



OPEN

DATA DESCRIPTOR

# A global long-term ocean surface daily/0.05° net radiation product from 1983–2020

Hui Liang<sup>1,2</sup> , Bo Jiang<sup>1,2</sup> , Shunlin Liang<sup>3</sup> , Jianghai Peng<sup>1,2</sup>, Shaopeng Li<sup>1,2</sup>, Jiakun Han<sup>1,2</sup>, Xiuwan Yin<sup>1,2</sup>, Jie Cheng<sup>1,2</sup>, Kun Jia<sup>1,2</sup>, Qiang Liu<sup>4</sup>, Yunjun Yao<sup>1,2</sup>, Xiang Zhao<sup>1,2</sup> & Xiaotong Zhang<sup>1,2</sup>

The all-wave net radiation ( $R_n$ ) on the ocean surface characterizes the available radiative energy balance and is important to understand the Earth's climate system. Considering the shortcomings of available ocean surface  $R_n$  datasets (e.g., coarse spatial resolutions, discrepancy in accuracy, inconsistency, and short duration), a new long-term global daily  $R_n$  product at a spatial resolution of 0.05° from 1983 to 2020, as part of the Global High Resolution Ocean Surface Energy (GHOSE) products suite, was generated in this study by fusing several existing datasets including satellite and reanalysis products based on the comprehensive *in situ* measurements from 68 globally distributed moored buoy sites. Evaluation against *in-situ* measurements shows the root mean square difference, mean bias error and correlation coefficient squared of 23.56 Wm<sup>-2</sup>, 0.88 Wm<sup>-2</sup> and 0.878. The global average ocean surface  $R_n$  over 1983–2020 is estimated to be 119.71 ± 2.78 Wm<sup>-2</sup> with a significant increasing rate of 0.16 Wm<sup>-2</sup> per year. GHOSE  $R_n$  product can be valuable for oceanic and climatic studies.

## Background & Summary

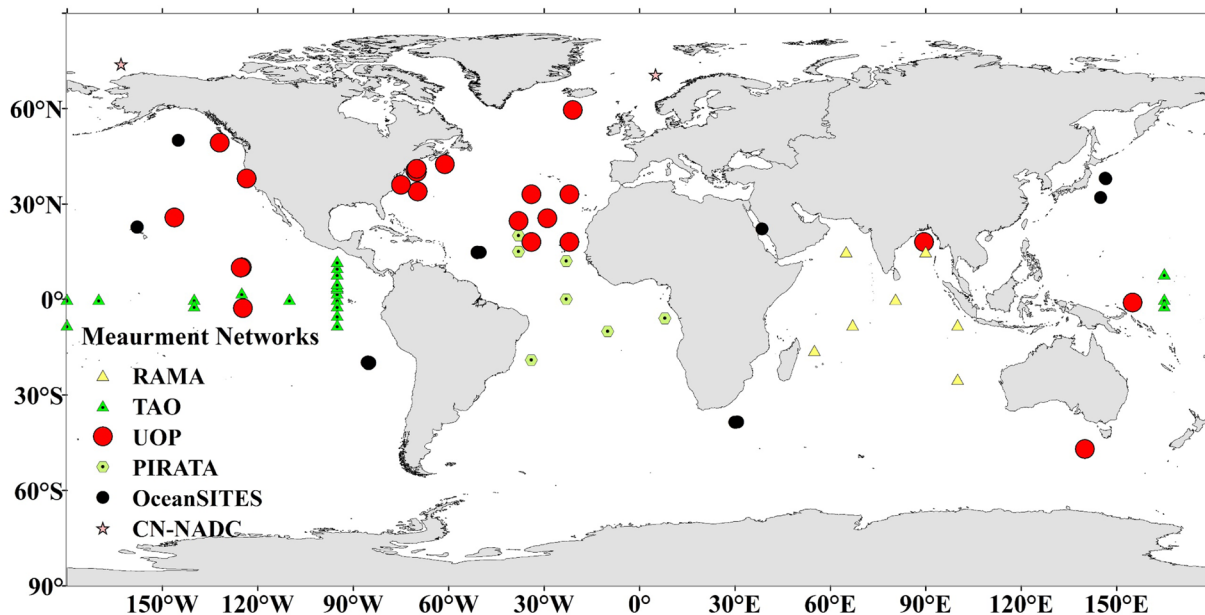
The all-wave net radiation ( $R_n$ ) at the ocean surface, which characterizes the available radiative energy balance on the ocean surface, is the difference between radiation from the sun and the atmosphere, and that emitted and reflected from the ocean surface<sup>1</sup>. The ocean surface  $R_n$  is mathematically represented as the sum of the net shortwave radiation ( $R_{ns}$ ) and the net longwave radiation ( $R_{nl}$ ) on the ocean surface, as shown in Eq. (1):

$$\begin{aligned} R_n &= R_{ns} + R_{nl} \\ R_{ns} &= R_{si} - R_{so} \\ R_{nl} &= R_{li} - R_{lo} \end{aligned} \quad (1)$$

where  $R_{si}$  is the incoming shortwave radiation (negative upward, Wm<sup>-2</sup>),  $R_{so}$  is the reflected outgoing shortwave radiation (Wm<sup>-2</sup>),  $R_{li}$  is the incoming longwave radiation (Wm<sup>-2</sup>), and  $R_{lo}$  is the outgoing longwave radiation (Wm<sup>-2</sup>).

The ocean is the largest body of heat storage in the Earth's climate system because of its large specific heat capacity in seawater and high occupancy of the Earth's surface area (71%). Ocean heat flux, which is composed of the ocean surface radiative heat flux ( $R_n$ ) and turbulent heat flux (latent heat flux and sensible heat flux)<sup>2</sup>, can regulate the heat balance in the Earth's system through frequent air-sea energy exchange and by transferring the heat into the underlying body of water and the overlying air<sup>3</sup>. The ocean heat flux is closely related to the Earth's Energy Imbalance (EEI)<sup>4</sup>, which is represented by the difference between the incident solar radiation and the outgoing longwave radiation of the Earth, and is thought to characterize the planet's climate and global warming caused by human activities<sup>5</sup>. Hence, as it is a major component of ocean heat flux, variations in the ocean surface  $R_n$  can directly affect it and, in turn, influence atmospheric and oceanic circulations, and even climate change at

<sup>1</sup>The State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and Institute of Remote Sensing and Digital Earth of Chinese Academy of Sciences, Beijing, 100875, China. <sup>2</sup>China and Beijing Engineering Research Center for Global Land Remote Sensing Products, Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing, 100875, China. <sup>3</sup>Department of Geographical Sciences, University of Maryland, College Park, MD, 20742, USA. <sup>4</sup>College of Global Change and Earth System Science, Beijing Normal University, Beijing, 100875, China. <sup>5</sup>e-mail: [bojiang@bnu.edu.cn](mailto:bojiang@bnu.edu.cn)



**Fig. 1** The distribution of the 68 moored buoys from six observation networks.

different spatio-temporal scales<sup>1,6,7</sup>. In addition, the ocean surface  $R_n$  plays an important role in the oceanic and climate-related research as an essential input to most models<sup>8</sup>. Therefore, an accurate estimation of the ocean surface  $R_n$  is important for various applications and understanding the state of the Earth's climate system.

However,  $R_n$  or the four radiative components (see Eq. (1)) are not routinely measured at the ocean surface, and the ocean surface  $R_n$  is usually estimated by several parameterizations or empirical formulas, or is obtained from various products by summing their radiative components. And almost all available products have coarse spatial resolutions, only with the finest spatial resolution of  $0.25^\circ$ . Moreover, some products cover only the ice-free oceans (e.g., the Japanese Ocean Flux Data Sets with Use of Remote Sensing Observations, version3 <J-OFURO3><sup>2</sup>, the Objectively Analyzed Air-Sea Flux <OAFflux><sup>9,10</sup>, and TropFlux<sup>11</sup>), and the uncertainty in the  $R_n$  over the coastal seas is large for most products owing to their coarse spatial resolutions<sup>12</sup>. Furthermore, other problems in these products, such as the discrepancy in accuracy, inconsistency<sup>13</sup>, and short duration, also limit the wide use of the ocean surface  $R_n$  from them. Therefore, a reliable, long-term time series of the ocean surface  $R_n$  dataset with a high accuracy and high spatio-temporal resolution is needed.

Data fusion, as an effective means of improving parameter estimation, has been applied widely, and works well as long as a sufficient number of datasets are available. It can combine estimates from multiple algorithms or high-level products to obtain new and better estimates, which are highly accurate and have high spatio-temporal resolution, through statistical methods<sup>14,15</sup>, including classic methods (e.g., regression, Bayesian Model Averaging (BMA)<sup>16,17</sup>, Simple Taylor Skill (STS)<sup>18</sup>, Empirical Orthogonal Function (EOF)<sup>19</sup>), and various machine learning methods (e.g., Multi-resolution Tree (MRT)<sup>20</sup>, Supporting Vector Machine (SVM)<sup>6</sup>, and artificial neural networks (ANN)<sup>21</sup>). A large number of datasets with various parameters has been successfully generated by data fusion, such as those for the land surface  $R_{nl}$ <sup>22</sup>, terrestrial latent heat flux<sup>23</sup>, the surface radiation budget (SRB) over the Tibetan Plateau<sup>24</sup>, land surface broadband emissivity<sup>25</sup>, Leaf Area Index (LAI)<sup>19</sup>, and surface broadband albedo<sup>26</sup>. Overall, data fusion is a very useful technique for improving product quality regardless of the parameter at hand. Therefore, by considering a plenty of available ocean surface  $R_n$  datasets, data fusion was applied for generating a new better one.

In this study, after exploring four commonly used data fusion methods including stepwise regression (SR), STS, BMA and back-propagation neural network (BP) thoroughly, a new daily ocean surface  $R_n$  dataset at a spatial resolution of  $0.05^\circ$  from 1983 to 2020, which was included in the Global High resolution Ocean Surface Energy (GHOSE) data suite, was generated with the best forming SR method. Afterwards, GHOSE daily  $R_n$  was evaluated comprehensively.

## Methods

**Calculation of the *in situ* ocean surface  $R_n$ .** In this study, 68 moored buoy sites from six observation networks were collected from 1988 to 2019. The spatial distribution of the buoy sites is shown in Fig. 1. They were located in global oceans, and were mainly distributed over seas at mid-to-low latitudes, especially near the Equator. However, few buoys were located in seas at high latitudes ( $>50^\circ$ ), especially in the Southern Hemisphere.

Table 1 lists detailed information on the six observation networks to which the 68 moored buoy sites belonged, including the Pilot Research Moored Array in the Tropical Atlantic (PIRATA), the Tropical Atmosphere Ocean (TAO), the Research Moored Array for African-Asian-Australian Monsoon (RAMA), the Ocean Sustained Interdisciplinary Timeseries Environment Observation System (OceanSITES), the Upper Ocean Processes Group (UOP), and the Chinese National Arctic and Antarctic Center (CN-NADC). Among them, PIRATA,

TAO, and RAMA are the major components of the Global Tropical Moored Buoy Array (GT MBA) Program, whose target of measurements are the tropical seas, including the Atlantic<sup>27</sup>, Pacific<sup>28</sup>, and Indian Oceans<sup>29</sup>. Their measurements are highly accurate owing to strict quality control procedures<sup>30</sup>. OceanSITES is a worldwide system for collecting and providing data at fixed locations in the open ocean<sup>31</sup>. The UOP network offers data on the upper ocean and at the air–sea interface in the north Atlantic. Measurements of three parameters ( $R_{si}$ ,  $R_{li}$ , and the sea surface temperature <SST>) provided by 66 sites from the above five networks were used to calculate  $R_n$ . Specifically at the moored buoys, two radiative measurements were taken directly using the Eppley Laboratory's precision spectral pyranometer (PSP) and the precision infrared radiometer (PIR)<sup>32</sup>. The SST was the adjusted sea bulk temperature measured 2 m under the sea surface using Sea Bird Electronics (SBE37/39), and adding  $-0.17\text{ K}$  to the result when the wind speed exceeded  $4\text{ ms}^{-1}$ <sup>33–35</sup>. Unlike the other 66 moored buoy sites, the two moored sites from CN-NADC, located in the North Polar seas, could measure four radiative components, but their periods of observation were short (five days in 1999 at CNNADC\_NP site, and 158 days from 2012 to 2014 at CNNADC\_NW\_sea site). The measurements provided by the six networks have been widely used, such as on research of climate variation<sup>36,37</sup>, heat budget<sup>38,39</sup>, and air–sea radiation fluxes<sup>40</sup>. Even still, the high quality measurements identified by their publishers<sup>41</sup> were manually checked, and the final selected measurements were used. All the selected *in situ* measurements were transferred into the Coordinated Universal Time (UTC) format, and the daily means were then calculated as long as all measurements were available for a given day.

