heat fluxes remains limited. Among these studies, some
93 have focused on solely improving the accuracy of LHF, while the remaining studies
94 have mostly considered independent modeling of SHF and LHF (Bourras et al., 2007;
95 Cummins et al., 2024; Cummins et al., 2023; Zhou et al., 2024). In both cases, however,
96 most studies have not produced long-term flux products. The only publicly available
97 machine learning-based global air-sea turbulent heat fluxes product, released by the
98 National Oceanic and Atmospheric Administration (NOAA) Ocean heat flux CDR

99 (hereafter dubbed OHF), simultaneously modeled SHF and LHF using a Neural
100 Network (NN) technique (Clayson and Brown, 2016). Although it performed well when
101 validated against the observations from the tropical buoys, it failed to capture the
102 regional characteristics, particularly in areas where air-sea turbulent heat exchange is
103 intense (e.g. oceans with latitudes beyond 45° for SHF and subtropical highs for LHF)
104 (Tang et al., 2024). Additionally, it exhibited different patterns of temporal evolution of
105 global annual mean and opposite inter-annual trends at both regional and global scales
106 to most widely-used physical model-based products, likely due to unreasonable
107 construction of observation datasets [with data before and after 2007 coming from
108 SeaFlux in-situ datasets and ICOADS (International Comprehensive Ocean-
109 Atmosphere Data Set) datasets, respectively]. Furthermore, the product likely suffers
110 from unphysical estimates of the β due to neglecting the interrelations among SHF, LHF
111 and β during the model construction.

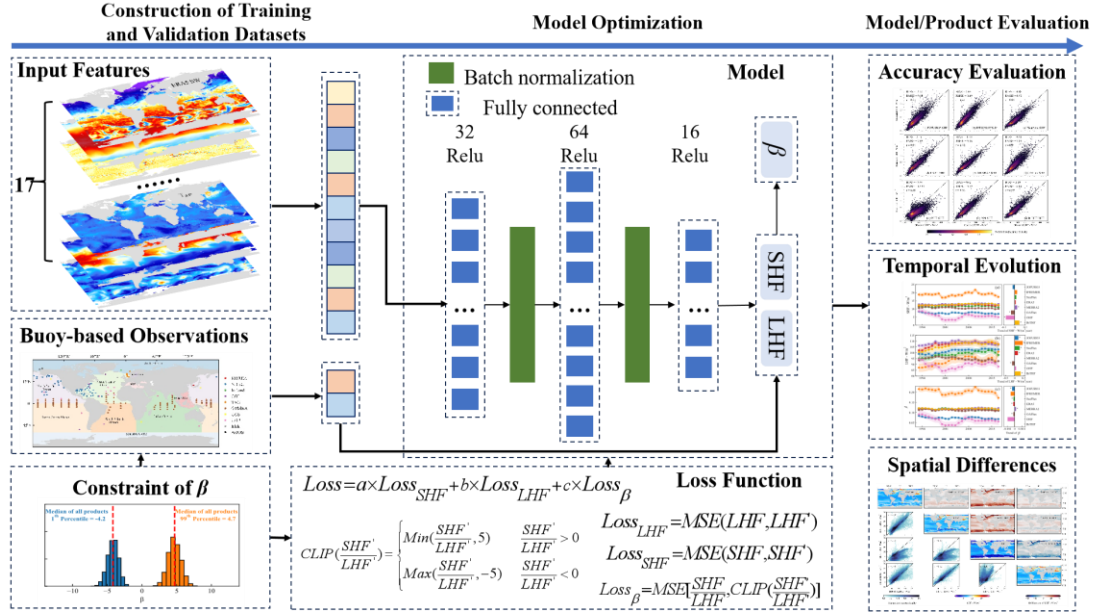
112 To improve the estimation of SHF, LHF, and β in a coordinative framework, we
113 recently proposed an innovative Bowen ratio-informed data-driven model by
114 considering the synergistic changes [on the one hand, ensuring physical consistency
115 (i.e., $\text{SHF/LHF} = \beta$); on the other hand, achieving high-accuracy estimations of SHF,
116 LHF, and β simultaneously] using a Random Forest (RF) technique (Wang et al., 2024).
117 Validation against hourly eddy covariance (EC) flux measurements from 53 historical
118 cruises demonstrated the model's superior performance, achieving high accuracy in
119 estimating SHF, LHF, and β , with results that are physically consistent. Wang et al.
120 (2024) highlights the feasibility of simultaneously estimating SHF, LHF, and β with
121 high accuracy using machine learning techniques, offering strong potential for global
122 mapping that aligns with physical consistency. However, since EC observations are
123 difficult to obtain at sea due to platform motion and airflow distortion (Bourras et al.,
124 2019; Bourras et al., 2009)—their limited spatio-temporal coverage constrains the
125 application of the model for global mapping. Buoy-based flux observations provide a
126 viable alternative. Buoy data offer globally representative flux samples with adequate

127 volume and acceptable accuracy, which have been widely used to evaluate the
128 performance of global products (Bentamy et al., 2017; Bourras, 2006; Tang et al., 2024;
129 Weller et al., 2022; Zhou et al., 2020) and support global modeling (Chen et al., 2020a)
130 and analysis (Song et al., 2024; Yan et al., 2024).

131 The primary objectives of this study are three fold: (1) to develop an innovative
132 Bowen ratio-constrained statistical model for improving the air-sea SHF, LHF and β
133 estimates (referred to as the BrTHF model hereafter) using the machine learning
134 technique and global buoy-based air-sea turbulent heat fluxes observations; (2) to
135 demonstrate the superiority of the statistical model through an inter-comparison with
136 seven widely used global products and the estimates from the physics-free machine
137 learning-based model; (3) to produce a global daily 0.25° dataset based on the BrTHF
138 model over ice-free oceans covering the period from 1993 to 2017. The flux
139 observations from 197 global distributed buoys, along with multi-source satellite-based
140 and reanalysis-based inputs, were collected to construct the models and further produce
141 the global air-sea turbulent heat fluxes dataset. The accuracy and spatio-temporal
142 patterns of the SHF, LHF and β estimates were inter-compared with seven widely used
143 products, including the remote sensing-based JOFURO v3, IFREMER v4.1 and
144 SeaFlux v3, as well as reanalysis-based ERA5 and MERRA2, hybrid-based OAFflux v3
145 and machine learning-based OHF v2 products.

146 **2. Data and Methods**

147 The following sub-sections provide an overview of the development of the BrTHF
148 product, detailing the construction of air-sea turbulent heat fluxes observation datasets,
149 learning datasets and the BrTHF model, as well as the evaluation strategies used in this
150 study, as indicated in Figure 1.



151

152 **Figure 1. flowchart of the generation of a global product of air-sea SHF, LHF and β by the**
 153 **BrTHF model**

154 **2.1 Air-sea turbulent heat fluxes observation datasets**

155 To obtain the buoy-derived air-sea turbulent heat fluxes observations, the hourly
 156 or sub-hourly oceanic and atmospheric measurements including sea surface
 157 temperature (T_s), sea surface air temperature (T_a), sea surface wind speed (WS) and
 158 relative humidity (RH) were firstly collected at 268 buoys covering a variety of ocean
 159 basins from 13 organizations or networks. Detailed information about the buoy sources
 160 and the number of buoys from each provider is summarized in Table 1.

161 For certain buoys lacking RH measurements [e.g. buoys from NDBC (National
 162 Data Buoy Center) provided dew point temperature (DEW) rather than RH], the RH
 163 was computed using DEW and T_a . To ensure the quality of the measurements, we
 164 filtered the records based on the quality control recommendations provided by the data
 165 providers. Further refinement was also made by removing the questionable values that
 166 exceeded three standard deviations (3σ) for each variable at individual buoys.

167 Once the data was cleaned, daily mean aggregation was applied to the oceanic and
 168 atmospheric measurements. Given the varying temporal resolutions of the
 169 measurements (e.g. NDBC provided hourly observations before 2005 and 10-min

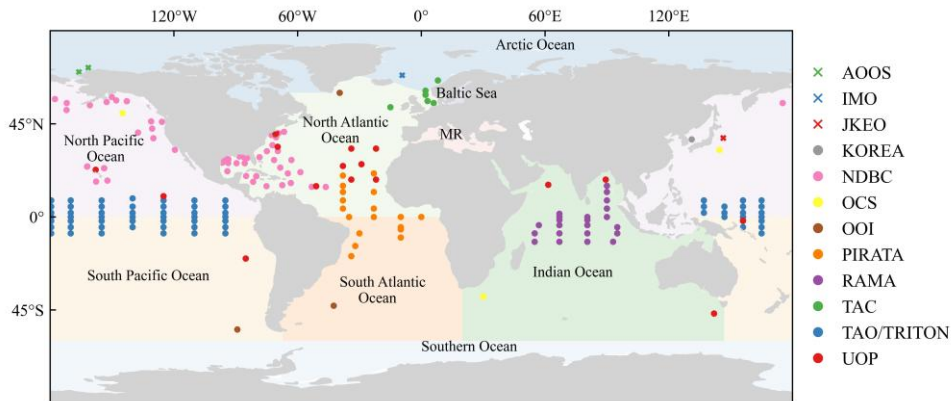
170 observations thereafter), we only retained the daily mean data when the fraction of the
 171 valid hourly or sub-hourly observations exceeded 80% on a given day.

172 **Table 1 Summary of f buoys used in this study**

Network/Organization	Number of Buoys	Network/Organization	Number of Buoys
TAO/TRITON (Tropical Atmosphere Ocean / TRITON)	67	KOREA (Korea Meteorological Administration)	24
PIRATA (Prediction and Research Moored Array)	20	OOI (Ocean Observatories Initiative)	6
RAMA (Research Moored Array for Monsoon Analysis)	23	AOOS (Alaska Ocean Observing System)	2
NDBC (National Data Buoy Center)	73	JKEO (Japan Agency for Marine-Earth Science and Tech.)	1
Copernicus Marine In Situ TAC	19	Irish Weather Buoy Network	6
UOP (Upper Ocean Processes Group)	23	Icelandic Meteorological Office (IMO)	1
OCS (Ocean Climate Stations Project)	3		

173 After the above mentioned data preprocessing, the daily buoy-derived air-sea
 174 turbulent heat fluxes (SHF and LHF) observations were then calculated using the daily
 175 oceanic and atmospheric measurements combined with the version 3.5 of Coupled
 176 Ocean-Atmosphere Response Experiment (COARE3.5) model (Edson, 2013)
 177 (available at <https://github.com/NOAA-PSL/COARE-algorithm>) with the cool-skin
 178 and warm-layer calculation switched off. The configuration follows the practice
 179 adopted by Pacific Marine Environmental Laboratory for producing daily air-sea
 180 turbulent heat flux products
 181 (<https://www.pmel.noaa.gov/tao/drupal/flux/documentation-lw.html>) at Global
 182 Tropical Moored Buoy Array (TAO/TRITON, PIRATA, and RAMA).ss Note that
 183 outliers of β present in the observations are likely associated with uncertainties in the
 184 model-derived estimates and input data. Considering that such outliers can severely
 185 impede model training and evaluation, it was necessary to constrain β within a
 186 reasonable range to enable simultaneous high-accuracy estimation of SHF, LHF, and β .
 187 Therefore, following the air-sea turbulent heat fluxes computations, we further made a

188 quality control on the derived SHF and LHF observations to exclude the abnormal
 189 records, by filtering the observations based on the range of daily β values determined
 190 from seven widely-used flux products. Specifically, we calculated the cumulative
 191 distribution of daily β for each product and their ensemble (across all products). The
 192 medians of the 1st and 99th percentiles, approximately -5 and 5, respectively, were
 193 selected as the minimum and maximum of valid daily β , as shown in Figure S1. In total,
 194 this study compiled 463,585 observations of valid daily air-sea turbulent heat flux from
 195 197 buoy stations (Figure 2 and Table S1) in the Arctic Ocean, Pacific Ocean, Atlantic
 196 Ocean and Indian Ocean.



197
 198 **Figure 2. Geographic locations of 197 buoy sites from 12 organizations or networks involved**
 199 **in this analysis including TAO/TRITON, PIRATA, RAMA, NDBC, TAC, UOP, OOI, AOOS,**
 200 **KOREA, OCS, JKEO and IMO. The boundaries of global land and open oceans were sourced**
 201 **from the Natural Earth dataset (<https://www.naturalearthdata.com/downloads/>) and the**
 202 **Global Oceans and Seas dataset (<https://www.marinerregions.org/sources.php>), respectively.**
 203 **Abbreviations MR refers to the Mediterranean Region. It should be noted that the Caspian**
 204 **Sea was not included within the boundaries of the open oceans and is shown in white.**

205 Finally, the quality-controlled observations were collected to train and validate the
 206 BrTHF model. Note that the COARE-based observations at the buoy stations have
 207 already widely applied as a benchmark for global air-sea turbulent heat flux product
 208 development and evaluation (Bentamy et al., 2017; Chen et al., 2020b; Tang et al., 2024;
 209 Weller et al., 2022)

210 **2.2 Learning datasets and state-of-the-art products**

211 **2.2.1 Learning datasets for training the neural network**

212 Learning variables were carefully selected based on their potential impacts on the
213 variations of the air-sea turbulent heat fluxes (Grist et al., 2016; Kudryavtsev et al.,
214 2014; Myslenkov et al., 2021; Song, 2020, 2021; Yan et al., 2024) to conduct the feature
215 selection (see section 2.3.1). These variables include T_a , sea surface air specific
216 humidity (Q_a), Mean Sea Level Pressure (SLP), Downward Long Wave Radiation Flux
217 (LW), Downward Short Wave Radiation Flux (SW), T_s , sea surface specific humidity
218 (Q_s), Absolute Dynamic Topography (ADT), Sea Level Anomaly (SLA), Sea Surface
219 Salinity (SSS), Sea Surface Density (SSD), Ocean Mixed Layer Current Velocity (CS),
220 WS , Significant Wave Height (SWH), Wave period (T_p), as well as gradient of
221 temperature ($diff_T$) calculated using the T_s and T_a , and gradient of humidity ($diff_Q$)
222 calculated using the Q_s and Q_a .