As mentioned above, neither  $R_n$  nor the four radiative components (see Eq. (1)) are routinely measured, but the downward radiation ( $R_{si}$  and  $R_{li}$ ) were measured instead at almost all moored buoys. Hence, the *in situ* ocean surface  $R_n$  at the moored buoys needed to be first calculated based on the moored measurements and ancillary information ( $R_{si}$ ,  $R_{li}$ , and SST) in Eq. (2), except at the two sites from CN-NADC:

$$\begin{aligned} R_n &= R_{ns} + R_{nl} \\ R_{ns} &= R_{si} \times (1 - \alpha_{ocean}) \\ R_{nl} &= \varepsilon_{ocean} \times (R_{li} - \delta \times SST^4) \end{aligned} \quad (2)$$

where  $\delta$  is the Stefan-Boltzmann constant ( $5.6697 \times 10^{-8} [W/(m^2 \cdot K^4)]$ ),  $\alpha_{ocean}$  is the ocean surface broadband albedo,  $\varepsilon_{ocean}$  is the ocean surface broadband emissivity, and  $\alpha_{ocean}$  and  $\varepsilon_{ocean}$  were obtained from the GHOSE Ocean Water Albedo (OWA) and the broadband emissivity (BBE) products, respectively. The GHOSE OWA was developed based on satellite reflectance data via the Moderate-Resolution Imaging Spectroradiometer (MODIS) and ancillary data from MERRA2, obtained by Feng *et al.*<sup>42</sup> and Wang *et al.*<sup>43</sup> by using a three-component reflectance model at a resolution of 10 km from 1982 to 2019. It performed well after being validated with other parameterization methods. The GHOSE BBE is the first global ocean surface emissivity product with a resolution of 10 km from 1981 to 2019, and was generated from the MERRA2 reanalysis product developed by Cheng *et al.*<sup>44</sup>. It considers the effects of wind speed and sea foam. The authors claimed that uncertainty in the daily  $\varepsilon_{ocean}$  was less than 0.003 under wind-free conditions<sup>45</sup>.  $\alpha_{ocean}$  and  $\varepsilon_{ocean}$  were extracted directly according to the locations of the 66 moored buoys. According to Peng *et al.*<sup>12</sup>, the uncertainty in the calculated *in situ* daily ocean surface  $R_n$  could be tolerated even given uncertainties in the other parameters.

In summary, 69,923 samples of the *in situ* ocean surface  $R_n$  at the daily scale were calculated and used in this study. Specifically, 70% samples from 1988 to 2016 at each moored buoy were randomly selected for the fusion model development, and the other 30% samples were used for independent validation. Moreover, all samples in 2017–2019 (No. of samples = 4,734) were used for evaluation.

**Products for fusion.** The products that can provide the ocean surface  $R_n$  can be roughly divided into four categories: remotely sensed products, reanalysis products, reconstructed products, and ship-based products. Remotely sensed products are usually based on satellite observations, such as the Clouds and Earth's Radiant Energy Systems synoptic Edition4A<sup>46</sup> (CERES SYN1deg\_Ed4A, CERES4 hereinafter), the International Satellite Cloud Climatology Project radiative flux D-series<sup>29</sup> and H-series<sup>47,48</sup> (ISCCP-FD and ISCCP-FH), and Global Energy and Water Cycle Experiment - Surface Radiation Budget<sup>49,50</sup> (GEWEX-SRB); reanalysis products are derived from a data assimilation system by merging the available observations with model simulations<sup>51</sup>, such as the fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric Re-Analysis<sup>52</sup> (ERA5) and its previous version ERA-Interim<sup>7</sup>, the Japanese 55-year Reanalysis<sup>53</sup> (JRA-55), the Modern-Era Retrospective analysis for Research and Applications Version2<sup>54</sup> (MERRA2), and the National Centers for Environmental Prediction Department Reanalysis2<sup>55</sup> (NCEP R2). These two kinds of products have mostly global coverage with a relative coarse spatial resolution. Although the reanalysis products cover more years than remotely sensed products, the accuracy of the latter in terms of the radiative components is generally superior to that of most reanalysis products for land<sup>56</sup> and ocean surfaces<sup>57</sup>. Reconstructed products are generated by combining different kinds of products (e.g., remotely sensed and reanalysis products) through interpolation or other methods<sup>11</sup>, such as the aforementioned TropFlux, OAFlux, and J-OFURO3, and their performance is majorly determined by the original products. Ship-based products are calculated by means of parameterizations by using *in situ* measurements and then gridded by interpolation<sup>9</sup>, such as NOCS Surface Flux Dataset v2.0<sup>9</sup> (NOCS V2.0), hence, their accuracies heavily depend on the quality of observations, the performance of the parameterization methods, and the re-gridding methods.

Take characteristics of all products into account, values of the ocean surface  $R_n$  from 11 products were applied in this study. Specifically, six products covering 1988–2016 were used for fusion including two remotely sensed products (CERES4 and ISCCP-FH) and four reanalysis products (JRA-55, MERRA2, NCEP R2, and ERA5), and five products were used for comparison including one remotely sensed product (GEWEX-SRB), three reconstructed products (TropFlux, OAFlux, and J-OFURO3), and one ship-based product (NOCS V2.0).

Network	No. of sites	Variables measuring	Observation frequency	Site URL
OceanSITES	8	$R_{sp}$ , $R_{fb}$ , SST	1 hour	<a href="http://www.oceansites.org/">http://www.oceansites.org/</a> (last accessed: 11 Feb. 2021)
TAO	21	$R_{sp}$ , $R_{fb}$ , SST	10 min	<a href="https://www.pmel.noaa.gov/gtmba/">https://www.pmel.noaa.gov/gtmba/</a> (last accessed: 21 Apr. 2022)
RAMA	7	$R_{sp}$ , $R_{fb}$ , SST	10 min	<a href="https://www.pmel.noaa.gov/gtmba/">https://www.pmel.noaa.gov/gtmba/</a> (last accessed: 21 Apr. 2022)
PIRATA	7	$R_{sp}$ , $R_{fb}$ , SST	10 min	<a href="https://www.pmel.noaa.gov/gtmba/">https://www.pmel.noaa.gov/gtmba/</a> (last accessed: 21 Feb. 2022)
UOP	23	$R_{sp}$ , $R_{fb}$ , SST	1 hour	<a href="http://uop.whoi.edu/index.html">http://uop.whoi.edu/index.html</a> (last accessed: 11 Feb. 2021)
CN-NADC	2	$R_{sp}$ , $R_{fb}$ , $R_{so}$ , $R_{lo}$	1 hour	<a href="http://www.chinare.org.cn">http://www.chinare.org.cn</a> (last accessed: 11 Feb. 2021)

**Table 1.** The information of global six observing networks.

Product	Spatial Resolution	Temporal Resolution	Period	Coverage	URL
CERES4	1°	1 hourly	2000-	Global	<a href="https://ceres.larc.nasa.gov">https://ceres.larc.nasa.gov</a> (last accessed: 18 Jun. 2021)
ISCCP-FH	1°	3 hourly	1983–2017	Global	<a href="https://isccp.giss.nasa.gov/">https://isccp.giss.nasa.gov/</a> (last accessed: 18 Jun. 2021)
GEWEX-SRB	1°	3 hourly	1983–2007	Global	<a href="http://www.gewex.org/">http://www.gewex.org/</a> (last accessed: 18 Jun. 2021)
ERA5	31 km(T639)	hourly	1950-	Global	<a href="https://www.ecmwf.int/en/">https://www.ecmwf.int/en/</a> (last accessed: 18 Jun. 2021)
JRA55	T319(~55 km)	3 hourly	1958-	Global	<a href="https://rda.ucar.edu/datasets/ds628.0/">https://rda.ucar.edu/datasets/ds628.0/</a> (last accessed: 18 Jun. 2021)
MERRA2	1/2° × 2/3°	hourly	1979-	Global	<a href="https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/">https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/</a> (last accessed: 18 Jun. 2021)
NCEP R2	T62(200 km)	6 hourly	1979-	Global	<a href="https://rda.ucar.edu/datasets/ds091.0/">https://rda.ucar.edu/datasets/ds091.0/</a> (last accessed: 18 Jun. 2021)
OAFlux	1°	daily	1985–2009	Global ice-free ocean	<a href="http://oafux.whoi.edu/">http://oafux.whoi.edu/</a> (last accessed: 18 Jun. 2021)
TropFlux	1°	daily	1979-	30°S–30°N ocean	<a href="https://incois.gov.in/tropflux/">https://incois.gov.in/tropflux/</a> (last accessed: 18 Jun. 2021)
J-OFURO3	0.25°	daily	1988–2013	Global ice-free ocean	<a href="https://j-ofuro.scc.u-tokai.ac.jp/en">https://j-ofuro.scc.u-tokai.ac.jp/en</a> (last accessed: 18 Jun. 2021)
NOCS v2.0	1°	monthly	1973–2014	Global ocean	<a href="http://www.noc.soton.ac.uk/">http://www.noc.soton.ac.uk/</a> (last accessed: 18 Jun. 2021)

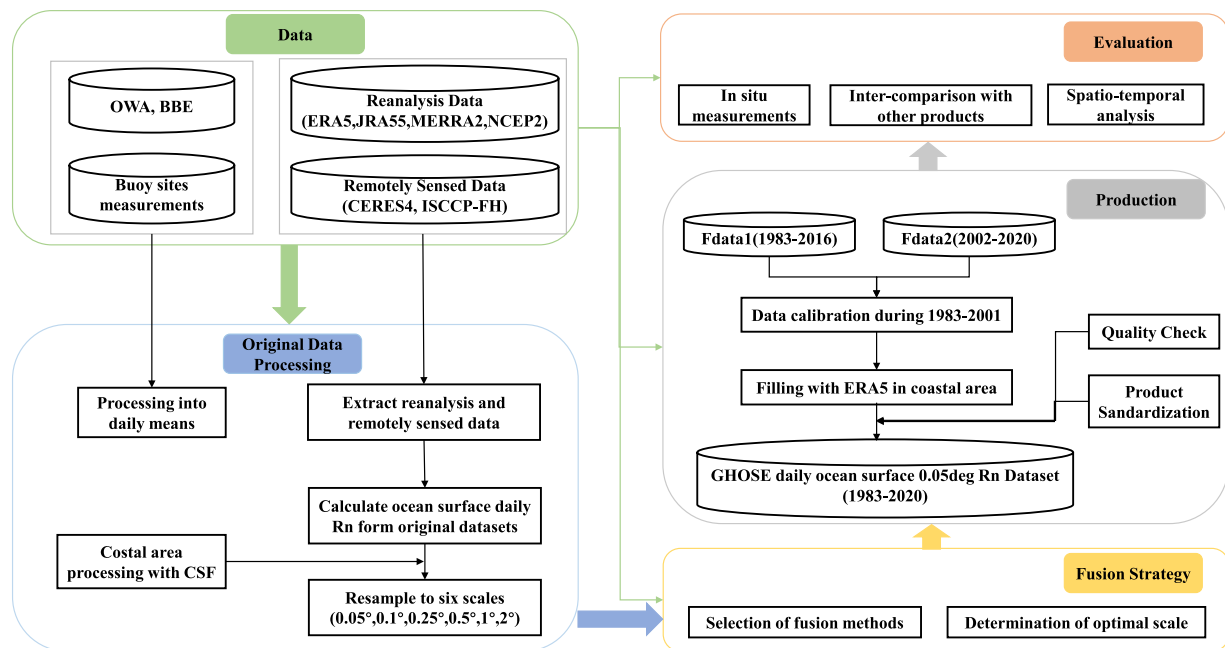
**Table 2.** The characteristics of the 11 used ocean surface  $R_n$  products.

Note that J-OFURO3 which was thought to be one of the best products in ocean surface  $R_n$ <sup>13</sup> was mainly considered as a reference. Table 2 lists the characteristics of the 11 used ocean surface  $R_n$  products.