223 Datasets of these learning variables were collected from multiple publicly
224 available sources, as summarized in Table 2 and were used as the input features for
225 training the neural network. The Multi Observation Global Ocean Sea Surface Salinity
226 and Sea Surface Density (MOGOSD) dataset and Global Ocean Waves (GOW)
227 Reanalysis dataset were spatially resampled to a 0.25° resolution using mean
228 aggregation, while temporal mean aggregation to daily values was applied to the GOW
229 dataset (originally at 3-hour resolution) and Cross-Calibrated Multi-Platform (CCMP)
230 wind vector analysis dataset (6-hour resolution). Additionally, a daily ERA5 sea-ice
231 mask was applied to the datasets to mitigate the impact of sea ice.

232 **2.2.2 State-of-the-art products for inter-comparison**

233 Seven widely used air-sea turbulent heat flux products, involving remote sensing-
234 based JOFURO3, IFREMER and SeaFlux, as well as reanalysis-based ERA5 and
235 MERRA2, hybrid-based OAFflux and machine learning-based OHF products were
236 selected for inter-comparison.

237 The remote sensing-based JOFURO3, IFREMER, and SeaFlux products were

238 developed by the Japanese Ocean Flux Data Sets under the Remote Sensing
239 Observations (J-OFURO) research project, the Institute Français de Recherche pour
240 l'Exploitation de la Mer (IFREMER), and the NASA Global Hydrology Resource
241 Center (GHRC) Distributed Active Archive Center (DAAC), respectively. The
242 reanalysis-based ERA5 and MERRA2 products were developed by the ECMWF and
243 NASA Global Modeling and Assimilation Office (GMAO), respectively. The hybrid-
244 based OAFlux and machine learning-based OHF products were developed or published
245 by the Woods Hole Oceanographic Institution (WHOI) and NOAA Ocean Surface
246 Bundle (OSB) Climate Data Record (CDR), respectively. With the exception of the
247 OHF product calculating SHF and LHF simultaneously using a NN model without a
248 constraint, all other products employed bulk aerodynamic methods to estimate SHF and
249 LHF. The JOFURO3, IFREMER, and OAFlux products used the COARE3.0 model,
250 while the SeaFlux used the COARE3.5 model. Differently, the ERA5 adopted the bulk
251 aerodynamic method used by the ECMWF, and the MERRA2 used the method by the
252 GEOS. These products provide SHF and LHF estimates at a 0.25° spatial resolution,
253 except for the MERRA2 ($0.5^\circ \times 0.625^\circ$) and OAFlux (1°). Additionally, most products
254 provide daily SHF and LHF estimates, while only the OHF product provides estimates
255 at a 3-hour interval. For further inter-comparison, the daily mean aggregation was
256 applied to the OHF products. More details about the seven products can be found in the
257 review of Tang et al. (2024).

Table 2 Summary of learning datasets used in this study

Dataset source	Resolution	Variables	Urls
ERA5	0.25°/daily	Sea surface air temperature (T_a), sea surface air specific humidity (Q_a), mean sea level pressure (SLP), downward long wave radiation flux (LN) and downward short wave radiation flux (SN)	https://cds.climate.copernicus.eu/datasets/derived-era5-single-levels-daily-statistics?tab=overview
OSCAR	0.25°/daily	Ocean mixed layer current velocity (CS)	https://podaac.jpl.nasa.gov/dataset/OSCAR_L4_OC_FINAL_V2.0
CCMP	0.25°/6-hour	Wind speed (WS)	https://data.remss.com/ccmp/v03.0/daily/
MOGOSD	0.125°/daily	Sea surface salinity (SSS) and sea surface density (SSD),	https://data.marine.copernicus.eu/product/MULTIOBS_GLO_PHY_S_SURFACE_MYNRT_015_013/description
SSH	0.25°/daily	Absolute dynamic topography (ADT) and sea level anomaly (SLA)	https://data.marine.copernicus.eu/product/SEALEVEL_GLO_PHY_CLIMATE_L4_MY_008_057/description
GOW	0.2°/3-hour	Significant wave height (SWH) and wave period (T_p)	https://data.marine.copernicus.eu/product/GLOBAL_MULTITYEAR_WAV_001_032/description
OISST	0.25°/daily	Sea surface temperature (T_s) and sea surface specific humidity (Q_s)	https://www.ncei.noaa.gov/data/sea-surface-temperature-optimum-interpolation/v2.1/access/avhrr/
OISST – ERA5	0.25°/daily	Gradient of temperature ($diff_T$) and gradient of humidity ($diff_Q$)	-

Table 3 Summary of the state-of-the-art air-sea turbulent heat flux products used for inter-comparison in this study

Dataset source	Resolution	Model	Variables	Urls
JOFURO3	0.25°/daily	COARE3.0		https://www.j-ofuro.com/en/
IFREMER	0.25°/daily	COARE3.0		ftp://ftp.ifremer.fr/ifremer/cersat/data/heat-flux/ifremer/v4.1/daily
SeaFlux	0.25°/daily	COARE3.5		https://www.earthdata.nasa.gov/data/catalog/ghrc-daac-seaflux-1
MERRA2	0.5° × 0.625°/daily	GEOS	Latent heat flux (LHF), sensible heat flux (SHF) and Bowen ratio ($\beta =$ SHF/LHF)	https://developers.google.com/earth-engine/datasets/catalog/NASA_GSFC_MERRA_flux_2?hl=zh-cn#bands
ERA5	0.25°/daily	ECMWF	Bowen ratio ($\beta =$ SHF/LHF)	https://cds.climate.copernicus.eu/datasets/derived-era5-single-levels-daily-statistics?tab=overview
OAFlux	1°/daily	COARE3.0		ftp://ftp.whoi.edu/pub/science/oaflux/data_v3
OHF	0.25°/3-hour	Neural Network model		https://www.ncei.noaa.gov/products/climate-data-records/ocean-heat-fluxes

263 2.3 Construction of the BrTHF model

264 2.3.1 Feature selection

265 The study employed a random forest (RF) model to evaluate the importance scores
266 of 17 oceanic and atmospheric learning variables (with datasets collected in Section 2.2)
267 for target variables (SHF and LHF), aiming to filter out less influential variables. As
268 shown in Table S2, the variable importance assessment revealed that $diff_T$ and $diff_Q$
269 showed the highest importance score (71.56% and 49.93%) for SHF and LHF
270 modelling, respectively; additionally, WS exhibited the second highest importance for
271 both SHF (10.19%) and LHF (27.59%) modelling. Building upon the importance
272 evaluation and through careful screening of highly correlated variables, we ultimately
273 selected 11 key environmental features for subsequent air-sea turbulent heat flux
274 modelling including SLP , LW , SW , SSS , ADT , CS , WS , SWH , T_p , $diff_Q$, and $diff_T$.

275 2.3.1 Model construction and optimization

276 We selected the NN technique to build the BrTHF model due to its strong ability
277 to capture the complex nonlinear relationships between the multiple inputs and
278 multiple-target variables with high accuracy (Zhou et al., 2024; Fu et al., 2023;
279 Cummins et al., 2023; Cummins et al., 2024). Additionally, the technique enables the
280 seamless integration of physical constraints, improving the reasonableness of the results
281 (Zhou et al., 2024; Zhao et al., 2019; Shang et al., 2023).

282 In order to estimate the SHF and LHF with high accuracy in a physics-consistency
283 framework, the β (= SHF/LHF) physical constraint was incorporated into the NN model
284 using the custom multiple-objects (SHF, LHF and β) loss function as follows:

$$285 \quad LOSS = a \times LOSS_{SHF} + b \times LOSS_{LHF} + c \times LOSS_{\beta} \quad (1)$$

286 $LOSS_{SHF}$, $LOSS_{LHF}$ and $LOSS_{\beta}$ represent the Mean Squared Error (MSE) of SHF, LHF and
287 β , respectively. They were weighted using the factors of a (SHF), b (LHF) and c (β) to
288 balance the different magnitudes of loss during optimization. To prevent potential
289 gradient explosion during model training, predicted β [SHF'/LHF' , calculated using the
290 predicted SHF (SHF') and LHF (LHF')] values were clipped within the observed range

291 of β (from -5 to 5) during training:

$$292 \quad CLIP\left(\frac{SHF'}{LHF'}\right) = \begin{cases} \text{Min}\left(\frac{SHF'}{LHF'}, 5\right) & \frac{SHF'}{LHF'} > 0 \\ \text{Max}\left(\frac{SHF'}{LHF'}, -5\right) & \frac{SHF'}{LHF'} < 0 \end{cases} \quad (2)$$

$$293 \quad Loss_{\beta} = MSE\left(\frac{SHF'}{LHF'}, CLIP\left(\frac{SHF'}{LHF'}\right)\right) \quad (3)$$

294 Finally, after optimization, the final weights (a, b and c) for SHF, LHF, and β were
295 set to 5, 1, and 250, respectively. The model was constructed consisting of one input
296 layer, three hidden layers, two BatchNormalization layers, and one output layer using
297 the Python TensorFlow library. The numbers of neurons in the three hidden layers were
298 32, 64, and 16, respectively and the activation function of Leaky Rectified linear unit
299 (ReLU) was used throughout the model.

300 To illustrate the superiority of the BrTHF model in terms of accuracy and physical
301 consistency, another physics-free NN models, trained without integrating the β
302 constraint, were also constructed to predict SHF and LHF separately for further
303 comparison, where β was calculated to be SHF/LHF.

304 **2.4 Evaluation strategy**

305 A spatial 10-fold cross-validation was employed to assess the performance of the
306 BrTHF model for estimating air-sea SHF, LHF and β . Compared to the traditional 10-
307 fold cross-validation, which randomly split all samples into ten folds and thus may
308 result in overlapping spatial samples between training and validating datasets, the
309 spatial 10-fold cross-validation was conducted in a relatively independent spatial
310 distribution and can provide a more robust and generalizable evaluation.

311 Specifically, first, all buoy sites were randomly split into ten folds. Then, each fold
312 was in succession selected as the validation dataset and the remaining nine folds was
313 used as the training dataset.

314 The metrics used to evaluate the performance of the models include: (1) the mean
315 bias error (BIAS); (2) the root mean squared error (RMSE); (3) the correlation
316 coefficient (r).

317
$$BIAS = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i) \quad (4)$$

318
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (5)$$

319
$$r = \frac{\sum_{i=1}^n [(\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y})]}{\sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (6)$$

320 where n is the number of samples, \hat{y}_i and y_i are the estimated value and reference
 321 truth, $\bar{\hat{y}}$ and \bar{y} are the mean of \hat{y}_i and y_i , respectively. These metrics—BIAS,
 322 RMSE, and r —comprehensively evaluate model performance, representing systematic
 323 deviation, dispersion between observations and estimates, and the strength and
 324 direction of the linear relationship, respectively. Note that RMSE and r can be sensitive
 325 to extreme values, particularly for the β , which is a ratio-based variable and may exhibit
 326 unrealistically large magnitudes under near-zero flux conditions. To ensure a fair
 327 evaluation, model performance is assessed both with all samples retained and with
 328 extreme β values excluded in subsequent analyses. This dual evaluation allows us to
 329 quantify overall model performance while explicitly accounting for the influence of
 330 rare extreme cases.

331 **3. Results and discussion**

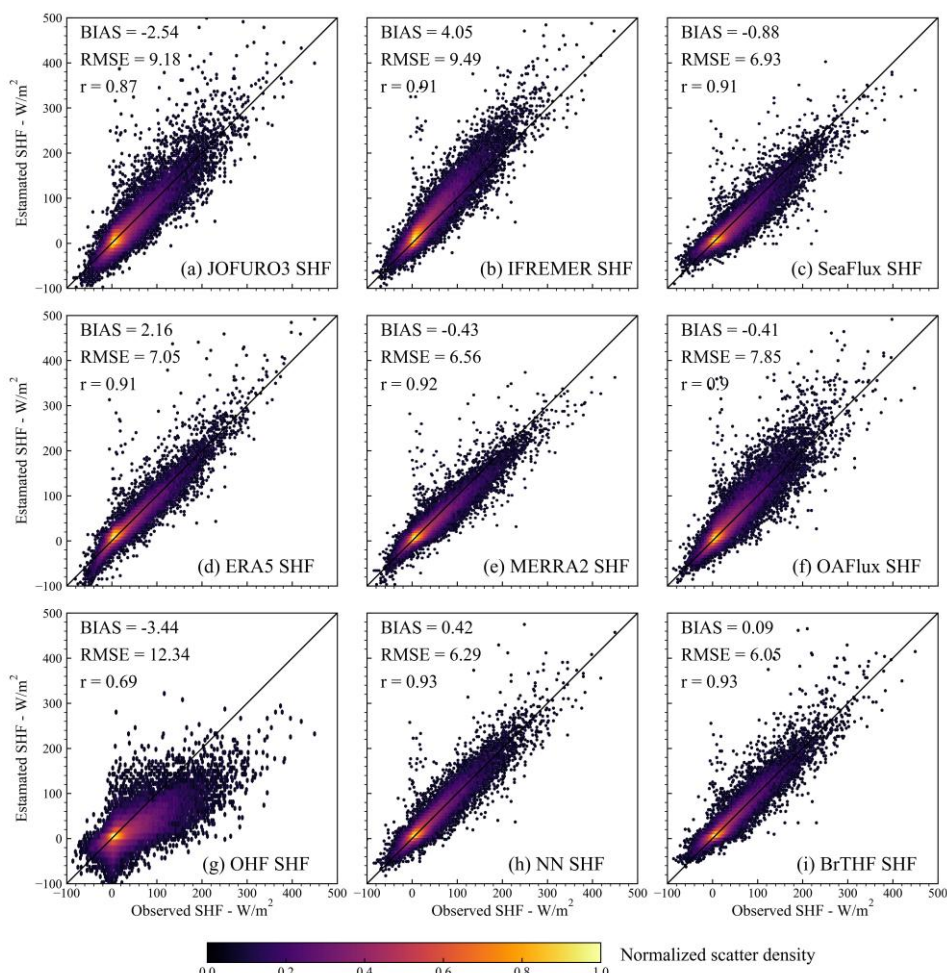
332 **3.1 Spatial ten-fold cross-validation of the models**

333 **3.1.1 Overall accuracy**

334 Figures 3, 4 and 5 present the normalized scatter density plots of the estimated
 335 daily SHF, LHF and β from the BrTHF and physics-free NN models, as well as the
 336 seven air-sea turbulent heat flux products against the observations obtained from 197
 337 global distributed buoys by the spatial ten-fold cross-validation strategy.

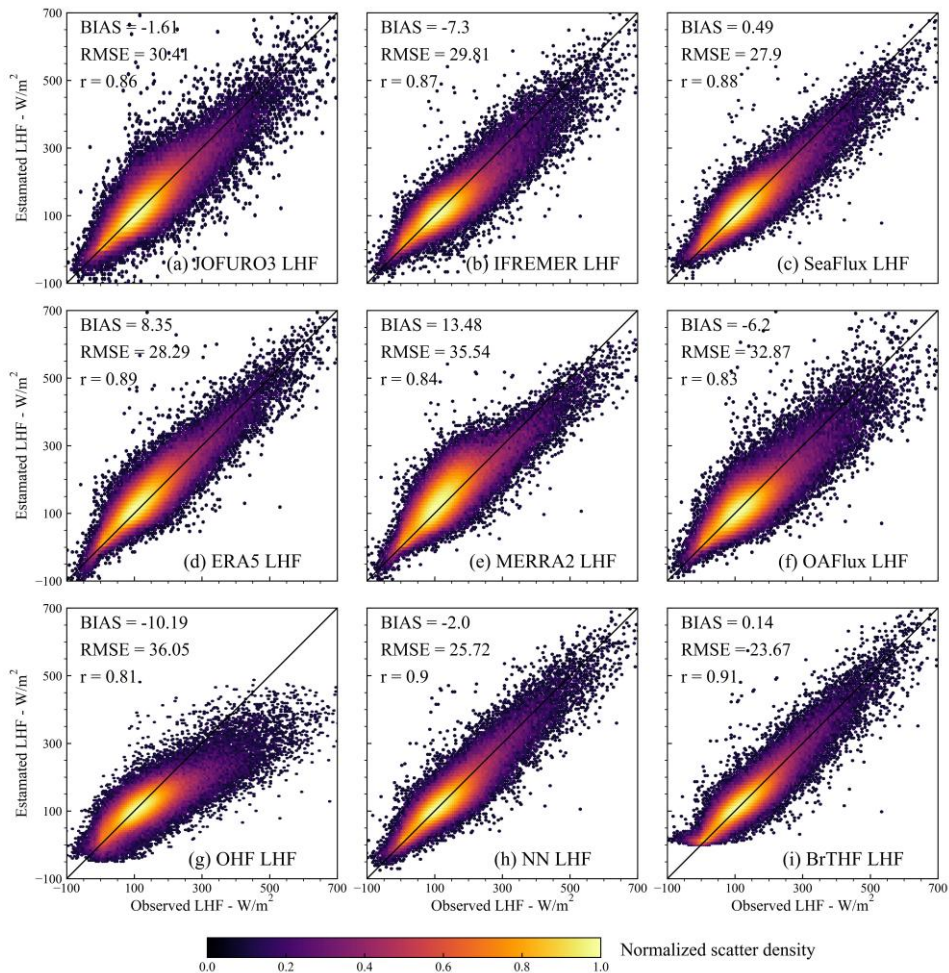
338 Most models and products showed data distributions closely aligned with the
 339 observed SHF, with the majority of samples clustered around the 1:1 line. The BrTHF
 340 model slightly overestimated SHF with a BIAS of 0.09 W/m², whereas the physics-free
 341 NN models, and ERA5 and IFREMER products showed more pronounced
 342 overestimations (from 0.42 W/m² to 4.05 W/m²). In contrast, the remaining five

343 products exhibited notable underestimations (from -3.44 W/m^2 to -0.41 W/m^2). As
 344 illustrated in Figure 6, the variability of estimated SHF from the BrTHF and the
 345 physics-free NN models and ERA5 product closely matched the observed SHF, all with
 346 a Standard Deviation (STD) of approximately 16 W/m^2 . Notably, the BrTHF model
 347 achieved the lowest RMSE (6.05 W/m^2), outperforming both the physics-free NN
 348 models (6.29 W/m^2) and the seven air-sea turbulent heat flux products (ranging up to
 349 12.34 W/m^2 for OHF). Additionally, the BrTHF model combined with the physics-free
 350 NN models yielded the highest values of r (0.93), surpassing all seven other products.
 351 In summary, the BrTHF model showed the best overall performance in estimating SHF
 352 among all the models and products.



353
 354 **Figure 3. Normalized scatter density plots of estimated SHF from the BrTHF model, the**
 355 **physics-free NN models and seven air-sea turbulent heat flux products against the observed**
 356 **SHF obtained from 197 global distributed buoys.**

357 For LHF, similar to the results for SHF, the BrTHF model demonstrated the best
 358 agreement with observations, achieving the lowest RMSE (23.67 W/m^2) and the highest
 359 value of r (0.91). Compared to the physics-free NN models and seven products, the
 360 BrTHF model reduced RMSE by 2.05 W/m^2 (physics-free NN models) to 12.38 W/m^2
 361 (OHF) and improved r by 0.01 (physics-free NN model) to 0.1 (OHF). Additionally, the
 362 BrTHF model showed a slight overestimation of LHF (BIAS = 0.14 W/m^2), lower than
 363 that of the SeaFlux, MERRA2, and ERA5 products. In contrast, the remaining products
 364 (JOFURO3, IFREMER, OAFflux, and OHF), along with the physics-free NN models,
 365 underestimated LHF, with the BIAS values ranging from -10.19 W/m^2 (OHF) to -1.61
 366 W/m^2 (JOFURO3).

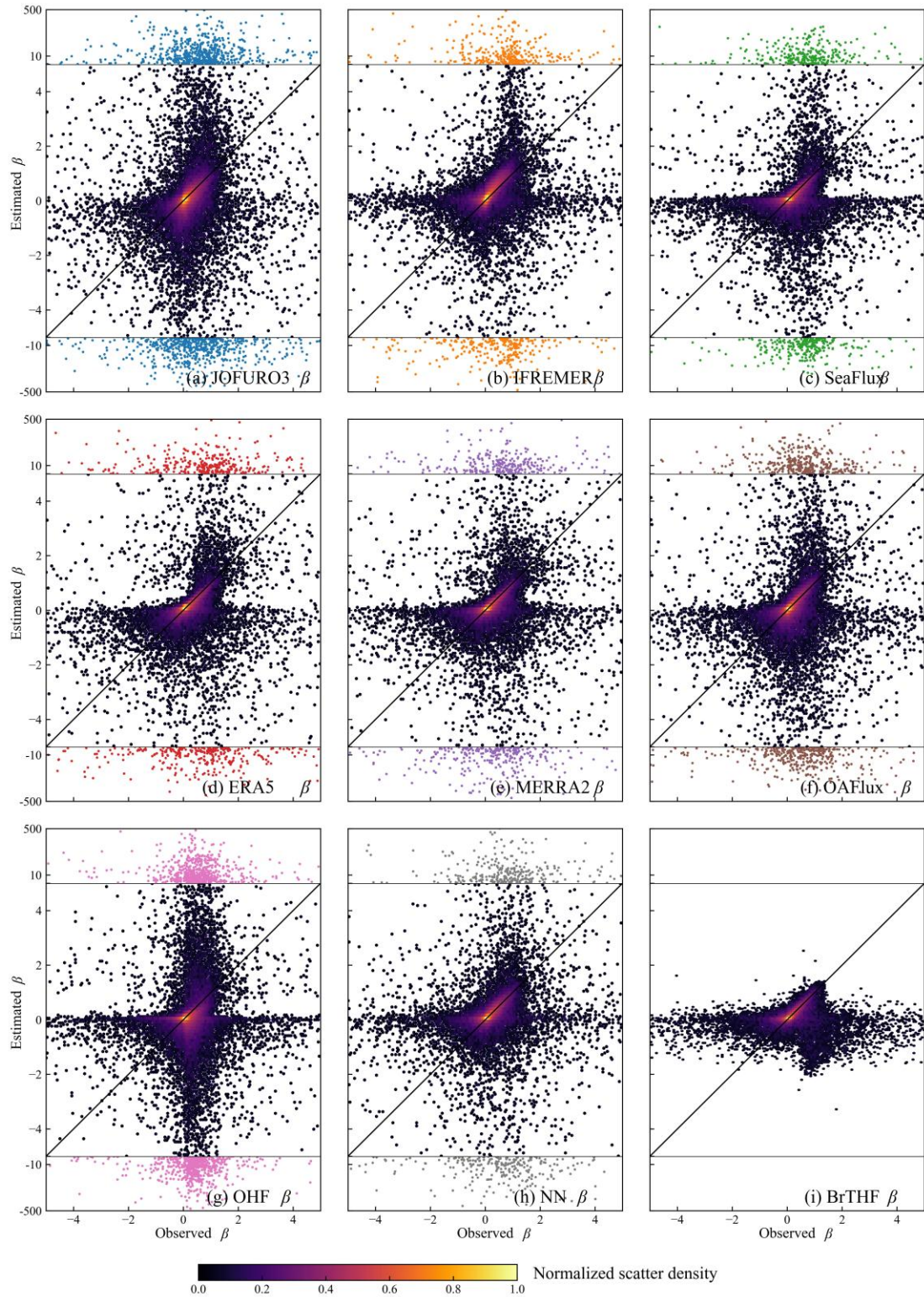


367

368 **Figure 4. Same as Figure 3 but for LHF.**

369 The BrTHF model exhibited a significantly different distribution of β compared to
 370 the physics-free NN models and the seven products, as shown in Figure 5. The β

371 estimates from the BrTHF model consistently fell within the observed range of -5 to 5,
372 while the physics-free NN model and the seven products occasionally produced
373 estimates outside this range. Specifically, approximately 0.9% of β estimates from both
374 the physics-free NN model and the seven products were out of range. The extreme
375 positive and negative β estimates were found in the OHF ($\beta = 14997$) and physics-free
376 NN models ($\beta = -25703$) products, respectively. The abnormal β estimates significantly
377 impacted the accuracy of the physics-free NN models and the seven products as Figure
378 6 indicated. When excluding the abnormal β samples from the physics-free NN models
379 and seven products, the RMSEs ranged from 0.17 (physics-free NN models and
380 SeaFlux) and 0.26 (OHF), with values of r ranging from 0.13 (OHF) to 0.46
381 (IFREMER), as shown in Figure 6 and Table S3. However, when all estimates were
382 considered, the performances of these model and products deteriorated sharply, with
383 RMSEs rising from 0.87 (SeaFlux) to 39.21 (physics-free NN models), and values of r
384 dropping from 0.06 (SeaFlux) to 0 (JOFURO3, MERRA2 and OHF). In contrast, the
385 BrTHF model maintained robust outperformance, with the lowest RMSEs of 0.22 and
386 0.15, and higher r values of 0.25 and 0.43, both before and after removing the abnormal
387 β samples from the physics-free NN models and the seven products. Notably, the BIAS
388 values remained stable (ranging from -0.04 to 0.04) for all models and products,
389 regardless of whether the abnormal samples were excluded.



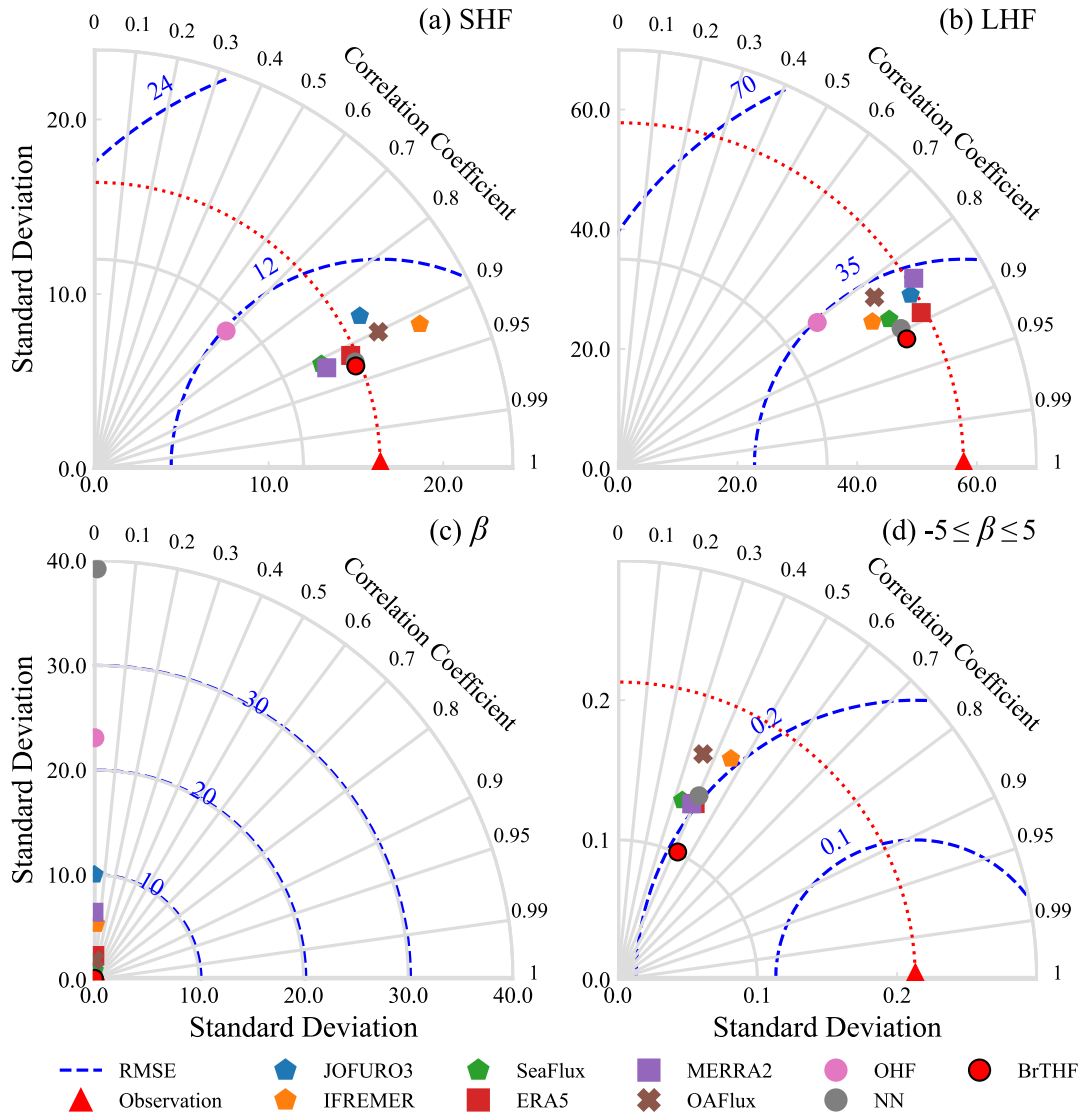
390

391 **Figure 5. Same as Figure 3 but for β . The samples out of the ranges of observed β ($-5 \leq \beta \leq 5$)**

392 **were colored in blue, orange, green, red, purple, brown, pink and gray for JOFURO3,**

393 **IFREMER, SeaFlux, ERA5, MERRA2, OAFflux, OHF products and the physics-free NN**

394 **models, respectively. The statistical metrics could be found in Table S3 and Figure 6.**



395

396 **Figure 6. Taylor diagrams of the validation of estimated daily SHF (a), LHF (b), β (c) and β ($-5 \leq \beta \leq 5$, d) from the BrTHF model, the physics-free NN models and the seven products against**
 397 **the in-situ observations.**
 398

399

400 3.1.2 Accuracies across oceans

401 To better understand the accuracy of SHF, LHF and β estimates from the BrTHF
 402 and physics-free NN models, as well as the seven products in different oceans, we
 403 conducted an additional evaluation by categorizing the observations according to the
 404 belonging ocean basins, as shown in Figure 7. The major ocean boundaries, obtained
 405 from Marine Regions (<https://www.marineregions.org/>), were used to define the ocean

406 basins, which include the Arctic Ocean, South Pacific Ocean, North Pacific Ocean,
407 South Atlantic Ocean, North Atlantic Ocean, and Indian Ocean.