Since only radiative components could be provided by the 11 products, the ocean surface  $R_n$  should be calculated by adding these radiative components at first. Only except the three reconstructed products and NOCS V2.0 that providing  $R_{ns}$  and  $R_{nl}$ , all of the other products providing four radiative components ( $R_{si}$ ,  $R_{so}$ ,  $R_{fb}$ , and  $R_{lo}$ ). After that, the ocean surface  $R_n$  from all the 11 products were pre-processed into daily means in UTC format, and the six products used for fusion were resampled into six scales (0.05°, 0.1°, 0.25°, 0.5°, 1°, and 2°) using bilinear interpolation<sup>58,59</sup>. Specifically, for fusion modeling, these six resampled products were extracted according to the locations of all moored buoys one by one; for production, the areas of the ocean of these products were first identified by the ocean-land mask originally developed at Boston University based on MODIS data<sup>60</sup> before resampling. To obtain more reasonable values of  $R_n$  over coastal seas, the pixels of these six products along land-sea boundaries were further processed by the creeping sea fill (CSF) method<sup>61</sup>, using which the targeted points were filled through the weight of the eight nearest pixels: a weight of one at the four diagonal points and a weight of two at up, down, left, and right points that have values, the target point is filled if the weights sum to at least three. The CSF method has been widely used for coastline processing<sup>2,62,63</sup> to reduce the influence of contaminations from land.

**Workflow of data generation and evaluation.** Figure 2 shows a flowchart of the generation and evaluation of the GHOSE daily ocean surface  $R_n$ . It includes three parts: the development of a fusion strategy, the generation and evaluation of the GHOSE daily ocean surface  $R_n$ . The fusion strategy involved determining the fusion methods, products of fusion, and the optimal spatial scale based on the *in situ* ocean surface  $R_n$  samples from 1988 to 2016, followed by the application of the fusion strategy to generate the GHOSE daily ocean surface  $R_n$  from 1983 to 2020. And the evaluation includes direct validation and inter-comparison with other products in accuracy and spatio-temporal variations. The details are provided below.

**Fusion model development.** In order to generate a long term ocean surface  $R_n$  dataset covering 1983–2020, the fusion was decided to conduct for two periods (1988–2016, 2002–2016) separately because of the different lasting years of the products for fusion (see Table 2) after multiple experiments. Four commonly used



**Fig. 2** The flowchart of generating and evaluating the GHOSE daily ocean surface  $R_n$ .

data fusion methods, SR, STS, BMA and BP, were explored one by one at six spatial scales (0.05°, 0.1°, 0.25°, 0.5°, 1°, and 2°). ISCCP-FH and the four reanalysis products were used for the fusion model building for 1988–2016 (called Fmod1), CERES4 and the four reanalysis products were used for establishing the fusion model for 2002–2016 (called Fmod2). Note that the beginning year was set to 2002 in Fmod2 because of the unstable performance in CERES4 in 2000–2001<sup>64</sup>. To determine the optimal fusion method for the ocean surface  $R_n$ , five-fold cross-validation (5-CV) was applied in this study. All training samples were randomly divided into five subsets, of which one subset was used for validation and the remaining four subsets for model fitting. This procedure was repeated five times until each of the five subset had been used for validation and training once, and the statistics of the results of validation were then compiled<sup>65</sup> and taken as the final results. Three statistical measures were used to represent the accuracy of validation: the root mean-squared difference (RMSD,  $\text{Wm}^{-2}$ ), correlation coefficient squared ( $R^2$ ), and bias ( $\text{Wm}^{-2}$ ).

The results of 5-CV modeling for the four methods at the six spatial scales in Fmod1 and Fmod2 are presented and compared in Fig. 3. The results indicated that the performance of SR and BP was very similar and relatively better than other two methods, with much smaller biases and comparable values of RMSD and  $R^2$ . Combined with the performance of the models at the six spatial scales and their implementation, the SR method at a spatial scale of 0.05° was determined to be the most appropriate fusion strategy, both for Fmod1 and Fmod2, because of its high efficiency and accuracy (RMSDs of 26.6 and 21.57  $\text{Wm}^{-2}$ , biases of  $0.67 \times 10^{-3}$  and  $-0.132 \times 10^{-2}$   $\text{Wm}^{-2}$ , and  $R^2$  values of 0.83 and 0.88, respectively). The overall performance of Fmod2 was superior to that of Fmod1 for all four fusion methods, with an overall smaller RMSD and larger  $R^2$ . We think that the contribution of CERES4, considered to be one of the most reliable radiative products, was critical here.

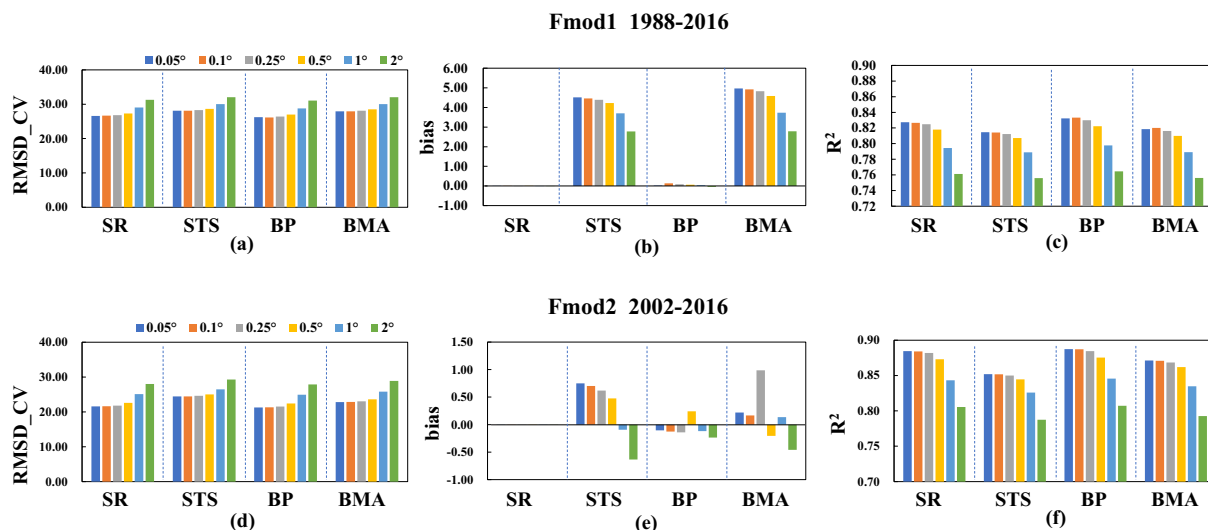
Table 3 gives the coefficients of the SR model in the finally determined Fmod1 and Fmod2. The two remotely sensed products (ISCCP-FH and CERES4) occupied a large portion, with larger values in the regression coefficient than other products in Fmod1 and Fmod2. This meant that the performance of the two remotely sensed products was closer to the *in situ* ocean surface  $R_n$ , than the other four reanalysis products.

**Data consistency processing.** As described above, two fused datasets could be obtained by using the two fusion models namely Fdata1 for 1983–2016 and Fdata2 for 2002–2016 respectively. Hence, to obtain a consistency long-term ocean surface  $R_n$ , it was necessary to reasonably combine the Fdata1 and Fdata2. The performance of the two fused datasets during the overlapping period of 2002–2016 was examined, and is shown in Fig. 4.

The results showed that Fdata2 was more accurate than Fdata1, with a smaller RMSD of 21.39  $\text{Wm}^{-2}$ , a smaller bias of  $-0.07$   $\text{Wm}^{-2}$ , and a higher  $R^2$  of 0.888 (Fig. 4a,b), although the two datasets were linearly related to each other well (Fig. 4c). Therefore, a simple calibration method was performed by regressing Fdata1 to Fdata2 as below:

$$Fdata1_r = 1.0116 \times Fdata1 - 1.459 \quad (3)$$

where  $Fdata1_r$  represents the calibrated Fdata1, and the two regression coefficients were determined based on all samples during 2002–2016. Then, Eq. (3) was used on Fdata1 in 1983–2001 by assuming that the relationship between the datasets was stable. The calibrated Fdata1 for 1983–2001 was then combined directly with Fdata2 from 2002–2016, while the Fdata2 from 2016 onward (2017–2020) could be also combined directly. Based on



**Fig. 3** Comparison of the CV accuracy results represented by RMSD ( $\text{Wm}^{-2}$ ), bias ( $\text{Wm}^{-2}$ ) and  $R^2$  of four fusion methods (SR, STS, BP, and BMA) at six scales (0.05°, 0.1°, 0.25°, 0.5°, 1°, and 2°) in (a–c) Fmod1 during 1988–2016 and (d–f) Fmod2 during 2002–2016.

Regression Coefficients						
Intercept	ISCCP-FH	CERES4	ERA5	JRA55	MERRA2	NCEP R2
<b>Fmod1 (1988–2016, No. of training samples: 45,629)</b>						
4.936	0.729*	\	0.137*	0.029*	0.017*	0.016*
<b>Fmod2 (2002–2016, No. of training samples: 38,034)</b>						
-1.302	\	1.001*	0.0235*	0.008*	-0.014*	-0.024*

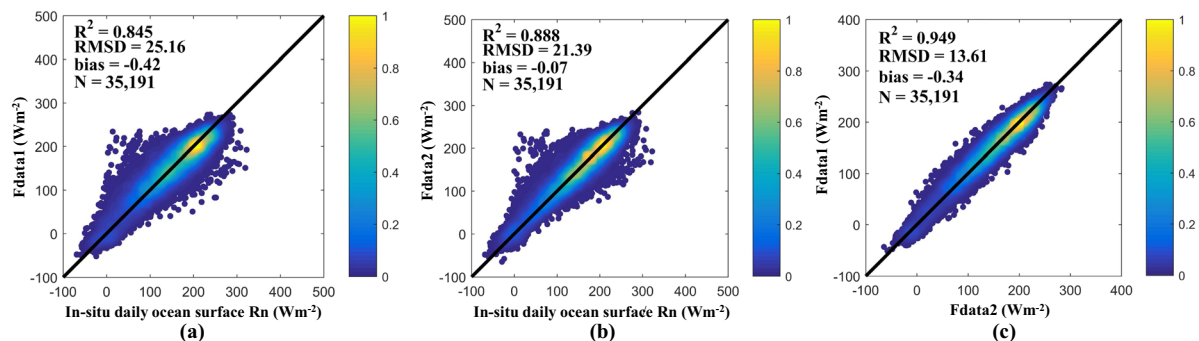
**Table 3.** The regression coefficients of the two fusion models.

this, the dataset of a long-term, fused, global daily ocean surface  $R_n$  from 1983 to 2020 at 0.05° over global was generated.