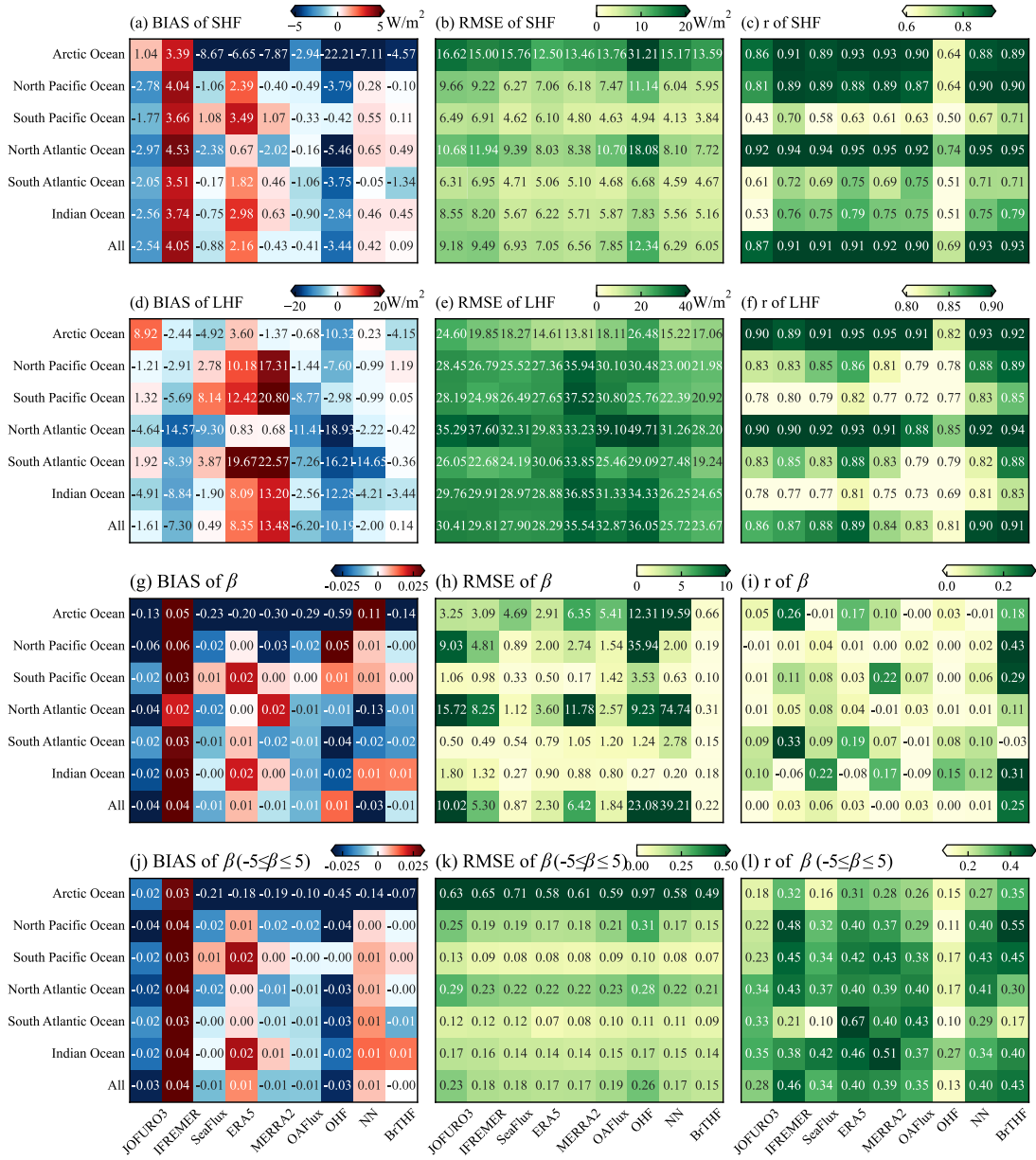
408 For SHF, the BrTHF model exhibited overestimations in the South Pacific Ocean,
409 North Atlantic Ocean, and Indian Ocean, while it underestimated SHF in the remaining
410 three ocean basins. The values of BIAS ranged from -4.57 W/m^2 in the Arctic Ocean to
411 0.49 W/m^2 in the North Atlantic Ocean. Furthermore, the BrTHF achieved the lowest
412 RMSEs in most ocean basins, ranging from 3.84 W/m^2 in the South Atlantic Ocean to
413 7.72 W/m^2 in the North Atlantic Ocean, except in the Arctic Ocean where the RMSE of
414 13.59 W/m^2 was higher than those of the ERA5 (12.5 W/m^2) and MERRA2 (13.46
415 W/m^2) products, as shown in Figure 7(b). Correlation analysis also demonstrated the
416 robust performance of the BrTHF model in estimating SHF, with values of r exceeding
417 0.89 in most ocean basins, except for ocean basins in the Southern Hemisphere (ranging
418 from 0.71 to 0.79) where the values of r for all models and products were reduced.

419 For LHF, the values of BIAS of the BrTHF model ranged from -4.15 W/m^2 in the
420 Arctic Ocean to 1.19 W/m^2 in the North Pacific Ocean. In comparison, the BrTHF
421 model showed more pronounced underestimations in the Arctic Ocean and Indian
422 Ocean. Additionally, the BrTHF model outperformed the physics-free NN models and
423 the seven products across most ocean basins, achieving the lowest RMSEs (ranging
424 from 17.06 W/m^2 in the Arctic Ocean to 28.20 W/m^2 in the North Atlantic Ocean) and
425 the highest values of r (ranging from 0.83 in the Indian Ocean to 0.94 in the North
426 Atlantic Ocean) except for the Arctic Ocean where the value of r was 0.01 less than the
427 physics-free NN models and the RMSE were 2.45 W/m^2 , 3.25 W/m^2 and 1.84 W/m^2
428 higher than the ERA5 and MERRA2 products and the physics-free NN models,
429 respectively. Combining the overall and regional evaluations, we find that, except for
430 the OHF product, the relative performance of SHF and LHF among the seven products
431 is generally consistent with Tang et al. (2024). The OHF product shows notable
432 degradation, particularly in high-latitude oceans, which may be due to limitations in the
433 training datasets (sparse and unevenly distributed in-situ observations) and model

434 training strategy (randomly splitting all observations into training, validation, and test
435 sets without accounting for spatial dependencies), leading to reduced generalizability.

436 The BrTHF model consistently performed better in estimating β across most ocean
437 basins, both before and after removing the abnormal β samples that deviated from the
438 observed range ($-5 \leq \beta \leq 5$). In contrast, the physics-free NN models and the seven
439 products did not perform as well. Specifically, the BrTHF model exhibited the lowest
440 RMSEs in almost all ocean basins except in the South Atlantic Ocean after removing β
441 outliers. Moreover, in terms of correlation analysis, the BrTHF model achieved higher
442 values of r in most ocean basins before and after the removal of abnormal β samples,
443 among all models and products.

444



445

446 **Figure 7. Heatmaps of BIAS, RMSE and r metrics for the validation of estimated daily SHF**

447 **(a - c), LHF (b - e), β (f - i) and β ($-5 \leq \beta \leq 5$, j - l) from the BrTHF model, the physics-free NN**

448 **models and the seven products against the in-situ observations across different ocean basins.**

449 **It should be noted that the statistical metrics for each ocean basin were calculated using**

450 **observations from the available buoys within the corresponding basin.**

451

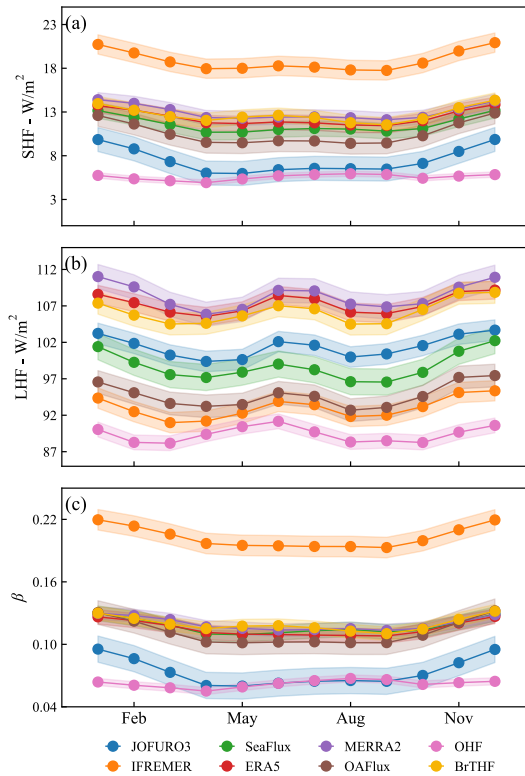
452 **3.2 Temporal variations in SHF, LHF and β**

453 **After spatial ten-fold cross-validation, we produced the daily 0.25° global air-sea**

454 **turbulent heat flux products from 1993 to 2017 using a combination of the BrTHF**

455 model and learning datasets, and further made a comparison of the temporal variation
456 (in this section), spatial distribution (in Section 3.3) and annual trend (in Section 3.4)
457 of SHF, LHF and β estimates from the BrTHF product and those with the seven state-
458 of-the-art global products. The selected period (from 1993 to 2017) was determined by
459 the overlapping availability of input learning datasets.

460 Figure 8 illustrates the monthly area-weighted global means of SHF, LHF and β
461 from 1993 to 2017 for the BrTHF product and seven state-of-the-art products. The
462 BrTHF product exhibited similar bimodal patterns for SHF, LHF and β as the seven
463 products, with peaks in December-January and May-June-July-August. In addition, the
464 peak in May-June-July-August was less pronounced for SHF and β compared to that
465 for LHF. The monthly area-weighted global means of SHF and β from the BrTHF
466 product were higher than those of most products, except for the MERRA2 product in
467 January, February, March, April, July, August and September, and the IFREMER
468 product in all months. For LHF, the BrTHF showed lower values than the ERA5 and
469 MERRA2 products across all months. Notably, the patterns of SHF and β from the OHF
470 product, with the highest peak occurring in August and smoother intra-annual cycles,
471 differed from those of the corresponding BrTHF product and the other six products
472 developed using the bulk aerodynamic methods.

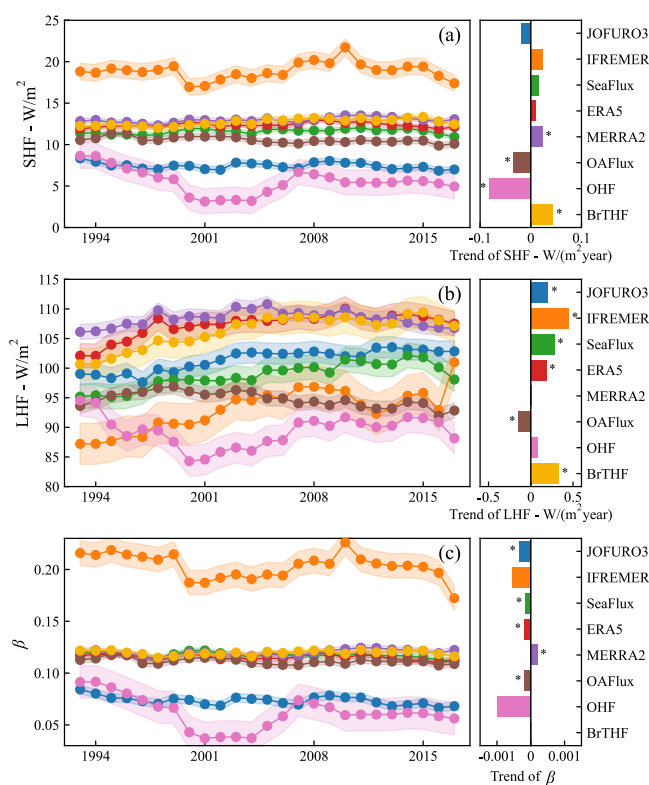


473

474 **Figure 8. Intra-annual cycles of area-weighted global monthly mean of SHF (a), LHF (b) and**
 475 **β (c) from the eight products from 1993 to 2017. The shaded areas indicate ± 1 standard**
 476 **deviation around the mean.**

477 Figure 9 presents the temporal evolution of the area-weighted annual global mean
 478 of SHF, LHF and β from 1993 to 2017 for the eight products over ice-free oceans. The
 479 global mean annual SHF of the BrTHF product was 12.7 W/m^2 , which was close to
 480 those of SeaFlux (11.6 W/m^2), OAFflux (10.6 W/m^2), MERRA2 (13 W/m^2) and ERA5
 481 (12.4 W/m^2), whereas it significantly lower than that of IFREMER (18.8 W/m^2) and
 482 higher than those of JOFURO3 (7.5 W/m^2) and OHF (5.6 W/m^2). Meanwhile, the
 483 BrTHF product exhibited the largest significant increase of SHF with the trend of 0.04
 484 $\text{W}/(\text{m}^2 \cdot \text{year})$ among all eight products, and showed similar temporal evolution as
 485 SeaFlux, MERRA2, ERA5 and OAFflux during the period from 1993 to 2017. As for
 486 LHF, the BrTHF exhibited a relatively high global mean annual value of 106.2 W/m^2 ,
 487 which was close to those of the ERA5 (107.3 W/m^2) and MERRA2 (108.3 W/m^2), and
 488 it was significantly higher than the rest of the products. Moreover, the growth of the
 489 BrTHF LHF was significant with a trend of $0.33 \text{ W}/(\text{m}^2 \cdot \text{year})$, which was lower than

490 the IFREMER but higher than the OAFflux, MERRA2, OHF, ERA5, JOFURO3 and
 491 SeaFlux, ranging from $-0.14 \text{ W}/(\text{m}^2 \cdot \text{year})$ to $0.4 \text{ W}/(\text{m}^2 \cdot \text{year})$. Note that only the
 492 OAFflux product showed a negative trend of LHF from 1993 to 2017. For β , the BrTHF
 493 showed a similar temporal pattern to that of SHF, and most products concentrated
 494 within the narrow range of 0.11 to 0.12 for the annual values. The magnitude of annual
 495 β of the BrTHF was about 0.11, which was close to the OAFflux, SeaFlux, MERRA2
 496 and ERA5, but significantly lower than the IFREMER and higher than the JOFURO3
 497 and OHF. Moreover, in contrast to the significant increasing trends of LHF and SHF,
 498 negative trends of β were shown for most products. However, the BrTHF product
 499 exhibited a weak positive trend, which may be attributed to the relatively smaller
 500 differences between the SHF and LHF trends in BrTHF compared to those in other
 501 products.

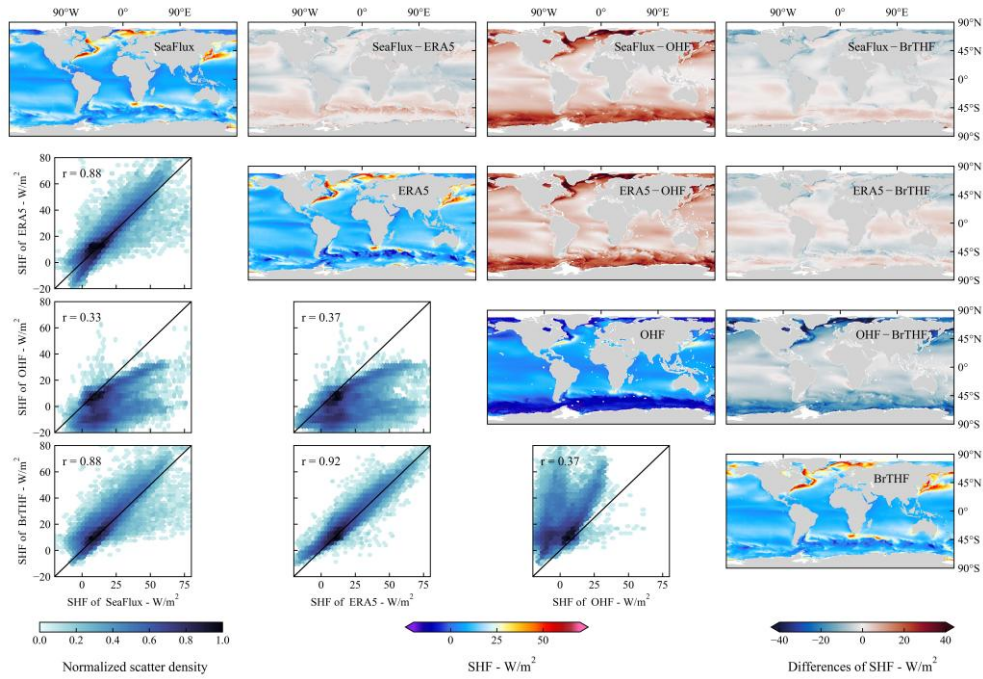


502
 503 **Figure 9. Inter-annual evolution of area-weighted global mean SHF (a - b), LHF (c - d) and β**
 504 **(e - f) from 1993 to 2017. The trends were calculated based on the Sen's slope method. The ***
 505 **in the sub-figures (b, d and f) represent the trend passed the Mann-Kendall significant test (p**
 506 **< 0.05). The shaded areas indicate ± 1 standard deviation around the mean.**

507 3.3 Inter-comparison of the spatial pattern

508 We selected three representative products including the (reanalysis-based) ERA5,
509 (remote sensing-based) SeaFlux, and (the only publicly available machine learning-
510 based) OHF products to evaluate the BrTHF product's ability in simulating global air-
511 sea turbulent heat fluxes (SHF, LHF, and β) from 1993 to 2017. These products were
512 chosen because they demonstrated relatively high accuracy within their respective
513 categories (as shown in Section 3.1) and shared the same 0.25° spatial resolution with
514 the BrTHF product.

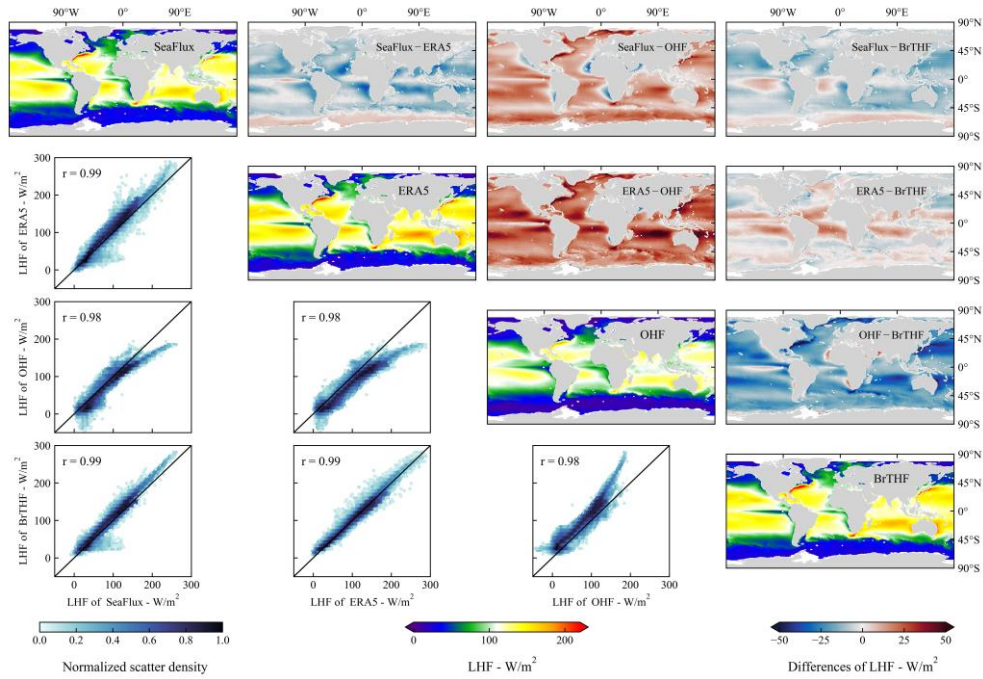
515 Figure 10 presents the spatial distribution of multi-year mean of SHF from the
516 SeaFlux, ERA5, BrTHF, and OHF products, along with their cross-comparisons.
517 Overall, the BrTHF product exhibited strong consistency with ERA5 and SeaFlux
518 products, with values of r exceeding 0.88, which was significantly higher than the
519 consistency between SeaFlux and OHF ($r = 0.33$) and between ERA5 and OHF ($r =$
520 0.37). Spatially, the BrTHF, SeaFlux and ERA5 products all showed higher SHF (over
521 50 W/m^2) in the Western Boundary Currents (WBCs, e.g. Kuroshio, Gulf Stream, Brazil
522 Current and Agulhas Current) regions, whereas the OHF product yielded much lower
523 SHF ($\sim 25 \text{ W/m}^2$). Additionally, the former three products captured pronounced SHF
524 gradients in the Southern Ocean, features that were absent in the OHF product. SHF
525 differences between BrTHF and SeaFlux/ERA5 remained within $\pm 10 \text{ W/m}^2$ in most
526 oceans. The BrTHF product exhibited slightly higher SHF values than SeaFlux in the
527 Northern Hemisphere, whereas in the Southern Hemisphere—particularly over the
528 Southern Ocean—the BrTHF showed relatively lower SHF. Compared to the ERA5
529 product, the BrTHF product yielded lower SHF in the equatorial zone, subtropical high-
530 pressure regions and the Southern Ocean, but higher SHF in other areas, particularly in
531 the North Pacific and the southern Indian Ocean.



532

533 **Figure 10. Inter-comparison of the spatial distributions of multi-year means of SHF among**
 534 **the SeaFlux, ERA5, OHF and BrTHF products from 1993 to 2017.**

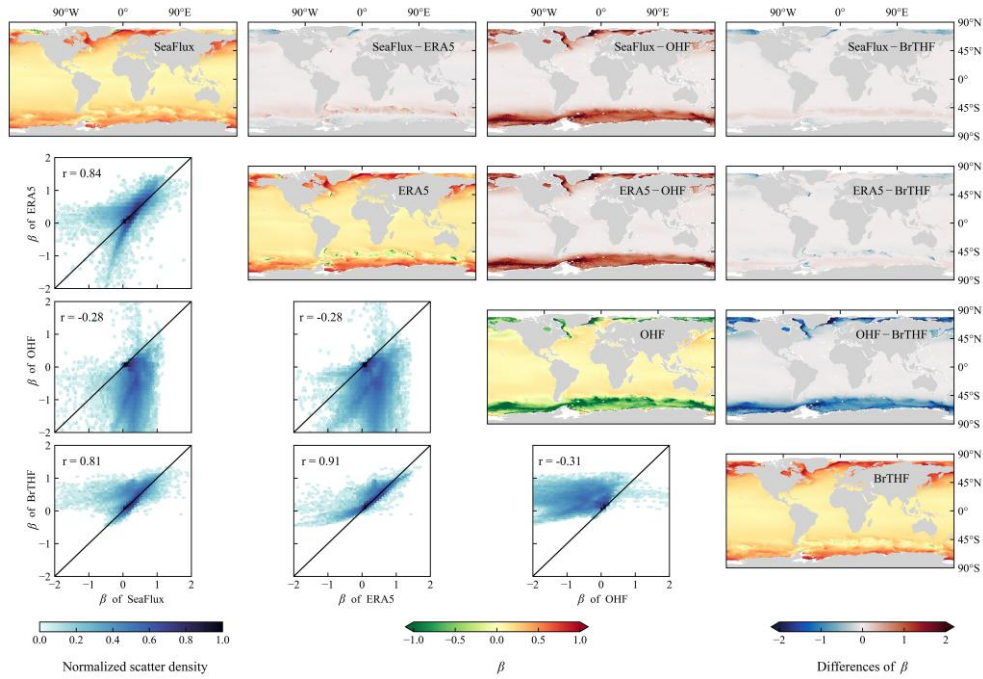
535 For LHF, the BrTHF and the three reference products exhibited more similar
 536 spatial distribution patterns, with the values of r exceeding 0.98, compared to the results
 537 for the SHF, as shown in Figure 11. The higher LHF (over 150 W/m^2) primarily
 538 occurred around the regions of WBCs and the sub-tropic highs, while lower LHF
 539 (below 50 W/m^2) appeared in the Eastern Equatorial Pacific and Atlantic Warm Tongue
 540 and the oceans with latitudes higher than 45° . The spatial distribution of LHF in the
 541 BrTHF product generally agreed better with that of the ERA5 product, though the
 542 BrTHF showed significantly lower LHF in sub-tropic highs. Additionally, the BrTHF
 543 exhibited relatively lower LHF than the ERA5 over the Southern Ocean and the central
 544 North Atlantic. Compared to the SeaFlux, the BrTHF yielded slightly higher LHF in
 545 most oceans except the Southern Ocean and equatorial zones.



546

547 **Figure 11. Same as Figure 10 but for LHF.**

548 For β , the BrTHF product demonstrated strong spatial correlation with the ERA5
 549 and SeaFlux in multi-year mean distributions, with values of r exceeding 0.81. In
 550 contrast, the OHF showed a markedly different spatial pattern of β , exhibiting negative
 551 correlations when compared to the other three products. Spatially, the BrTHF product's
 552 β distribution aligned more closely with the SeaFlux, both displaying higher β (up to 1)
 553 in high-latitude oceans particularly in the Northern Hemisphere and the similar
 554 wavelike textures of β over the Southern Ocean's Antarctic Circumpolar Current zone.
 555 The differences between the BrTHF and OHF products were more evident. Specifically,
 556 the BrTHF product showed overall overestimation of β in the oceans where latitudes
 557 were larger than 45° compared to the OHF product.