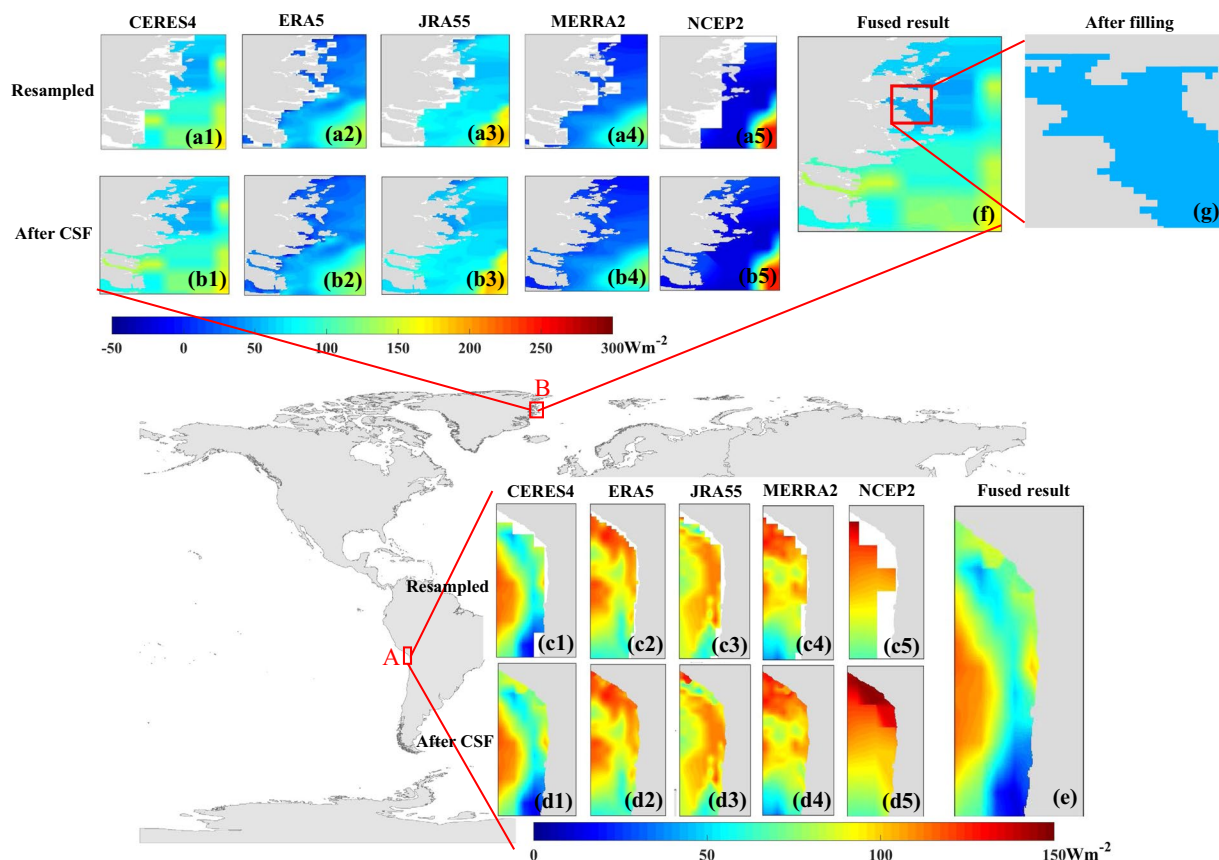
**Post-processing.** Missing data still persisted after fusion. We use the fused daily ocean surface  $R_n$  on a randomly selected day, 150, in 2005 as an example. The details of two areas with (box B) and without (box A) missing data are shown in Fig. 5. The missing data usually appeared over seas very close to coastlines that were irregular, such as the area in box B near Greenland (Fig. 5f), while no missing data were found over coastal seas, where their coastlines were relatively smooth, such as the area in box A in Fig. 5e. For better illustration, the five original products used for fusion, and their processed and resampled results using the CSF method for the two areas are presented in Fig. 5a–d. A comparison of the resampled results of the six products in area A (Fig. 5d1–d5) and area B (Fig. 5b1–b5) suggested that the missing data due to resampling could not be completely filled because of the complicated and irregular coastline, the coarse resolution of the original products, and the limitation of the CSF method itself<sup>61</sup>. All missing data in the final, fused dataset were filled with the resampled daily ocean surface  $R_n$  (0.05°) from ERA5 by considering its large spatial resolution and reliable performance. Figure 5g shows the results of filling of the magnified box in Fig. 5f. Finally, the GHOSE daily ocean surface  $R_n$  covering the globe at 0.05° from 1983 to 2020 without any missing data was generated.

### Data Records

All GHOSE daily ocean surface  $R_n$  from 1983–2020 are freely available from the Big Earth Data Repository (<http://data.casearth.cn/>)<sup>66</sup>, through which users can link to the specific data entities of each year as well as the user note. The total amount of this dataset is up to 356 Gigabyte. It was stored in thirty-eight folders named after the year, in which each file is stored in the Hierarchical Data File (HDF) format containing the values of daily ocean surface  $R_n$  and the corresponding quality check (QC) flags. Daily files are named as “GHOSE07B03.V10.AYYYYDDD.yyyyddd.hdf” where “GHOSE07B03”, “V10”, “YYYY”, and “DDD” denote the product name, version number, year, and Julian day of year (doy), while the lower case “yyyddd” represents the year and do of the processing date. All daily values are provided in UTC time in WGS84 geographic coordinate system and stored as 32-bit integer data type in the unit of  $\text{Wm}^{-2}$  through the conversion by multiplying a scale factor of 0.01, and their valid range was from -150 to 350  $\text{Wm}^{-2}$  only with the value of -9999 representing land surface. The QC flags are dimensionless and stored as 8-bit unsigned integer data type for providing information about the sources and quality of the estimates for users reference, and the detailed information is shown in Table 4.



**Fig. 4** The 5-CV results against the *in-situ* data of (a) Fdata1 and (b) Fdata2 during 2002–2016, and (c) is the scatter plot between Fdata1 and Fdata2 during 2002–2016.



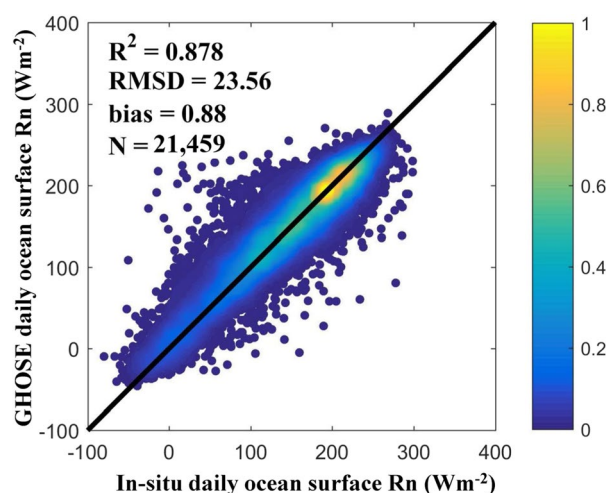
**Fig. 5** The comparison of two areas with (box B) or without (box A) missing data after fusing: (a1–a5) and (c1–c5) are the daily ocean surface  $R_n$  from the five resampled ( $0.05^\circ$ ) products (CERES4, ERA5, JRA55, MERRA2, and NCEP2) used for fusion, (b1–b5) and (d1–d5) are their corresponding processed results with the CSF method, (e) and (f) are the fused results, and (g) is the final result after filling with the resampled ocean surface  $R_n$  from ERA5 within the magnified box in (f).

## Technical Validation

**Direct evaluation.** A scatter plot between the GHOSE daily ocean surface  $R_n$  and the *in situ*  $R_n$  during 1988–2019 is shown in Fig. 6. The overall evaluation of accuracy yielded an  $R^2$  of 0.878, an RMSD of  $23.56 \text{ Wm}^{-2}$ , and a bias of  $0.88 \text{ Wm}^{-2}$ . As mentioned above, the fusion model was established based on various datasets from 1988 to 2016; hence, the accuracy of evaluation was further examined for the forward period of 2017–2019. Comfortingly, the uncertainties in the estimated ocean surface  $R_n$  for the two periods were similar to or even better than those during 1988–2019, with  $R^2$  values of 0.861 and 0.921, RMSDs of  $24.21$  and  $21.09 \text{ Wm}^{-2}$ , and biases of  $0.37$  and  $2.72 \text{ Wm}^{-2}$ , respectively. Hence, the forecasting performance of the fusion model was satisfactory for the short term.

Digits	Description	Bit combination	Detailed information
1~2	ocean/land	00	land
		10	ocean
		11	ocean and ice mixing area
3~5	quality	000	missing (no estimated values due to missing input data in some cases)
		001	missing (inland water)
		011	missing (no estimated values caused by interpolation algorithm, filled with neighboring pixel value)
		010	good
		100	accept
		110	uncertainty
6~8	algorithm	000	fusion
		001	filling with CSF before fusion
		011	Filling with ERA5 of the merged vacancy value in coastline area

**Table 4.** Explanations of the QC flag in the GHOSE daily ocean surface  $R_n$ .



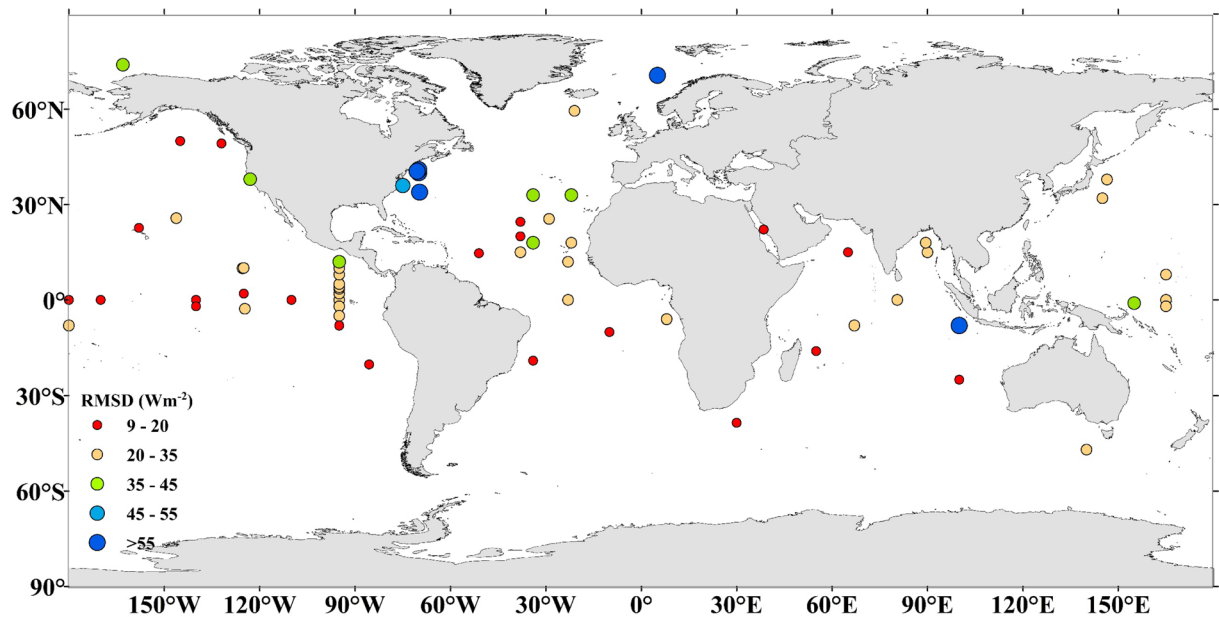
**Fig. 6** The scatter plot between the daily ocean surface  $R_n$  from GHOSE and *in situ* samples during 1988–2019.

Meanwhile, the spatial distribution of the performance of the GHOSE daily ocean surface  $R_n$  represented by the RMSD at each moored site was calculated and presented in Fig. 7. It shows that the closer to the coastal area a moored buoy was, the larger was its estimated uncertainty. Thus, GHOSE should be used with caution in the coastal seas.

And then, the performance of GHOSE was inter-compared with other products. At first, the evaluation accuracy in GHOSE ocean surface  $R_n$  and its data sources, the six products used for fusion (CERES4, ISCCP-FH, ERA5, JRA55, MERRA2, and NCEP R2), was calculated against the independent samples during the two fusion periods of 1988–2016 and 2002–2016 and given in Table 5. The daily ocean surface  $R_n$  from GHOSE outperformed the original six products for both periods. Specifically, for 1988–2016, the uncertainty of GHOSE was significantly lower than that of ISCCP-FH, with an RMSD of  $2.82 \text{ Wm}^{-2}$  and a magnitude of bias as large as  $9.11 \text{ Wm}^{-2}$ . For 2002–2016, the uncertainty in GHOSE was the lowest, with an  $R^2$  of 0.886, an RMSD of  $21.88 \text{ Wm}^{-2}$ , and a bias of  $0.35 \text{ Wm}^{-2}$ , similar to but still better than those of CERES4 — an RMSD lower by  $0.27 \text{ Wm}^{-2}$  and a bias of nearly  $2 \text{ Wm}^{-2}$ . Combined with the results of evaluation of the six products used for fusion, it seemed that their positive and negative biases could be offset by fusion to obtain a smaller magnitude of the final bias. Moreover, the uncertainty of GHOSE was similar to those in the two remotely sensed products (ISCCP-FH and CERES4) for the two periods. This again demonstrated the significant contribution of the two products to fusion.