558

559 **Figure 12. Same as Figure 10 but for β .**

560

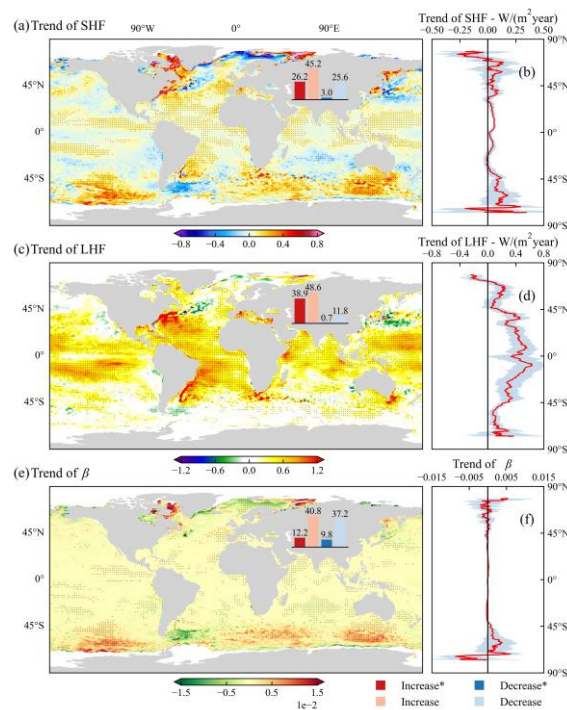
561 3.4 Spatial pattern of trends in SHF, LHF and β from the BrTHF product

562 Figure 13 illustrates the spatial distribution of inter-annual trends of SHF, LHF and
 563 β in the BrTHF product from 1993 to 2017. The SHF showed increasing trends across
 564 71.4% of the oceans, with statistically significant increases in 26.2% of regions. In
 565 contrast, decreasing trends were observed in 28.6% of the oceans, with only 3%
 566 showing significant reductions. Overall, the trends of zonal annual averages of SHF
 567 remained stable between 60°N to 45°S, with significant increases occurring southward
 568 and decreases northward. Specifically, moderate increases ($\sim 0.2 \text{ W}/(\text{m}^2 \text{ year})$)
 569 dominated between 45°N and 45°S, while more pronounced increases ($> 0.8 \text{ W}/(\text{m}^2$
 570 $\text{ year})$) were observed in mid- and high-latitude oceans, including the Kara Sea, Gulf
 571 Stream, Baffin Bay, Brazil Current, Sea of Okhotsk, and Sea of Japan. Notable
 572 decreases ($< -0.8 \text{ W}/(\text{m}^2 \text{ year})$) were concentrated in the Barents Sea and the central
 573 North Atlantic.

574 The LHF exhibited markedly different characteristics of the spatial distribution,
 575 with 87.5% of oceans showing increasing trends (38.9% were significant), versus 12.5%
 576 decreasing (0.7% were significant). In contrast to those of the SHF, the trends of zonal

577 annual averages for LHF weakened poleward from the oceans of Equator. The
 578 substantial increases ($>0.6 \text{ W}/(\text{m}^2 \text{ year})$) occurred in the oceans between 45°N to 45°S ,
 579 particularly in the Gulf Stream, Brazil Current, and Agulhas Current systems, while
 580 notable decreases (lower than $-0.3 \text{ W}/(\text{m}^2 \text{ year})$) were observed in the central North
 581 Atlantic and Kuroshio extension regions.

582 For β , approximately 53% of the oceans showed increasing trends, with 12.2% of
 583 these being statistically significant. Conversely, about 47% of the oceans showed
 584 decreasing trends, with 9.8% being significant. Most oceans between 45°N to 45°S
 585 exhibited near-zero trends, while significant trends were concentrated in the high-
 586 latitude oceans. Notable increases were found in Baffin Bay, Kara Sea, and the Southern
 587 Ocean, while decreases were observed in the Barents Sea and the Southern Ocean near
 588 South America.



589
 590 **Figure 13. Spatial maps of inter-annual trends for SHF (a), LHF (c), and β (e) from the BrTHF**
 591 **product for the period 1993 to 2017. The trends were calculated using the Sen’s slope method.**
 592 **Dotted areas indicate oceans where the p-value of the Mann-Kendall significance test is less**
 593 **than 0.05. Panels (b), (d) and (f) represent the inter-annual trends of zonal annual averages**
 594 **for SHF, LHF and β , respectively.**

595 **3.5 Discussion**

596 Advancing our understanding of the air-sea interaction and achieving the global
597 closure of the ocean surface energy budget require accurate global-scale simulations of
598 air-sea turbulent heat fluxes (Yu, 2019). Existing global air-sea turbulent heat flux
599 products, primarily generated using the semi-empirical bulk aerodynamic methods and
600 data-driven machine learning approach, are often weak in accuracy and physical
601 rationality, arising from uncertainties in environmental forcings and inappropriate
602 parameterizations (Brodeau et al., 2017; Jiang et al., 2024a; Wang et al., 2024). To
603 improve the simulation of the global air-sea turbulent heat fluxes, this study presents
604 the BrTHF product, generated using a Bowen ratio-constrained NN technique with a
605 custom multiple-objective loss function, as well as observations from 197 globally
606 distributed buoys along with multi-source remote sensing and reanalysis inputs.

607 **3.5.1 Advantages**

608 The primary advantage of the BrTHF model lies in its accurate estimation of β ,
609 which shows the most pronounced improvement among all flux components. As a key
610 indicator of surface energy partitioning, β is widely used within the surface energy
611 balance framework to ensure physically consistent and reliable estimates of SHF and
612 LHF (Jo et al., 2004; Yang and Roderick, 2019; Yang et al., 2025). In addition, β serves
613 as an effective diagnostic variable in the studies of large-scale climate variability (e.g.,
614 El Niño) (Jo, 2002; Weisberg and Wang, 1997) and in investigations of how surface
615 energy constraints regulate the hydrological cycle (e.g., precipitation) (Cai and Lu,
616 2009; Wang et al., 2021). With its enhanced representation of β , the BrTHF product is
617 expected to provide more reliable support for such applications.

618 To achieve these improvements in β and flux, the BrTHF model was designed
619 differently from our previous study (Wang et al., 2024), which simultaneously predicted
620 SHF, LHF and β in the constructed RF model. Specifically, this study employed an NN
621 model constrained by the Bowen ratio to jointly estimate SHF and LHF. The new
622 approach avoided the issue of selection of β derived from either the calculated β [β_{cal}

623 equals predicted SHF (SHF_{pre}) divided by predicted LHF (LHF_{pre})] or the predicted β
624 (β_{pre}), as reported by Wang et al. (2024). Furthermore, the custom loss function in our
625 new approach provides a flexible approach to adjust the weights of SHF, LHF, and β ,
626 allowing the model to balance attention among these variables. As a result, the accuracy
627 of SHF, LHF, and β from our newly developed BrTHF model outperformed that of the
628 mainstream air-sea turbulent heat flux products and the physics-free NN models on both
629 global and regional scales. In contrast, the accuracy of SHF and LHF in the model
630 constructed by Wang et al. (2024) was marginally lower than that of the physics-free
631 RF model.

632 **3.5.2 Generalizability**

633 Based on Figure 2 and Table S8, we observe that the spatial coverage of
634 observations varies across different ocean regions: the Northern Hemisphere generally
635 has higher coverage than the Southern Hemisphere, with the Northern Pacific Ocean
636 exhibiting the highest coverage, while the Arctic Ocean shows the lowest. Comparing
637 spatial coverage with accuracy metrics reveals a more complex relationship between
638 model performance and data coverage. Specifically, the values of r tend to be lower in
639 regions with lower coverage — a pattern consistent across SHF, LHF, and β . However,
640 RMSE does not follow this trend. For SHF and β , RMSEs in the Northern Hemisphere
641 are generally higher than those in the Southern Hemisphere. Similarly, for LHF, RMSEs
642 are higher in the Northern Hemisphere except in the Indian Ocean, where the pattern
643 differs.

644 We applied a spatial 10-fold cross-validation, which provides a more generalized
645 assessment than traditional random cross-validation, to evaluate the BrTHF model.
646 However, it is important to acknowledge that the spatial distribution of the training
647 dataset is inherently imbalanced, with a heavy concentration of observations in the
648 Tropics and the Northern Hemisphere. In contrast, the Southern Hemisphere—
649 particularly the Southern Ocean—suffers from sparse or even missing observational
650 coverage. Given that the environmental conditions in these underrepresented or data-

651 sparse regions may differ significantly from those captured in the training dataset, the
652 selected input variables for the observations may lead to large uncertainty in the model's
653 performance in these areas. To further assess the model's ability to extrapolate to such
654 regions, we conducted an additional targeted cross-validation. Specifically, we
655 excluded stations from the Southern Ocean [i.e., Southern Ocean Flux Station (SOFS)
656 and Global Southern Ocean Station (GSOS)] from the training dataset and used them
657 solely for validation. Results presented in Tables S4 and S5 show that the BrTHF model
658 achieved the best performance in terms of LHF and β at the SOFS with lower RMSE
659 of 15.6 W/m² and 0.73 and higher values of r of 0.96 and 0.34, respectively, while its
660 SHF was slightly outperformed by ERA5 and the physics-free NN model. At the GSOS,
661 BrTHF yielded more accurate estimates for SHF and β with RMSEs of 6.38 W/m² and
662 0.74 and values of r of 0.95 and 0.16, respectively, compared to other products, while
663 its LHF was marginally less accurate than that of SeaFlux and the physics-free NN
664 model. Moreover, under both spatially-informed cross-validation and targeted cross-
665 validation, the model demonstrates comparable accuracy at the two sites, as shown in
666 Figures S4–S7. These findings suggest that BrTHF retains competitive accuracy of SHF,
667 LHF and β even in regions entirely excluded from training, reflecting promising
668 generalization. While these results are encouraging, it is important to note that the
669 validation remains limited to a small number of sites with available observations.
670 Therefore, the reported r values and RMSE reflect model performance in these specific
671 locations and do not necessarily guarantee similar accuracy in broader, unobserved
672 ocean regions. Consequently, the BrTHF product should be viewed as being optimized
673 within the geographical coverage of existing buoy networks. In remote regions far from
674 the observation-rich regions, such as the high-latitude Southern Ocean, the model
675 inputs may fall outside the range of values represented in the training data, leading to
676 increased uncertainty. Users should therefore exercise caution when interpreting the
677 global-scale performance, particularly in data-sparse basins where spatial sampling
678 limitations are most pronounced.

679 The generalization capability of the model can also affect the accuracy of
680 simulated long-term trends. In Figure 13, we present the spatial distributions of long-
681 term trends for SHF, LHF, and β simulated by the BrTHF product. Considering the
682 scarcity of training data in high-latitude oceans, the simulated long-term trends in these
683 regions may be associated with larger uncertainties. However, due to the lack of long-
684 term observations in high-latitude oceans, we cannot validate the simulated trends using
685 observational records as has been done in previous studies for mid- and low-latitude
686 regions (Weller et al., 2022; Tang et al., 2024). To address this, we examined the spatial
687 distribution of long-term trends from the other seven widely used products. Specifically,
688 in these high-latitude regions, the trends simulated by the BrTHF are largely consistent
689 with those of most other products—for example, SHF exhibits a pronounced increase
690 in the Kara Sea, Gulf Stream, Baffin Bay, Brazil Current, Sea of Okhotsk, and Sea of
691 Japan, with differences mainly in magnitude.

692 **3.5.3 Sensitivity and uncertainty**

693 Given the key role of the β constraint in the BrTHF model, it is also important to
694 assess the sensitivity of the estimated SHF and LHF to the imposed β range. A series of
695 sensitivity experiments with progressively relaxed β constraints indicate that the BrTHF
696 model exhibits weak sensitivity to the specific choice of the β range (see Table S9). The
697 resulting variations in RMSE are small relative to their mean values, while values of r
698 and BIAS remain largely unchanged across different β configurations. These results
699 suggest that the improved performance of BrTHF is not driven by a particular
700 predefined range of β , but instead reflects the robustness of the Bowen ratio-
701 constrained machine learning framework.

702 Beyond the sensitivity to the imposed β range, we further examined the robustness
703 of the BrTHF model to uncertainties in its environmental inputs, using one-at-a-time
704 perturbation experiments (Table S10). The results indicate that the model is generally
705 insensitive to perturbations (from ± 5 to $\pm 20\%$) in most auxiliary inputs, with changes
706 in SHF and LHF RMSE typically remaining within 1%. Noticeable sensitivities are

707 mainly associated with WS , $diff_Q$, and $diff_T$, which are physically expected given their
708 direct roles in controlling air–sea turbulent heat exchanges. Even under extreme
709 perturbations, the resulting variations in SHF and LHF remain within physically
710 reasonable ranges, suggesting that the BrTHF model does not rely excessively on
711 precise tuning of individual inputs. For β , most perturbations lead to minor changes;
712 however, perturbations of $diff_Q$ can induce large variability due to the ratio-based nature
713 of β , further highlighting the necessity of the β constraint.

714 To further address potential methodological inconsistencies between buoy-derived
715 fluxes and flux products, we conducted an additional baseline experiment. Specifically,
716 the COARE3.5 model was forced with the same subset of daily meteorological
717 variables used to drive the BrTHF model, thereby providing a reference under identical
718 forcing conditions. The results (Table S12) show that the COARE3.5-driven estimates
719 exhibit substantially larger errors for SHF and LHF, with RMSEs of approximately 10.3
720 W/m^2 and 34.4 W/m^2 , respectively, accompanied by systematic underestimations
721 relative to buoy observations. More importantly, physically unrealistic values of the β
722 emerge in the COARE3.5 estimates, leading to a degradation in β accuracy (RMSE =
723 6.35). These findings suggest that, even when driven by the same meteorological inputs,
724 traditional bulk formulations remain highly sensitive to forcing uncertainties. In
725 contrast, the BrTHF model demonstrates improved robustness by explicitly
726 incorporating physical constraints within the machine-learning framework.

727 Finally, we examined the potential impact of temporal aggregation, as the buoy-
728 derived daily fluxes used for model training and evaluation were calculated from daily-
729 averaged meteorological variables via COARE3.5, which may introduce biases due to
730 the nonlinearity of bulk flux formulations. By recalculating fluxes from high-frequency
731 buoy observations and aggregating to daily means, we found that although absolute
732 errors differ to some extent, the relative performance among different flux products and
733 models remains unchanged, with the BrTHF model still achieving the best overall
734 accuracy (Table S11). This confirms that our conclusions are robust to the choice of

735 daily flux calculation strategy.

736 **3.5.4 Limitations and recommendations**

737 Although the results demonstrated significant improvements in the accuracy and
738 physical consistency of SHF, LHF, and β estimates from the BrTHF model compared
739 to those from the physics-free NN models and the seven products, the BrTHF product
740 also has some limitations. First, extreme values of β may arise either from physically
741 plausible but rare ocean conditions or from numerical instability or poorly constrained
742 estimates associated with uncertainties in model-derived fluxes and input variables. In
743 practice, these two sources are difficult to distinguish. To ensure robust model training
744 and stable performance across the vast majority of β conditions (approximately the 1-
745 99% of the β distribution), we applied a conservative β constraint (-5 to 5) during
746 training. This compresses the standard deviation and constrains the extreme tails of the
747 predicted β , but substantially improves model stability and accuracy for the majority of
748 ocean conditions. Although this choice leads to a narrower dynamic range compared to
749 some other products, it ensures that the 1-99% of the β distribution is well-represented
750 for most practical applications. It should be noted that this strategy represents an interim
751 solution rather than a final one. With future improvements in the quality and
752 spatiotemporal representativeness of observational datasets, physically plausible
753 extremes (e.g., 0-1% and 99-100% of the β distribution) could be better constrained and
754 incorporated into model training, allowing expansion of the dynamic range of β .
755 Secondly, the estimated SHF and LHF values exhibited a narrower distribution
756 compared to the observations. This issue possibly stems from the uncertainty of the
757 BrTHF model that was constructed from the uneven distribution of SHF and LHF in
758 the observation datasets, which contain a low proportion of extreme samples, especially
759 negative LHF values. Moreover, due to the insufficient observations, validation and
760 modelling in high-latitude oceans, especially in the Southern Ocean, was limited. To
761 address these problems, more experiments are highly recommended to collect
762 observations covering more regions of oceans with better spatial and temporal

763 representativeness, which could further enhance the product.

764 The BrTHF model demonstrated the feasibility and potential of jointly estimating
765 multiple interrelated air-sea variables through a machine learning model that
766 incorporates appropriate physical constraints to account for their interrelations. In the
767 future, the predicted variables in the BrTHF model could be expanded to include
768 surface radiation, heat storage, and precipitation over the ocean, by integrating the
769 physical mechanisms of energy and water exchange. This would allow for the
770 collaborative optimization of estimates across all components of the air-sea energy and
771 water budgets, potentially contributing to achieving global closure of the air-sea energy
772 and water budgets.

773

774 **4. Data and code availability**

775 The daily 0.25° BrTHF product, consisting of SHF and LHF estimates from 1993
776 to 2017, can be freely accessed from the National Tibetan Plateau Data Center (TPDC)
777 [<https://doi.org/10.11888/Atmos.tpdc.302578>, Tang and Wang (2025)]. The code for
778 developing the product can be found on the GitHub platform
779 (<https://github.com/zhezhe1996/BrTHF>).

780

781 **5. Summary and Conclusion**

782 In this study, we generated a daily 0.25° air-sea turbulent heat fluxes product for
783 the period from 1993 to 2017 using our developed BrTHF model and multi-source
784 remote sensing and reanalysis data. A comprehensive validation was performed against
785 observations from 197 buoys and inter-comparisons were made with seven
786 representative gridded products. The key findings are as follows:

787 The BrTHF model demonstrated the most significant improvement in estimating
788 the β , while its performance in estimating SHF and LHF was generally comparable to
789 or slightly better than that of the physics-free NN models and the seven widely used
790 air-sea turbulent heat products (including the JOFURO3, IFREMER, SeaFlux, ERA5,

791 MERRA2, OAFflux and OHF products). Through the spatial ten-fold cross-validation
792 against the observations from the 197 buoys, the BrTHF model achieved RMSEs of
793 6.05 W/m^2 for SHF, 23.67 W/m^2 for LHF and 0.22 for β , and showed values of r of 0.93 ,
794 0.91 , and 0.25 for SHF, LHF, and β , respectively. Additionally, the BrTHF model
795 performed better in evaluations across six major ocean basins, with lower RMSEs and
796 higher values of r , in comparison to the physics-free NN models and the seven products.
797 Notably, the BrTHF model significantly improved the rationality of β estimates,
798 successfully eliminating the outliers observed in both the physics-free NN models and
799 the seven products. Furthermore, the global distributions for SHF, LHF, and β from the
800 BrTHF product closely matched those of the physically-based ERA5 and SeaFlux
801 products. The global mean annual estimates of SHF, LHF, and β from the BrTHF
802 product from 1993 to 2017 were 12.7 W/m^2 , 106.2 W/m^2 and 0.11 , respectively, all
803 within the ranges of the seven products. The BrTHF product exhibited similar intra-
804 annual cycles for SHF, LHF and β , with bimodal patterns featuring lower and higher
805 peaks in May-June-July-August and December-January, respectively, which was
806 consistent with the results of the seven state-of-the-art products. Additionally, the
807 BrTHF product exhibited significant increasing trends for global SHF and LHF, with
808 rates of $0.04 \text{ W/(m}^2 \text{ year)}$ and $0.33 \text{ W/(m}^2 \text{ year)}$, respectively, which were consistent to
809 most of the seven products. In contrast, the BrTHF product displayed weak growth in
810 β , with a trend approaching 0 , which was opposite to the results of the seven products
811 except for the MERRA2 product. The increasing (significant increasing) trends
812 dominated the oceans, with areas of 71.4% (26.2%) for SHF, 87.5% (38.9%) for LHF,
813 53% (12.2%) for β in the BrTHF product.

814 The BrTHF product shows significant advantages in the accuracy and rationality
815 of estimates for key parameters (SHF, LHF, and β) related to air-sea interaction and the
816 global energy and water budgets compared to the existing products. It holds great
817 potential for quantifying the global air-sea energy and water budgets, enhancing our
818 understanding of the air-sea interaction, and projecting climate change under global

819 warming.

820 **Author contribution**

821 YW and RT conceived the study and designed the experimental framework. YW
822 performed the experiment and prepared the initial manuscript draft. All authors
823 contributed to manuscript revision, and approved the final version of the manuscript.

824 **Competing interests**

825 The contact author has declared that none of the authors has any competing
826 interests.

827 **Acknowledgement**

828 We thank the flux datasets and learning datasets provided by the J-OFURO project,
829 IFREMER, ECMWF, NASA, WHOI, NOAA, CEMES and RSS. Moreover, we thank
830 the observations provided by the TAO/TRITON, PIRATA, RAMA, NDBC, TAC, UOP,
831 OOI, AOOS, KOREA, OCS, JKEO and IMO networks or organizations. This work is
832 supported by the National Natural Science Foundation of China [42271378,42071332],
833 and the Strategic Priority Research Program of the Chinese Academy of Sciences
834 (Grant No. XDB0740202).

835

836 **References**

- 837 Bentamy, A., Piollé, J. F., Grouazel, A., Danielson, R., Gulev, S., Paul, F., Azelmat, H.,
838 Mathieu, P. P., von Schuckmann, K., Sathyendranath, S., Evers-King, H., Esau, I.,
839 Johannessen, J. A., Clayson, C. A., Pinker, R. T., Grodsky, S. A., Bourassa, M.,
840 Smith, S. R., Haines, K., Valdivieso, M., Merchant, C. J., Chapron, B., Anderson,
841 A., Hollmann, R., and Josey, S. A.: Review and assessment of latent and sensible
842 heat flux accuracy over the global oceans, *Remote Sensing of Environment*, 201,
843 196-218, 10.1016/j.rse.2017.08.016, 2017.
- 844 Berry, D. I. and Kent, E. C.: Air-Sea fluxes from ICOADS: the construction of a new
845 gridded dataset with uncertainty estimates, *International Journal of Climatology*,
846 31, 987-1001, 10.1002/joc.2059, 2011.
- 847 Bourras, D.: Comparison of five satellite-derived latent heat flux products to moored
848 buoy data, *Journal of Climate*, 19, 6291-6313, 2006.
- 849 Bourras, D., Reverdin, G., Caniaux, G., and Belamari, S.: A Nonlinear Statistical Model

850 of Turbulent Air–Sea Fluxes, *Monthly Weather Review*, 135, 1077-1089,
851 10.1175/mwr3335.1, 2007.

852 Bourras, D., Cambra, R., Marié, L., Bouin, M. N., Baggio, L., Branger, H., Beghoura,
853 H., Reverdin, G., Dewitte, B., Paulmier, A., Maes, C., Ardhuin, F., Pairaud, I.,
854 Fraunié, P., Luneau, C., and Hauser, D.: Air - Sea Turbulent Fluxes From a
855 Wave - Following Platform During Six Experiments at Sea, *Journal of*
856 *Geophysical Research: Oceans*, 124, 4290-4321, 10.1029/2018jc014803, 2019.

857 Bourras, D., Weill, A., Caniaux, G., Eymard, L., Bourlès, B., Letourneur, S., Legain,
858 D., Key, E., Baudin, F., Pignatelli, B., Traullé, O., Bouhours, G., Sinardet, B., Barrié,
859 J., Vinson, J. P., Boutet, F., Berthod, C., and Cléménçon, A.: Turbulent air - sea
860 fluxes in the Gulf of Guinea during the AMMA experiment, *Journal of*
861 *Geophysical Research: Oceans*, 114, 10.1029/2008jc004951, 2009.

862 Brodeau, L., Barnier, B., Gulev, S. K., and Woods, C.: Climatologically Significant
863 Effects of Some Approximations in the Bulk Parameterizations of Turbulent Air–
864 Sea Fluxes, *Journal of Physical Oceanography*, 47, 5-28, 10.1175/jpo-d-16-0169.1,
865 2017.

866 Brunke, M. A.: Uncertainties in sea surface turbulent flux algorithms and data sets,
867 *Journal of Geophysical Research*, 107, 10.1029/2001jc000992, 2002.

868 Cai, L., Wang, B., Wang, W., and Feng, X.: The Impact of Air–Sea Flux
869 Parameterization Methods on Simulating Storm Surges and Ocean Surface
870 Currents, *Journal of Marine Science and Engineering*, 13, 10.3390/jmse13030541,
871 2025.

872 Cai, M. and Lu, J.: Stabilization of the Atmospheric Boundary Layer and the Muted
873 Global Hydrological Cycle Response to Global Warming, *Journal of*
874 *Hydrometeorology*, 10, 347-352, 10.1175/2008jhm1058.1, 2009.

875 Chen, X., Yao, Y., Zhao, S., Li, Y., Jia, K., Zhang, X., Shang, K., Xu, J., and Bei, X.:
876 Estimation of High-Resolution Global Monthly Ocean Latent Heat Flux from
877 MODIS SST Product and AMSR-E Data, *Advances in Meteorology*, 2020, 1-19,
878 10.1155/2020/8857618, 2020a.

879 Chen, X., Yao, Y., Li, Y., Zhang, Y., Jia, K., Zhang, X., Shang, K., Yang, J., Bei, X., and
880 Guo, X.: ANN-Based Estimation of Low-Latitude Monthly Ocean Latent Heat
881 Flux by Ensemble Satellite and Reanalysis Products, *Sensors (Basel)*, 20,
882 10.3390/s20174773, 2020b.

883 Clayson, C. and Brown, J.: NOAA Climate Data Record Ocean Surface Bundle (OSB)
884 Climate Data Record (CDR) of Ocean Heat Fluxes, Version 2, *Clim. Algorithm*
885 *Theor. Basis Doc. C-ATBD Asheville NC NOAA Natl. Cent. Environ. Inf. Doi*,
886 10, V59K4885, 2016.

887 Crespo, J., Posselt, D., and Asharaf, S.: CYGNSS Surface Heat Flux Product
888 Development, *Remote Sensing*, 11, 10.3390/rs11192294, 2019.

889 Cummins, D. P., Guemas, V., Cox, C. J., Gallagher, M. R., and Shupe, M. D.: Surface
890 Turbulent Fluxes From the MOSAiC Campaign Predicted by Machine Learning,
891 *Geophysical Research Letters*, 50, 10.1029/2023gl105698, 2023.

892 Cummins, D. P., Guemas, V., Blein, S., Brooks, I. M., Renfrew, I. A., Elvidge, A. D.,
893 and Prytherch, J.: Reducing Parametrization Errors for Polar Surface Turbulent
894 Fluxes Using Machine Learning, *Boundary-Layer Meteorology*, 190,
895 10.1007/s10546-023-00852-8, 2024.

896 Edson, J. B. a. J., Venkata and Weller, Robert A and Bigorre, Sebastien P and
897 Plueddemann, Albert J and Fairall, Christopher W and Miller, Scott D and Mahrt,
898 Larry and Vickers, Dean and Hersbach, Hans: On the Exchange of Momentum
899 over the Open Ocean, *Journal of Physical Oceanography*, 43, 1589-1610,
900 10.1175/jpo-d-12-0173.1, 2013.

901 Fasullo, J. T., Trenberth, K. E., and Balmaseda, M. A.: Earth's Energy Imbalance,
902 *Journal of Climate*, 27, 3129-3144, 10.1175/jcli-d-13-00294.1, 2014.

903 Fu, S., Huang, W., Luo, J., Yang, Z., Fu, H., Luo, Y., and Wang, B.: Deep Learning -
904 Based Sea Surface Roughness Parameterization Scheme Improves Sea Surface
905 Wind Forecast, *Geophysical Research Letters*, 50, 10.1029/2023gl106580, 2023.

906 Gentemann, C. L., Clayson, C. A., Brown, S., Lee, T., Parfitt, R., Farrar, J. T., Bourassa,
907 M., Minnett, P. J., Seo, H., Gille, S. T., and Zlotnicki, V.: FluxSat: Measuring the
908 Ocean-Atmosphere Turbulent Exchange of Heat and Moisture from Space,
909 *Remote Sensing*, 12, 10.3390/rs12111796, 2020.

910 Grist, J. P., Josey, S. A., Zika, J. D., Evans, D. G., and Skliris, N.: Assessing recent air-
911 sea freshwater flux changes using a surface temperature-salinity space framework,
912 *Journal of Geophysical Research: Oceans*, 121, 8787-8806,
913 10.1002/2016jc012091, 2016.

914 Jiang, Y., Li, Y., Lu, Y., Wu, T., and Gao, Z.: Evaluating modifications to air-sea
915 momentum flux parameterizations under light wind conditions in CAM6, *Climate
916 Dynamics*, 62, 9687-9701, 10.1007/s00382-024-07415-8, 2024a.