Furthermore, four other products (GEWEX-SRB, J-OFURO3, OAFlux, and TropFlux) were also evaluated for reference (Table 6). By considering the different durations of the four products, GHOSE was evaluated according to each of their periods for better comparison. The results show that GHOSE outperformed each of the four products. Although the evaluation periods were different, GEWEX-SRB delivered the worst performance in terms of the daily ocean surface  $R_n$ , with an RMSD of  $42.74 \text{ Wm}^{-2}$  and a bias of  $14.58 \text{ Wm}^{-2}$ . GHOSE reduced the RMSD and bias by  $12.91$  and  $14.04 \text{ Wm}^{-2}$ , respectively. It also outperformed OAFlux and TropFlux, reconstructed from ISCCP-FD and ERA-Interim, respectively, in terms of daily ocean surface  $R_n$  by recording RMSD values lower by  $12.36$  and  $14.86 \text{ Wm}^{-2}$ , and the magnitudes of bias lower by  $7.71$  and  $2.32 \text{ Wm}^{-2}$ ,





**Fig. 7** Spatial distribution of RMSD value of the daily ocean surface  $R_n$  from GHOSE at 68 moored buoy sites from 1988–2019.

product	1988–2016			2002–2016		
	RMSD ( $\text{Wm}^{-2}$ )	bias ( $\text{Wm}^{-2}$ )	$R^2$	RMSD ( $\text{Wm}^{-2}$ )	bias ( $\text{Wm}^{-2}$ )	$R^2$
CERES4	\	\	\	22.15	2.34	0.88
ISCCP-FH	29.27	9.22	0.824	\	\	\
ERA5	34.73	−0.42	0.712	33.56	0.12	0.726
JRA55	44.53	−5.95	0.562	44.07	−6.06	0.566
MERRA2	43.19	−5.94	0.605	42.41	−6.06	0.613
NCEP R2	50.19	−9.55	0.479	54.04	−8.41	0.401
GHOSE	24.12	0.37	0.861	21.88	0.35	0.886

**Table 5.** The evaluation results of the daily ocean surface  $R_n$  from GHOSE and the six products used for fusing.

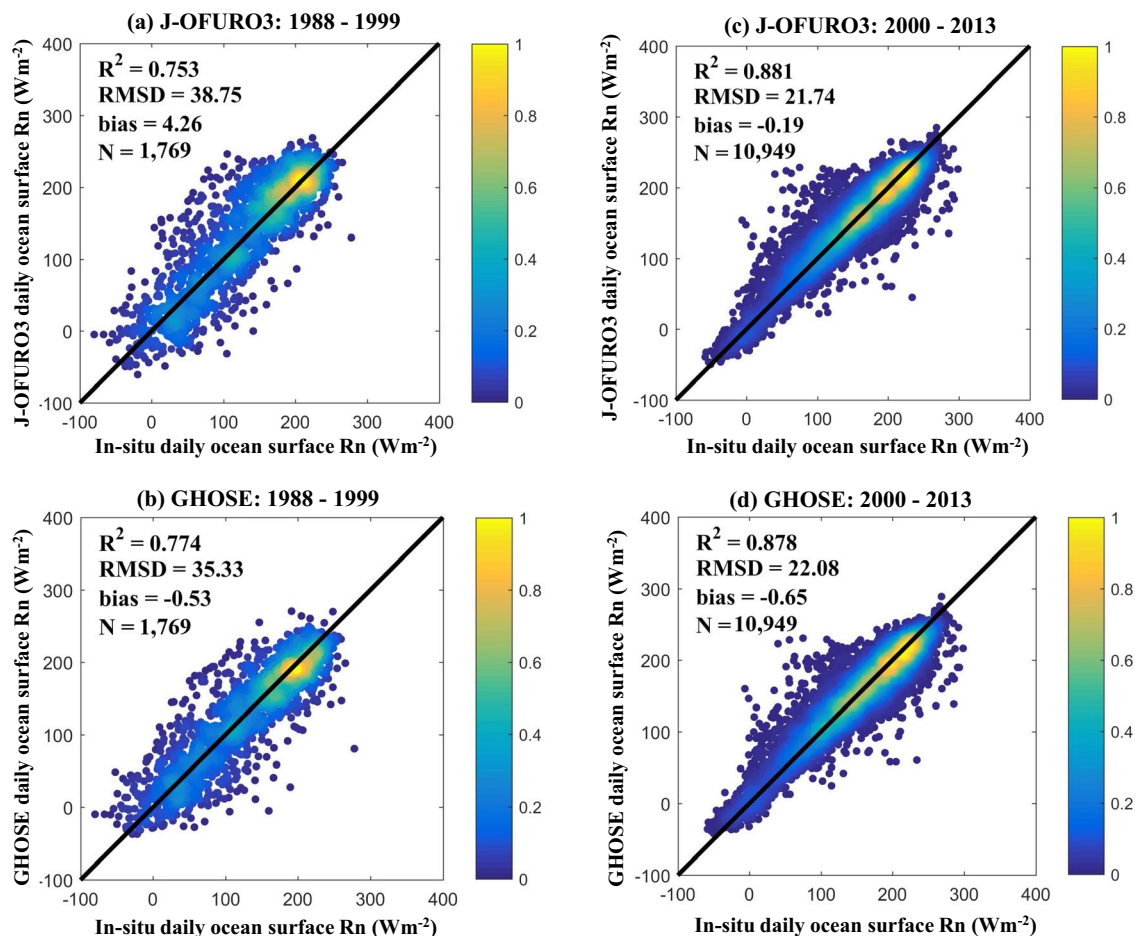
product	period	No. of samples	RMSD ( $\text{Wm}^{-2}$ )	bias ( $\text{Wm}^{-2}$ )	$R^2$	GHOSE		
						RMSD ( $\text{Wm}^{-2}$ )	bias ( $\text{Wm}^{-2}$ )	$R^2$
GEWEX-SRB	1988–2007	4,630	42.74	14.58	0.66	29.83	0.54	0.79
J-OFURO3	1988–2013	12,718	24.81	0.43	0.86	24.36	−0.63	0.86
OAFflux	1988–2009	6,969	36.82	7.72	0.71	24.46	−0.10	0.83
TropFlux	1988–2016	13,557	38.21	−3.28	0.57	23.35	0.96	0.81

**Table 6.** The evaluation results of GHOSE and four reference products in daily ocean surface  $R_n$ .

respectively. Among the four products, J-OFURO3 delivered the best performance, slightly worse than that of GHOSE, during its period. They had RMSDs of 24.81 and 24.36  $\text{Wm}^{-2}$  and biases of 0.43 and  $-0.63 \text{ Wm}^{-2}$ , respectively. This was coincident with the conclusion reported by Chen *et al.*<sup>13</sup>.

**Further comparison with J-OFURO3.** According to Chen *et al.*<sup>13</sup>, it is worth considering inconsistencies in the daily ocean surface  $R_n$  from J-OFURO3 around 2000, which was primarily caused by the change in input datasets from ISCCP-FD to CERES3. We compared GHOSE and J-OFURO3 in terms of the ocean surface  $R_n$  for 1988–1999 and 2000–2013 based on the independent evaluation samples. The results are shown in Fig. 8.

The evaluation results of both J-OFURO3 and GHOSE were worse before 2000 than after, with RMSD values of 38.75 and 35.33  $\text{Wm}^{-2}$ , respectively, before 2000 (Fig. 8a,b), and 21.74 and 22.58  $\text{Wm}^{-2}$  after 2000 (Fig. 8c,d), respectively. The quality of the radiative products was generally poor before 2000, possibly because of the limited available moored buoy measurements (No. of samples = 1,769) in this period. Moreover, a lower uncertainty in the GHOSE daily ocean surface  $R_n$  during 1988–1999 occurred possibly because ISCCP-FH was used as the major input in this period. Figure 9 presents the variations in the anomalies of the global averaged annual mean



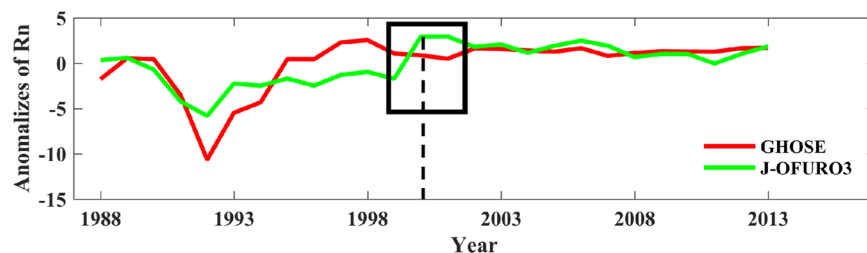
**Fig. 8** The evaluation results of GHOSE and J-OFURO3 in daily ocean surface  $R_n$  against the *in situ* samples during 1988–1999 ((a) and (b)) and 2000–2013 ((c) and (d)), respectively.

daily ocean surface  $R_n$  from GHOSE (red) and J-OFURO3 (green line), and the inconsistency (black box in Fig. 9) for J-OFURO3 was not observed in GHOSE around 2000.

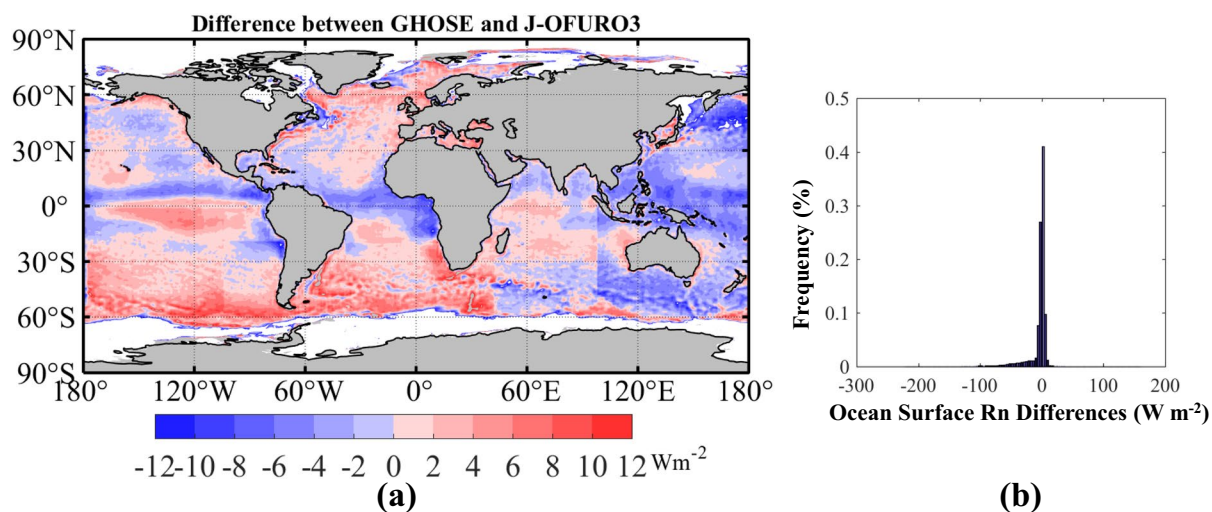
And the spatial distribution and histogram of the differences in the annual mean daily ocean surface  $R_n$  obtained by GHOSE minus that of J-OFURO3 during 1988–2013 over global ice-free ocean are shown in Fig. 10a,b. On the whole, the differences in the ocean surface  $R_n$  of the two products were mostly at around  $\pm 10 \text{ Wm}^{-2}$ , but that of GHOSE was mostly smaller than that of J-OFURO3, especially over seas at high latitudes, while the opposite case was obtained over the West Pacific.

Ten examples over global oceans for one year were randomly selected and are shown in Fig. 11. The variations in the daily ocean surface  $R_n$  from J-OFURO3, GHOSE, and moored buoys, and the scatter plot between the results of GHOSE and the *in situ* samples are presented for the 10 sites. Note that J-OFURO3 was not shown when the selected year was beyond its period of 1988–2013. For most cases, the daily ocean surface  $R_n$  from the two products and the *in situ* samples were in good agreement, especially at sites over open seas at mid-to-low latitudes, such as UOP\_PACS\_south (Fig. 11b2), RAMA\_01 (Fig. 11d2), RAMA\_06 (Fig. 11f2), and PIRATA\_07 (Fig. 11g2). For the other cases, the two products agreed well, but were not as close over coastal seas, such as site UOP\_CB\_A (Fig. 11a2)). Moreover, GHOSE performed generally better than J-OFURO3 before 2000, as exemplified by site UOP\_SUB\_CE (Fig. 11h2), and at high-latitudes, as in case of site UOP\_MLML91 (Fig. 11e2). However, the available *in situ* samples were too few to be used for comprehensive evaluation, such as in the winter season and before 2000, which means more research is needed to better quantify the uncertainty in the daily ocean surface  $R_n$  obtained by GHOSE, J-OFURO3, and other products.

**Spatio-temporal analysis and comparison.** Values of the global average annual mean ocean surface  $R_n$  from six long term products, including three products used for fusion (ISCCP-FH, CERES4, and EAR5) and three reference products (J-OFURO3, OAFlux, and NOCS2.0), were calculated according to their years of coverage during 1983–2020, and were compared with the results of GHOSE. Note that the area-weighted method, which considers different areas of each pixel by multiplying the cosine of its latitude, was applied to calculate the globally or regionally averaged ocean surface  $R_n$ . Only ice-free oceans were considered because J-OFURO3 was used. The results are presented in Fig. 12. It shows a large discrepancy among the seven products, especially before 2000.



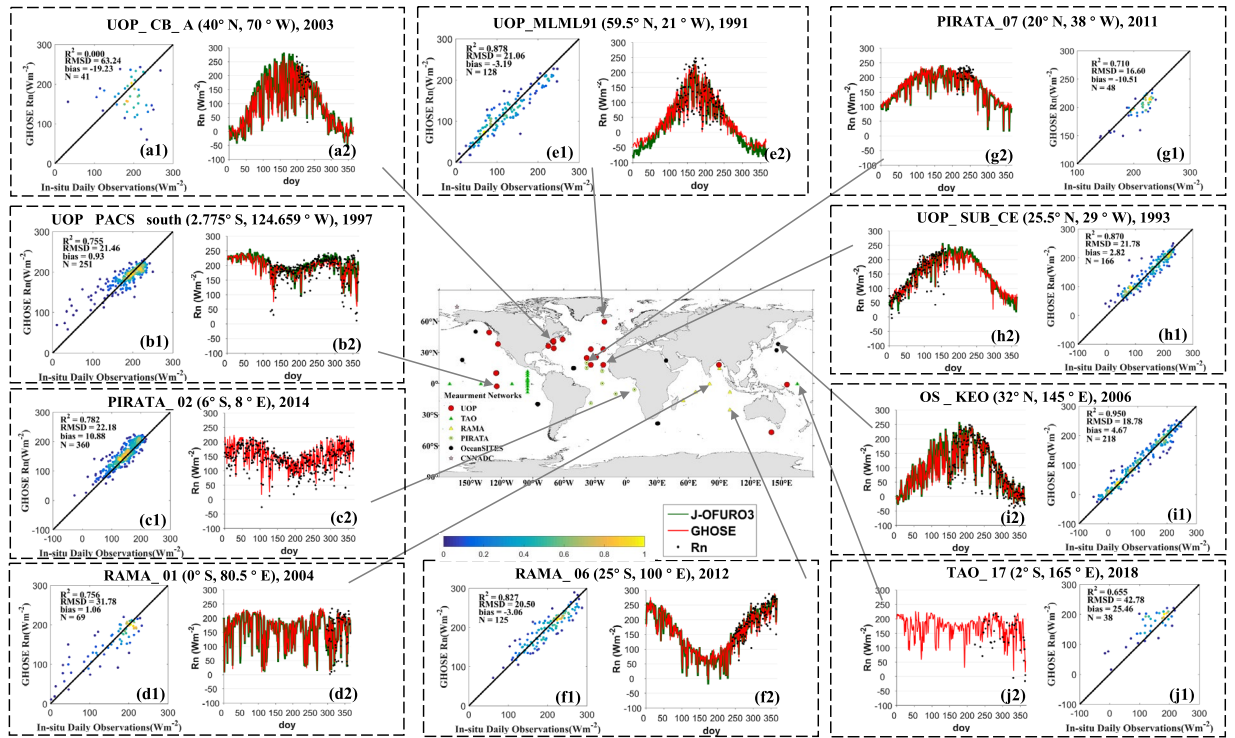
**Fig. 9** The variations in the anomalies of the global averaged annual mean daily ocean surface  $R_n$  from GHOSE and J-OFURO3 during 1988–2013 over the ice-free oceans determined by J-OFURO3. The dashed line indicates the year of 2000.



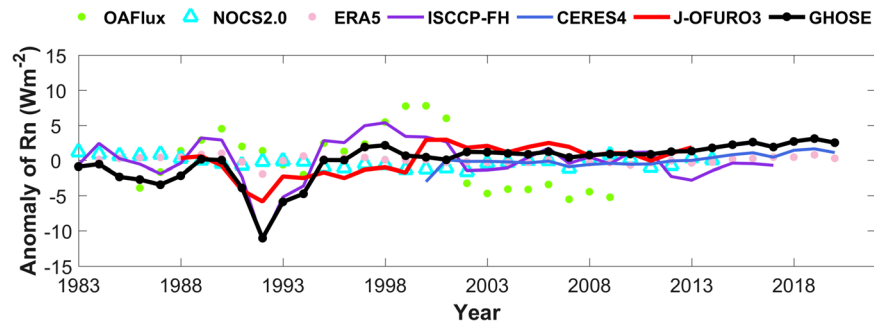
**Fig. 10** (a) The spatial distribution of the differences in the annual mean daily ocean surface  $R_n$  by GHOSE minus J-OFURO3 from 1988 to 2013 over global ice-free oceans, and (b) its corresponding histogram. White color represents the missing data.

The variations in the ocean surface  $R_n$  due to GHOSE were similar to those in ISCCP-FH before 2000 and those of CERES4 afterward. The variations in ERA5 and NOCS2.0 were the smoothest throughout, even around 1992, with a sudden drop caused by the eruption of Mount Pinatubo<sup>67</sup>. This was captured by almost all products, and especially well by NOCS2.0. The variations in J-OFURO3 were still significantly different from those of GHOSE, in addition to the inconsistency around 2000, although the results of their direct evaluation were similar.

The spatial distribution of long term trend ( $p < 0.05$ ) in annual mean GHOSE ocean surface  $R_n$  over global oceans during 1983–2020 is calculated using the linear regression method and shown in Fig. 13a. It indicates that most oceans underwent a significant change during 1983–2020 with mostly increasing at mid-high latitudes (red color) and only several increasing (e.g., the Indian Ocean and Tropical Center Pacific) and decreasing hot spots (e.g. the seas to northern Australia, western Africa and South America) over tropical seas, which were coincident with the results of previous studies<sup>13,68–70</sup>. The remarkable increased ocean surface  $R_n$  over most oceans at high latitudes where the rate was even as large as  $2 \text{ Wm}^{-2}$  per year was speculated caused by the increased absorbed solar radiation because of the reduced ocean surface albedo due to the sea ice or glacier melting caused by global warming<sup>71</sup>, but the case was unexpected over the Norway sea ( $68^\circ\text{N} - 80^\circ\text{N}$ ,  $10^\circ\text{W} - 45^\circ\text{E}$ ) where the ocean surface  $R_n$  overly decreased at a rate of  $\sim 0.5 \text{ Wm}^{-2}$  per year. According to Kim *et al.*<sup>72</sup> and Wen *et al.*<sup>73</sup>, the decreased ocean surface  $R_n$  occurred possibly because of the decreased SST owing to the inhibited northward flow of Atlantic Meridional Overturning Circulation (AMOC) in this area. This was prompted in turn by the increased freshwater from melting glaciers. In addition, the decrease in  $R_n$  over seas near to West Africa might have been caused by biomass burning aerosols (BBAs) over the continent, which cooled the ocean by absorbing shortwave radiation through complex interactions among the ocean, land, and the atmosphere<sup>74,75</sup>. Meanwhile, the long term trends in the area-averaged anomalies of the annual mean GHOSE ocean surface  $R_n$  for 1983–2020 and 2000–2020 for global and global ice-free oceans were calculated and given in Fig. 13b1–c1. The global averaged linear trend in Fig. 13b1 over 1983–2020 was in accordance with the results in Fig. 13a, with a significant rate of increase of  $0.16 \text{ Wm}^{-2}$  per year, which decreased by  $0.05 \text{ Wm}^{-2}$  per year after 2000, although the positive trend was still significant. Combined with the results in Fig. 13c1, the trend of a relatively smaller increase over global ice-free oceans than that over global oceans further indicated that seas with ice at high latitudes contributed significantly to the increased  $R_n$ .

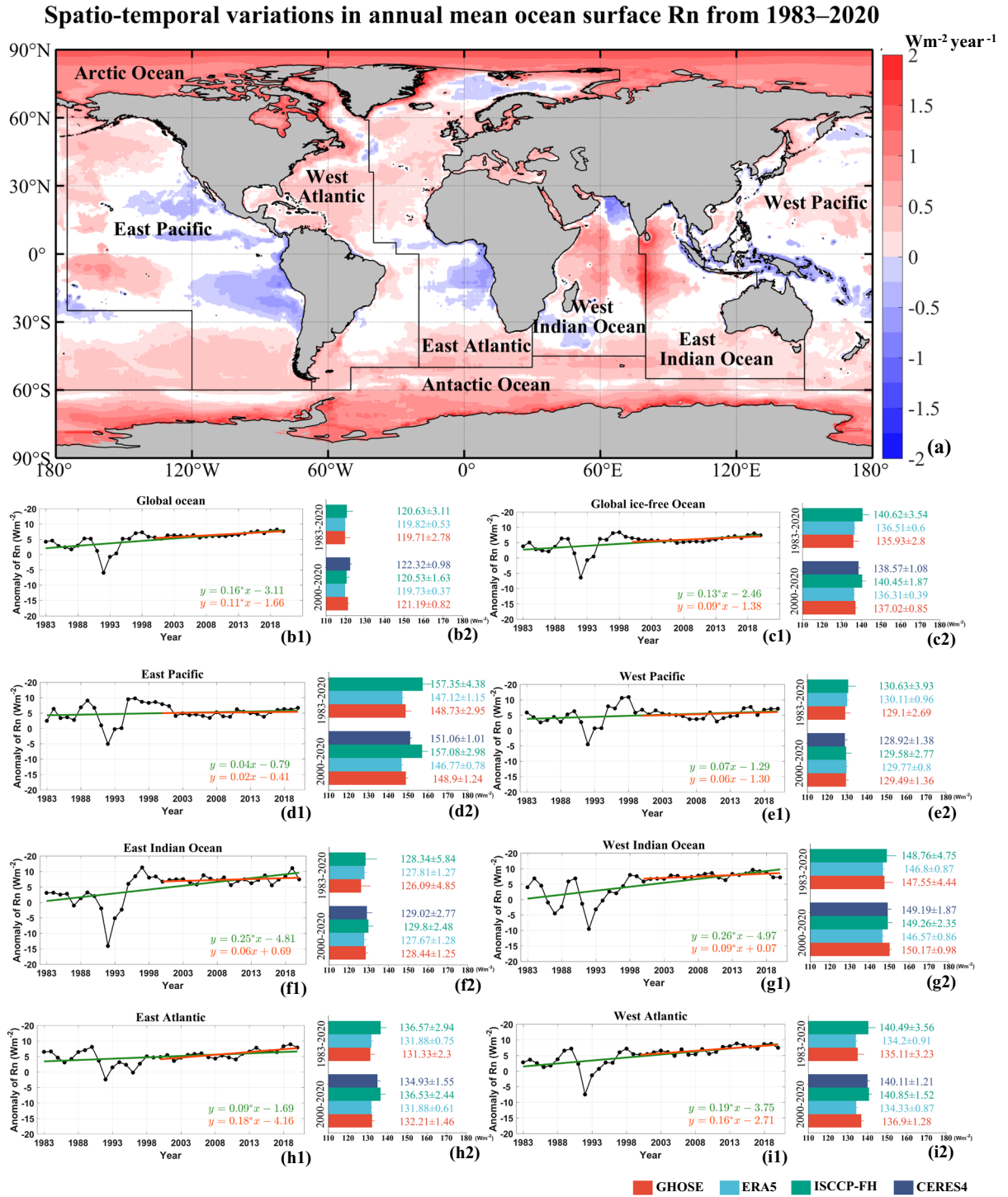


**Fig. 11** The temporal variations in daily ocean surface  $R_n$  from J-OFURO3 (green line), GHOSE (red line), and moored buoy (black dot) and the scatter plot between GHOSE and moored  $R_n$  for one randomly selected year at ten moored buoy sites: (a1–2) UOP\_CB\_A (40°N, 70°W) in 2003, (b1–2) UOP\_PACS\_south (2.775°S, 124.659°W) in 1997, (c1–2) PIRATA\_02 (6°S, 8°E) in 2014, (d1–2) RAMA\_01 (0°S, 80.5°E) in 2004, (e1–2) UOP\_MLML91 (59.5°N, 21°W) in 1991, (f1–2) RAMA\_06 (25°S, 100°E) in 2012, (g1–2) PIRATA\_07 (20°N, 38°W) in 2011, (h1–2) UOP\_SUB\_CE (25.5°N, 29°W) in 1993, (i1–2) OS\_KEO (32°N, 145°E) in 2006, (j1–2) TAO\_17 (2°S, 165°E) in 2018.