917 Jiang, Y., Li, Y., Lu, Y., Wu, T., Zhang, J., and Gao, Z.: Evaluating nine different air-sea
918 flux algorithms coupled with CAM6, *Atmospheric Research*,
919 10.1016/j.atmosres.2024.107486, 2024b.

920 Jo, Y.-H.: Calculation of the Bowen ratio in the tropical Pacific using sea surface
921 temperature data, *Journal of Geophysical Research*, 107, 10.1029/2001jc001150,
922 2002.

923 Jo, Y.-H., Yan, X.-H., Pan, J., Liu, W. T., and He, M.-X.: Sensible and latent heat flux
924 in the tropical Pacific from satellite multi-sensor data, *Remote Sensing of
925 Environment*, 90, 166-177, 10.1016/j.rse.2003.12.003, 2004.

926 Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L.:
927 Physics-informed machine learning, *Nature Reviews Physics*, 3, 422-440, 2021.

928 Kashinath, K., Mustafa, M., Albert, A., Wu, J., Jiang, C., Esmailzadeh, S.,
929 Azizzadenesheli, K., Wang, R., Chattopadhyay, A., and Singh, A.: Physics-
930 informed machine learning: case studies for weather and climate modelling,
931 *Philosophical Transactions of the Royal Society A*, 379, 20200093, 2021.

932 Kudryavtsev, V., Chapron, B., and Makin, V.: Impact of wind waves on the air - sea
933 fluxes: A coupled model, *Journal of Geophysical Research: Oceans*, 119, 1217-

934 1236, 10.1002/2013jc009412, 2014.

935 Liman, J., Schröder, M., Fennig, K., Andersson, A., and Hollmann, R.: Uncertainty
936 characterization of HOAPS 3.3 latent heat-flux-related parameters, *Atmospheric*
937 *Measurement Techniques*, 11, 1793-1815, 10.5194/amt-11-1793-2018, 2018.

938 Loeb, N. G., Johnson, G. C., Thorsen, T. J., Lyman, J. M., Rose, F. G., and Kato, S.:
939 Satellite and Ocean Data Reveal Marked Increase in Earth's Heating Rate,
940 *Geophysical Research Letters*, 48, 10.1029/2021GL093047, 2021.

941 Monin, A. S. and Obukhov, A. M.: Basic laws of turbulent mixing in the surface layer
942 of the atmosphere, *Contrib. Geophys. Inst. Acad. Sci. USSR*, 151, e187, 1954.

943 Myslenkov, S., Shestakova, A., and Chechin, D.: The impact of sea waves on turbulent
944 heat fluxes in the Barents Sea according to numerical modeling, *Atmospheric*
945 *Chemistry and Physics*, 21, 5575-5595, 10.5194/acp-21-5575-2021, 2021.

946 Nelson, J. A., Walther, S., Gans, F., Kraft, B., Weber, U., Novick, K., Buchmann, N.,
947 Migliavacca, M., Wohlfahrt, G., Šigut, L., Ibrom, A., Papale, D., Göckede, M.,
948 Duveiller, G., Knohl, A., Hörtnagl, L., Scott, R. L., Zhang, W., Hamdi, Z. M.,
949 Reichstein, M., Aranda-Barranco, S., Ardö, J., Op de Beeck, M., Billesbach, D.,
950 Bowling, D., Bracho, R., Brümmer, C., Camps-Valls, G., Chen, S., Cleverly, J. R.,
951 Desai, A., Dong, G., El-Madany, T. S., Euskirchen, E. S., Feigenwinter, I.,
952 Galvagno, M., Gerosa, G. A., Gielen, B., Goded, I., Goslee, S., Gough, C. M.,
953 Heinesch, B., Ichii, K., Jackowicz-Korczynski, M. A., Klosterhalfen, A., Knox, S.,
954 Kobayashi, H., Kohonen, K.-M., Korkiakoski, M., Mammarella, I., Gharun, M.,
955 Marzuoli, R., Matamala, R., Metzger, S., Montagnani, L., Nicolini, G., O'Halloran,
956 T., Ourcival, J.-M., Peichl, M., Pendall, E., Ruiz Reverter, B., Roland, M.,
957 Sabbatini, S., Sachs, T., Schmidt, M., Schwalm, C. R., Shekhar, A., Silberstein, R.,
958 Silveira, M. L., Spano, D., Tagesson, T., Tramontana, G., Trotta, C., Turco, F.,
959 Vesala, T., Vincke, C., Vitale, D., Vivoni, E. R., Wang, Y., Woodgate, W., Yopez,
960 E. A., Zhang, J., Zona, D., and Jung, M.: X-BASE: the first terrestrial carbon and
961 water flux products from an extended data-driven scaling framework,
962 *FLUXCOM-X, Biogeosciences*, 21, 5079-5115, 10.5194/bg-21-5079-2024, 2024.

963 O, S. and Orth, R.: Global soil moisture data derived through machine learning trained
964 with in-situ measurements, *Sci Data*, 8, 170, 10.1038/s41597-021-00964-1, 2021.

965 Peng, Z., Tang, R., Jiang, Y., Liu, M., and Li, Z.-L.: Global estimates of 500 m daily
966 aerodynamic roughness length from MODIS data, *ISPRS Journal of*
967 *Photogrammetry and Remote Sensing*, 183, 336-351,
968 10.1016/j.isprsjprs.2021.11.015, 2022.

969 Robertson, F. R., Roberts, J. B., Bosilovich, M. G., Bentamy, A., Clayson, C. A., Fennig,
970 K., Schröder, M., Tomita, H., Compo, G. P., Gutenstein, M., Hersbach, H.,
971 Kobayashi, C., Ricciardulli, L., Sardeshmukh, P., and Slivinski, L. C.:
972 Uncertainties in Ocean Latent Heat Flux Variations over Recent Decades in
973 Satellite-Based Estimates and Reduced Observation Reanalyses, *Journal of*
974 *Climate*, 33, 8415-8437, 10.1175/jcli-d-19-0954.1, 2020.

975 Shang, K., Yao, Y., Di, Z., Jia, K., Zhang, X., Fisher, J. B., Chen, J., Guo, X., Yang, J.,

976 Yu, R., Xie, Z., Liu, L., Ning, J., and Zhang, L.: Coupling physical constraints with
977 machine learning for satellite-derived evapotranspiration of the Tibetan Plateau,
978 *Remote Sensing of Environment*, 289, 10.1016/j.rse.2023.113519, 2023.

979 Shie, C.-L., Chiu, L. S., Adler, R., Nelkin, E., Lin, I. I., Xie, P., Wang, F.-C.,
980 Chokngamwong, R., Olson, W., and Chu, D. A.: A note on reviving the Goddard
981 Satellite-based Surface Turbulent Fluxes (GSSTF) dataset, *Advances in*
982 *Atmospheric Sciences*, 26, 1071-1080, 10.1007/s00376-009-8138-z, 2009.

983 Song, X.: The Importance of Relative Wind Speed in Estimating Air–Sea Turbulent
984 Heat Fluxes in Bulk Formulas: Examples in the Bohai Sea, *Journal of Atmospheric*
985 *and Oceanic Technology*, 37, 589-603, 10.1175/jtech-d-19-0091.1, 2020.

986 Song, X.: The Importance of Including Sea Surface Current when Estimating Air–Sea
987 Turbulent Heat Fluxes and Wind Stress in the Gulf Stream Region, *Journal of*
988 *Atmospheric and Oceanic Technology*, 38, 119-138, 10.1175/jtech-d-20-0094.1,
989 2021.

990 Song, X., Xie, X., Yan, Y., and Xie, S.-P.: Observed sub-daily variations in air–sea
991 turbulent heat fluxes under different marine atmospheric boundary layer stability
992 conditions in the Gulf Stream, *Monthly Weather Review*, 10.1175/mwr-d-24-
993 0003.1, 2024.

994 TANG, R. and WANG, Y.: Global dataset of air-sea turbulent heat fluxes (sensible heat
995 flux and latent heat flux) (1993–2017), National Tibetan Plateau Data Center
996 [dataset], <https://dx.doi.org/10.11888/Atmos.tpdc.302578>, 2025.

997 Tang, R., Wang, Y., Jiang, Y., Liu, M., Peng, Z., Hu, Y., Huang, L., and Li, Z.-L.: A
998 review of global products of air-sea turbulent heat flux: accuracy, mean, variability,
999 and trend, *Earth-Science Reviews*, 249, 10.1016/j.earscirev.2023.104662, 2024.

1000 Tomita, H., Hihara, T., Kako, S. i., Kubota, M., and Kutsuwada, K.: An introduction to
1001 J-OFURO3, a third-generation Japanese ocean flux data set using remote-sensing
1002 observations, *Journal of Oceanography*, 75, 171-194, 10.1007/s10872-018-0493-
1003 x, 2018.

1004 van der Westhuizen, S., Heuvelink, G. B., and Hofmeyr, D. P.: Multivariate random
1005 forest for digital soil mapping, *Geoderma*, 431, 116365, 2023.

1006 Wang, J., Tang, R., Jiang, Y., Liu, M., and Li, Z.-L.: A practical method for angular
1007 normalization of global MODIS land surface temperature over vegetated surfaces,
1008 *ISPRS Journal of Photogrammetry and Remote Sensing*, 199, 289-304,
1009 10.1016/j.isprsjprs.2023.04.015, 2023.

1010 Wang, W., Chakraborty, T. C., Xiao, W., and Lee, X.: Ocean surface energy balance
1011 allows a constraint on the sensitivity of precipitation to global warming, *Nat*
1012 *Commun*, 12, 2115, 10.1038/s41467-021-22406-7, 2021.

1013 Wang, Y., Tang, R., Huang, L., Liu, M., Jiang, Y., and Li, Z.-L.: A Bowen ratio-informed
1014 method for coordinating the estimates of air–sea turbulent heat fluxes,
1015 *Environmental Research Letters*, 19, 10.1088/1748-9326/ad9341, 2024.

1016 Weisberg, R. H. and Wang, C.: Slow variability in the equatorial west-central Pacific in
1017 relation to ENSO, *Journal of Climate*, 10, 1998-2017, 1997.

1018 Weller, R. A., Lukas, R., Potemra, J., Plueddemann, A. J., Fairall, C., and Bigorre, S.:
1019 Ocean Reference Stations: Long-Term, Open-Ocean Observations of Surface
1020 Meteorology and Air–Sea Fluxes Are Essential Benchmarks, *Bulletin of the*
1021 *American Meteorological Society*, 103, E1968-E1990, 10.1175/bams-d-21-0084.1,
1022 2022.

1023 Wild, M., Folini, D., Hakuba, M. Z., Schär, C., Seneviratne, S. I., Kato, S., Rutan, D.,
1024 Ammann, C., Wood, E. F., and König-Langlo, G.: The energy balance over land
1025 and oceans: an assessment based on direct observations and CMIP5 climate
1026 models, *Climate Dynamics*, 44, 3393-3429, 10.1007/s00382-014-2430-z, 2014.

1027 Yan, Y., Song, X., Wang, G., and Li, X.: Tropical Cool-Skin and Warm-Layer Effects
1028 and Their Impact on Surface Heat Fluxes, *Journal of Physical Oceanography*, 54,
1029 45-62, 10.1175/jpo-d-23-0103.1, 2024.

1030 Yang, Y. and Roderick, M. L.: Radiation, surface temperature and evaporation over wet
1031 surfaces, *Quarterly Journal of the Royal Meteorological Society*, 145, 1118-1129,
1032 10.1002/qj.3481, 2019.

1033 Yang, Y., Sun, H., Wang, J., Zhang, W., Zhao, G., Wang, W., Cheng, L., Chen, L., Qin,
1034 H., and Cai, Z.: Global ocean surface heat fluxes revisited: A new dataset from
1035 maximum entropy production framework with heat storage and Bowen ratio
1036 optimizations, *Earth System Science Data Discussions*, 2024, 1-44, 2024.

1037 Yang, Y., Sun, H., Wang, J., Zhang, W., Zhao, G., Wang, W., Cheng, L., Chen, L., Qin,
1038 H., and Cai, Z.: Global ocean surface heat fluxes derived from the maximum
1039 entropy production framework accounting for ocean heat storage and Bowen ratio
1040 adjustments, *Earth System Science Data*, 17, 1191-1216, 10.5194/essd-17-1191-
1041 2025, 2025.

1042 Yu, L.: Global Air–Sea Fluxes of Heat, Fresh Water, and Momentum: Energy Budget
1043 Closure and Unanswered Questions, *Annual Review of Marine Science*, 11, 227-
1044 248, 10.1146/annurev-marine-010816-060704, 2019.

1045 Yu, L. and Weller, R. A.: Objectively Analyzed Air–Sea Heat Fluxes for the Global Ice-
1046 Free Oceans (1981–2005), *Bulletin of the American Meteorological Society*, 88,
1047 527-540, 10.1175/bams-88-4-527, 2007.

1048 Zhang, R., Guo, Weihao, Wang, Xin, and Wang, Chunzai: Ambiguous Variations in
1049 Tropical Latent Heat Flux since the Years around 1998, *Journal of Climate*, 36,
1050 3403–3415, 2023.

1051 Zhao, W. L., Gentine, P., Reichstein, M., Zhang, Y., Zhou, S., Wen, Y., Lin, C., Li, X.,
1052 and Qiu, G. Y.: Physics - Constrained Machine Learning of Evapotranspiration,
1053 *Geophysical Research Letters*, 46, 14496-14507, 10.1029/2019gl085291, 2019.

1054 Zhou, S., Shi, R., Yu, H., Zhang, X., Dai, J., Huang, X., and Xu, F.: A Physical -
1055 Informed Neural Network for Improving Air - Sea Turbulent Heat Flux
1056 Parameterization, *Journal of Geophysical Research: Atmospheres*, 129,
1057 10.1029/2023jd040603, 2024.

1058 Zhou, X., Ray, P., Barrett, B. S., and Hsu, P.-C.: Understanding the bias in surface latent
1059 and sensible heat fluxes in contemporary AGCMs over tropical oceans, *Climate*

1060 Dynamics, 55, 2957-2978, 10.1007/s00382-020-05431-y, 2020.
1061