**Fig. 12** The temporal variations in the global averaged annual mean ocean surface  $R_n$  from seven products according to their lasting years during 1983–2020.

For better analysis, the global ice-free oceans (55°S–55°N) were divided into six sub-regions according to the definition provided by the Food and Agriculture Organization of the United Nations (<http://www.fao.org/geonetwork>, last accessed: 10 Jan. 2020): the East Pacific Ocean, West Pacific, East Atlantic, West Atlantic, East Indian Ocean, and West Indian Ocean, as shown in Fig. 13a. Therefore, the area-averaged linear trends of GHOSE annual mean ocean surface  $R_n$  for the two same periods were also calculated for the six sub-regions one by one and given in Fig. 13d1–i1. Based on the results, the ocean surface  $R_n$  of nearly all six sub-regions over the global ice-free oceans increased either for 1983–2020 or 2000–2020. Some had a higher rate of increase than the global average (e.g., the Atlantic and Indian Oceans). The ocean surface  $R_n$  increased the most quickly and significantly over the Indian Ocean, by more than 0.25  $\text{Wm}^{-2}$  per year, during 1983–2020 (Fig. 13f1,g1), especially the West Indian Ocean (Fig. 13g1) at 0.26  $\text{Wm}^{-2}$  per year. However, the trend has slowed since 2000 over the entire Indian Ocean, with rates of 0.06 and 0.09  $\text{Wm}^{-2}$  per year for the East and West Indian Oceans, respectively. Arora *et al.*<sup>69</sup> have noted that 70% of the missing heat from the atmosphere is trapped in the Indian Ocean as heat content, which in turn influences the SST. This might explain the increase in  $R_n$  in this region. The ocean surface  $R_n$  over the West Atlantic also had significant rates of increase of 0.19  $\text{Wm}^{-2}$  per year for



**Fig. 13** (a) The spatial distribution of the long term trend ( $p < 0.05$ ) in the annual mean ocean surface  $R_n$  from GHOSE during 1983–2020 over global oceans. (b1)–(i1) are the trends in the anomalies of the area-averaged annual mean daily ocean surface  $R_n$ , over global oceans, global ice-free oceans and six sub-regions within the global ice-free oceans ( $55^\circ\text{S}$ – $55^\circ\text{N}$ ) during 1983–2020 (green line) and 2000–2020 (red line), respectively, and (b2)–(i2) are the area averaged multi-annual mean and standard deviations in ocean surface  $R_n$  ( $\text{Wm}^{-2}$ ) from four products over the global oceans, global ice-free oceans and the six sub-regions during 1983–2020 and 2000–2020. White color in (a) represents the insignificant trend.

1983–2020 and  $0.16 \text{ Wm}^{-2}$  per year for 2000–2020 (Fig. 13i1). The value over the East Atlantic Ocean increased more slowly, at a rate of  $0.09 \text{ Wm}^{-2}$  per year over the past 38 years. This rate has been much higher since 2000, at  $0.18 \text{ Wm}^{-2}$  per year. Only observed in this region, this growth rate is also the highest among the six sub-regions for this period<sup>76</sup>. Over the East and West Pacific (Fig. 13d1,e1), the ocean surface  $R_n$  increased but insignificantly,

mostly because of the offset of the evenly distributed positive and negative trends in this region, as shown in Fig. 13a. Interestingly, the variations in the ocean surface  $R_n$  over the East and West Pacific, where the El Niño Southern Oscillation (ENSO) events occur, also reflected some information on it. For example, the ocean surface  $R_n$  increased during two ENSO events of 1987–1988 and 2010–2011 over the East Pacific and vice versa over West Pacific, which should be caused by the varied SST circulation around east-west Pacific<sup>77–79</sup>. Moreover, it seems that the influence of the eruption of Mount Pinatubo on the ocean surface  $R_n$  over all global oceans with a sudden drop around 1992.

Afterwards, the area-averaged multi-annual mean and standard deviations in GHOSE ocean surface  $R_n$  for global and global ice-free oceans, as well as the six sub-regions were calculated covering 1983–2020 and 2000–2020, respectively. For comparison, the same ones for three products (CERES4, ERA5, and ISCCP-FH) were also calculated. All results are given in Fig. 13b–i2. Note that the results of CERES4 only for 2000–2020 and that of ISCCP-FH for 1983–2017 and 2000–2017. During 1983–2020, the multi-annual mean ocean surface  $R_n$  over global oceans from GHOSE was  $119.71 \pm 2.78 \text{ Wm}^{-2}$ , close to that of ERA5 ( $119.82 \pm 0.53 \text{ Wm}^{-2}$ ), and this was also observed for all other regions; while they both were lower than that of ISCCP-FH, especially for the East Pacific (Fig. 13d2) which was thought to be one of the most sensitive regions containing the East Pacific Warm Pool and Cold Tongues with the difference value as large as  $10 \text{ Wm}^{-2}$ . For the period of 2000–2020, the global averaged multi-annual mean ocean surface  $R_n$  was  $121.19 \pm 0.82 \text{ Wm}^{-2}$ , close to that of ERA5 ( $119.73 \pm 0.37$ ) and a little bit lower than that of CERES4 ( $122.32 \pm 0.98$ ). For the six sub-regions, the multi-annual mean ocean surface  $R_n$  from GHOSE was mostly smaller than that of CERES4, except for the West Pacific, and was very close to that of ERA5 for most regions, with a larger value of  $\sim 3 \text{ Wm}^{-2}$  only over the West Indian Ocean. Similar to the results for 1983–2020, the largest discrepancies in the multi-annual mean ocean surface  $R_n$  among the four products were also observed in the East Pacific, followed by the entire Atlantic. Moreover, the standard deviations of the area-averaged multi-annual mean ocean surface  $R_n$  from ERA5 were the smallest of the products for all regions considered, which indicates its stability for the ocean surface  $R_n$ . This is coincident with the results in Fig. 12.

To sum up, the long-term daily ocean surface  $R_n$  from GHOSE was satisfactory overall because of robust results, consistent variations, fine spatial resolution, and global ocean coverage without missing data. The results of a preliminary spatio-temporal analysis highlighted its ability to adequately characterize and capture information on variations in the radiative balance on the ocean surface. Therefore, this new dataset offered promise for wide used in the near future.

### Usage Notes

Some abnormal values or mosaics were found in GHOSE ocean surface  $R_n$  over some regions for a few days, mainly in 1983 and 1988 (i.e., days 183–196 in 1983, and 153/214–230/ 237–244 in 1988). They occurred primarily because of ISCCP-FH, the quality of the daily ocean surface  $R_n$  of which was affected by issues related to the zenith angle of the satellite (<https://isccp.giss.nasa.gov/projects/flux.html>, last accessed: 8 Jan. 2022) and its  $R_{lo}$ . Hence, the daily ocean surface  $R_n$  from GHOSE on these days should be used with caution.

### Code availability

The MATLAB codes for generating and processing data and the CN-NADC *In situ* ocean surface  $R_n$  can be accessed at the figshare<sup>80</sup>.

Received: 17 March 2022; Accepted: 23 May 2022;

Published online: 14 June 2022

### References

- Fung, I. Y., Harrison, D. E. & Lacis, A. A. On the variability of the net longwave radiation at the ocean surface. *Reviews of Geophysics*. **22**, 177–193 (1984).
- Tomita, H., Hihara, T., Kako, S. I., Kubota, M. & Kutsuwada, K. An introduction to J-OFURO3, a third-generation Japanese ocean flux data set using remote-sensing observations. *Journal of Oceanography*. **75**, 171–194 (2019).
- Polavarapu, R. J. Estimation of net radiation at sea. *Atmosphere-Ocean*. **17**, 24–35 (1979).
- Baggenstos, D. *et al.* Earth's radiative imbalance from the Last Glacial Maximum to the present. *Proceedings of the National Academy of Sciences of the United States of America*. **116**, 14881–14886 (2019).
- von Schuckmann, K. *et al.* An imperative to monitor Earth's energy imbalance. *Nature Climate Change*. **6**, 138–144 (2016).
- Yao, Y. *et al.* Improving global terrestrial evapotranspiration estimation using support vector machine by integrating three process-based algorithms. *Agricultural and Forest Meteorology*. **242**, 55–74 (2017).
- Dee, D. P. *et al.* The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*. **137**, 553–597 (2011).
- Wild, M. *et al.* The energy balance over land and oceans: an assessment based on direct observations and CMIP5 climate models. *Climate Dynamics*. **44**, 3393–3429 (2014).
- Berry, D. I. & Kent, E. C. Air-Sea fluxes from ICOADS: the construction of a new gridded dataset with uncertainty estimates. *International Journal of Climatology*. **31**, 987–1001 (2011).
- Yu, L., Jin, X. & Weller, R. Multidecade Global Flux Datasets from the Objectively Analyzed Air-sea Fluxes (OAFlux) Project: Latent and Sensible Heat Fluxes, Ocean Evaporation, and Related Surface Meteorological Variables. (Woods Hole Oceanographic Institution, 2008).
- Praveen Kumar, B., Vialard, J., Lengaigne, M., Murty, V. S. N. & McPhaden, M. J. TropFlux: air-sea fluxes for the global tropical oceans—description and evaluation. *Climate Dynamics*. **38**, 1521–1543 (2011).
- Peng, J. *et al.* A New Empirical Estimation Scheme for Daily Net Radiation at the Ocean Surface. *Remote Sensing*. **13**, 4170 (2021).
- Chen, H. *et al.* Evaluation of the J-OFURO3 Sea Surface Net Radiation and Inconsistency Correction. *Remote Sensing*. **13**, 2403 (2021).
- Zhang, J. Multi-source remote sensing data fusion: status and trends. *International Journal of Image and Data Fusion*. **1**, 5–24 (2010).
- Pohl, C. & Van Genderen, J. L. Review article Multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing*. **19**, 823–854 (1998).

16. Duan, Q. & Phillips, T. J. Bayesian estimation of local signal and noise in multimodel simulations of climate change. *Journal of Geophysical Research: Atmospheres*. **115**, D18123 (2010).
17. Hoeting, J. A., Madigan, D., Raftery, A. E. & Volinsky, C. T. Bayesian Model Averaging A Tutorial Comment. *Statistical Science*. **14**, 382–401 (1999).
18. Yao, Y. *et al.* Estimation of high-resolution terrestrial evapotranspiration from Landsat data using a simple Taylor skill fusion method. *Journal of Hydrology*. **553**, 508–526 (2017).
19. Wang, D. & Liang, S. Integrating MODIS and CYCLOPES Leaf Area Index Products Using Empirical Orthogonal Functions. *IEEE Transactions on Geoscience and Remote Sensing*. **49**, 1513–1519 (2011).
20. Wang, D., & Liang, S. Using multiresolution tree to integrate MODIS and MISR-L3 LAI products. *Geoscience and Remote Sensing Symposium (IGARSS), IEEE International 1027–1030* (2010).
21. Wang, W. & Liang, S. Estimation of high-spatial resolution clear-sky longwave downward and net radiation over land surfaces from MODIS data. *Remote Sensing of Environment*. **113**, 745–754 (2009).
22. Wu, H., Zhang, X., Liang, S., Yang, H. & Zhou, G. Estimation of clear-sky land surface longwave radiation from MODIS data products by merging multiple models. *Journal of Geophysical Research: Atmospheres*. **117**, D22107 (2012).
23. Yao, Y. *et al.* Bayesian multimodel estimation of global terrestrial latent heat flux from eddy covariance, meteorological, and satellite observations. *Journal of Geophysical Research: Atmospheres*. **119**, 4521–4545 (2014).
24. Shi, Q. & Liang, S. Characterizing the surface radiation budget over the Tibetan Plateau with ground-measured, reanalysis, and remote sensing data sets: 1. Methodology. *Journal of Geophysical Research: Atmospheres*. **118**, 9642–9657 (2013).
25. Shi, L., Liang, S., Cheng, J. & Zhang, Q. Integrating ASTER and GLASS broadband emissivity products using a multi-resolution Kalman filter. *International Journal of Digital Earth*. **9**, 1098–1116 (2016).
26. He, T., Liang, S., Wang, D., Shuai, Y. & Yu, Y. Fusion of Satellite Land Surface Albedo Products Across Scales Using a Multiresolution Tree Method in the North Central United States. *IEEE Transactions on Geoscience and Remote Sensing*. **52**, 3428–3439 (2014).
27. Bourlès, B. *et al.* The Pirata Program: History, accomplishments, and future directions. *Bulletin of the American Meteorological Society*. **89**, 1111–1126 (2008).
28. McPhaden, M. J. *et al.* The Tropical Ocean-Global Atmosphere observing system: A decade of progress. *Journal of Geophysical Research: Oceans*. **103**, 14169–14240 (1998).
29. McPhaden, M. J. *et al.* RAMA: The Research Moored Array for African-Asian-Australian Monsoon Analysis and Prediction\*. *Bulletin of the American Meteorological Society*. **90**, 459–480 (2009).
30. Medovaya, M. Assessing ocean buoy shortwave observations using clear-sky model calculations. *Journal of Geophysical Research*. **107**(C2) (2002).
31. OceanSITES. OceanSITES Data Format Reference Manual NetCDF Conventions and Reference Tables. Version 1.4 July 16, 2020. (2020).
32. Cronin, M. F. & McPhaden, M. J. The upper ocean heat balance in the western equatorial Pacific warm pool during September-December 1992. *Journal of Geophysical Research: Oceans*. **102**, 8533–8553 (1997).
33. Alappattu, D. P. *et al.* Warm layer and cool skin corrections for bulk water temperature measurements for air-sea interaction studies. *Journal of Geophysical Research: Oceans*. **122**, 6470–6481 (2017).
34. Donlon, C. J. *et al.* Toward improved validation of satellite sea surface skin temperature measurements for climate research. *Journal of Climate*. **15**, 353–369 (2002).
35. Vanhellemont, Q. Automated water surface temperature retrieval from Landsat 8/TIRS. *Remote Sensing of Environment*. **237**, 111518 (2020).
36. Qin, H., Kawamura, H., Sakaida, F. & Ando, K. A case study of the tropical Hot Event in November 2006 (HE0611) using a geostationary meteorological satellite and the TAO/TRITON mooring array. *Journal of Geophysical Research*. **113**, C08045 (2008).
37. Tošić, I. *et al.* Annual and seasonal variability of precipitation in Vojvodina, Serbia. *Theoretical and Applied Climatology*. **117**, 331–341 (2013).
38. Jiang, C. Evaluation of a hybrid satellite- and NWP-based turbulent heat flux product using Tropical Atmosphere-Ocean (TAO) buoys. *Journal of Geophysical Research*. **110**, C09007 (2005).
39. Pinker, R. T., Bentamy, A., Katsaros, K. B., Ma, Y. & Li, C. Estimates of net heat fluxes over the Atlantic Ocean. *Journal of Geophysical Research: Oceans*. **119**, 410–427 (2014).
40. Zhou, X., Ray, P., Boykin, K., Barrett, B. S. & Hsu, P.-C. Evaluation of Surface Radiative Fluxes over the Tropical Oceans in AMIP Simulations. *Atmosphere*. **10** (2019).
41. Karnieli, A. *et al.* Use of NDVI and Land Surface Temperature for Drought Assessment: Merits and Limitations. *Journal of Climate*. **23**, 618–633 (2010).
42. Feng, Y., Liu, Q., Qu, Y. & Liang, S. Estimation of the Ocean Water Albedo From Remote Sensing and Meteorological Reanalysis Data. *IEEE Transactions on Geoscience and Remote Sensing*. **54**, 850–868 (2016).
43. Wang, S., Trishchenko, A. P., Khlopenkov, K. V. & Davidson, A. Comparison of International Panel on Climate Change Fourth Assessment Report climate model simulations of surface albedo with satellite products over northern latitudes. *Journal of Geophysical Research*. **111**, D21108 (2006).
44. Cheng, J. *et al.* A Lookup Table-Based Method for Estimating Sea Surface Hemispherical Broadband Emissivity Values (8–13.5 μm). *Remote Sensing*. **9**, 245 (2017).
45. Cheng, J., Cheng, X., Meng, X. & Zhou, G. A Monte Carlo Emissivity Model for Wind-Roughened Sea Surface. *Sensors (Basel)*. **19**, 2166- (2019).
46. Rutan, D. A. *et al.* CERES Synoptic Product: Methodology and Validation of Surface Radiant Flux. *Journal of Atmospheric and Oceanic Technology*. **32**, 1121–1143 (2015).
47. Zhang, Y. C., Rossow, W. B. & Lacis, A. A. Calculation of surface and top-of-atmosphere radiative fluxes from physical quantities based on ISCCP datasets: 1. Method and sensitivity to input data uncertainties. *Journal of Geophysical Research: Atmospheres*. **100**, 1149–1165 (1995).
48. Zhang, Y. C., Rossow, W. B., Lacis, A. A., Oinas, V. & Mishchenko, M. I. Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data. *Journal of Geophysical Research*. **109**, D19105 (2004).
49. Pinker, R. & Laszlo, I. Modeling surface solar irradiance for satellite applications on a global scale. *Journal of Applied Meteorology*. **31**(2), 194–211 (1992).
50. Fu, Q., Liou, K. N., Cribb, M. C., Charlock, T. P. & Grossman, A. Multiple scattering parameterization in thermal infrared radiative transfer. *Journal of the Atmospheric Sciences*. **54**, 2799–2812 (1997).
51. Decker, M. *et al.* Evaluation of the Reanalysis Products from GSFC, NCEP, and ECMWF Using Flux Tower Observations. *Journal of Climate*. **25**, 1916–1944 (2012).
52. Hans, H. *et al.* Global reanalysis: goodbye ERA-Interim, hello ERA5. *ECMWF Newsletter*. <https://doi.org/10.21957/vf291hehd7> (2019).
53. Kobayashi, S. *et al.* The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *Journal of the Meteorological Society of Japan. Ser. II*. **93**, 5–48 (2015).
54. Gelaro, R. *et al.* The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*. **30**, 5419–5454 (2017).

55. Kanamitsu, M. *et al.* Ncep-Doe Amip-Ii Reanalysis (R-2). *Bulletin of the American Meteorological Society*. **83**, 1631–1643 (2002).
56. Jia, A., Liang, S., Jiang, B., Zhang, X. & Wang, G. Comprehensive Assessment of Global Surface Net Radiation Products and Uncertainty Analysis. *Journal of Geophysical Research: Atmospheres*. **123**, 1970–1989 (2018).
57. Pinker, R. T., Zhang, B., Weller, R. A. & Chen, W. Evaluating Surface Radiation Fluxes Observed From Satellites in the Southeastern Pacific Ocean. *Geophysical Research Letters*. **45**, 2404–2412 (2018).
58. Han, D. Comparison of Commonly Used Image Interpolation Methods *International Conference on Computer Science and Electronics Engineering*. 1556–1559 (2013).
59. Parker, J., Kenyon, R. V. & Troxel, D. E. Comparison of interpolating methods for image resampling. *IEEE Transactions on Medical Imaging*. **2**, 31–39 (1983).
60. Salomon, J. *et al.* Global land water mask derived from MODIS nadir BRDF-adjusted reflectances (NBAR) and the MODIS land cover algorithm. **1** 241 (2004).
61. Hurlburt, H. E., Wallcraft, A. J. & Kara, A. B. A Correction for Land Contamination of Atmospheric Variables near Land–Sea Boundaries\*. *Journal of Physical Oceanography*. **37**, 803–818 (2007).
62. Wallcraft, A. J. *et al.* Comparisons of monthly mean 10 m wind speeds from satellites and NWP products over the global ocean. *Journal of Geophysical Research*. **114**, D16109 (2009).
63. Haza, A. C., Özgökmen, T. M., Griffa, A., Garraffo, Z. D. & Piterbarg, L. Parameterization of particle transport at submesoscales in the Gulf Stream region using Lagrangian subgridscale models. *Ocean Modelling*. **42**, 31–49 (2012).
64. Doelling, D. R. *et al.* Advances in Geostationary-Derived Longwave Fluxes for the CERES Synoptic (SYN1deg) Product. *Journal of Atmospheric and Oceanic Technology*. **33**, 503–521 (2016).
65. Geisser, S. The Predictive Sample Reuse Method with Applications. *Publications of the American Statistical Association*. **70**, 320–328 (1975).
66. Liang, H. & Jiang, B. Global Ocean Surface net radiation, 1982–2020. *CASEarth* <https://doi.org/10.12237/casearth.622ef74f819aec386070b4d5> (2022).
67. Minnis, P. *et al.* Radiative climate forcing by the mount pinatubo eruption. *Science*. **259**, 1411–1415 (1993).
68. Xu, J. *et al.* Generating a High-Resolution Time-Series Ocean Surface Net Radiation Product by Downscaling J-OFURO3. *IEEE Transactions on Geoscience and Remote Sensing*. **59**, 2794–2809 (2021).
69. Arora, A. *et al.* Role of Indian Ocean SST variability on the recent global warming hiatus. *Global and Planetary Change*. **143**, 21–30 (2016).
70. Moffa-Sanchez, P., Rosenthal, Y., Babila, T. L., Mohtadi, M. & Zhang, X. Temperature Evolution of the Indo-Pacific Warm Pool Over the Holocene and the Last Deglaciation. *Paleoceanography and Paleoclimatology*. **34**, 1107–1123 (2019).
71. He, T. *et al.* Greenland surface albedo changes in July 1981–2012 from satellite observations. *Environmental Research Letters*. **8**, 044043 (2013).
72. Kim, H. & An, S.-I. On the subarctic North Atlantic cooling due to global warming. *Theoretical and Applied Climatology*. **114**, 9–19 (2012).
73. Wen, Q., Yao, J., Döös, K. & Yang, H. Decoding Hosing and Heating Effects on Global Temperature and Meridional Circulations in a Warming Climate. *Journal of Climate*. **31**, 9605–9623 (2018).
74. Che, H., Stier, P., Gordon, H., Watson-Parris, D. & Deaconu, L.-T. *The significant role of biomass burning aerosols in clouds and radiation in the South-eastern Atlantic Ocean*. (2020).
75. Shi, Y., Zang, S., Matsunaga, T. & Yamaguchi, Y. A multi-year and high-resolution inventory of biomass burning emissions in tropical continents from 2001–2017 based on satellite observations. *Journal of Cleaner Production*. **270**, 122511 (2020).
76. Chen, X. & Tung, K. K. Global surface warming enhanced by weak Atlantic overturning circulation. *Nature*. **559**, 387–391 (2018).
77. Pinker, R. T., Grodsky, S. A., Zhang, B., Busalacchi, A. & Chen, W. ENSO impact on surface radiative fluxes as observed from space. *Journal of Geophysical Research: Oceans*. **122**, 7880–7896 (2017).
78. Pavlakis, K. G. *et al.* ENSO surface longwave radiation forcing over the tropical Pacific. *Atmospheric Chemistry and Physics*. **7**, 2013–2026 (2007).
79. Pavlakis, K. G., Hatzianastassiou, N., Matsoukas, C., Fotiadi, A. & Vardavas, I. ENSO surface shortwave radiation forcing over the tropical Pacific. *Atmospheric Chemistry and Physics*. **8**, 5565–5577 (2008).
80. Liang, H. CN-NADC *In situ* Ocean surface Rn data and production code. *figshare* <https://doi.org/10.6084/m9.figshare.19372823.v4> (2022).

## Acknowledgements

This work was funded by the Natural Science Foundation of China (Grant No. 41971291) and the Chinese Grand Research Program on Climate Change and Responses (Grant No. 2016YFA0600101). All data used in this paper are available freely online, as mentioned above. We thank NASA, Nagoya University, GMAO, ECMWF, JMA, JAXA, WCRP, and NCAR of NCEP2 for providing the remotely sensed and reanalysis data. The *in situ* measurements were provided by the Upper Ocean Processes Group (UOP), the GTMBA Project Office of NOAA/PMEL, the International OceanSITES project, and the Chinese National Arctic and Antarctic Center (CN-NADC). We thank the FAO for providing the divisional basis for the classification of oceans.

## Author contributions

B.J. and L.S.L. conceived the research. H.L. and B.J. designed the experiments and drafted the manuscript. H.L. performed the experiments and generated this product. H.L., J.P., P.S.L., J.H., X.Y. processed the input data, J.P., P.S.L., J.H., X.Y. provided code support. All authors discussed and revised the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to B.J.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.





**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022