Dear Reviewers and Editor,

This research was done with the support of China's National Key R&D Program. The purpose of the project is to build a long-term series of global major meteorological disaster remote sensing data sets with high spatio-temporal and consistency based on the current global multi-source remote sensing data and ground observation site data, and to provide key ocean temperature parameters (such as sea surface temperature) for marine meteorological disaster forecasting models, especially rapid forecasts of marine disasters such as typhoons, and provide early warning services for global fishing vessels and merchant ships.

We would like to express our sincere appreciation to Anonymous Referee #1 for his/her comprehensive comments and valuable suggestions on our manuscript which are very important to improve the quality of our manuscript. All the comments made by Anonymous Referee #1 have been carefully and individually addressed. The reply is in supplement.

 # Summary:

This is a review of “A New Global Gridded Sea Surface Temperature Data Product Based on Multisource Data” by Mengmeng Cao, Kebiao Mao, Yibo Yan, Jiancheng Shi, Han Wang, Tongren Xu, Shu Fang, and Zijin Yuan. The authors have merged the sea surface temperature (SST) data from multiple sources to create a new high-resolution global dataset of monthly SST. The new data product does not contain any missing values and has been shown to be more accurate than the unmerged datasets. However, I have some concerns about the usefulness of the new dataset and the assessment of accuracy (see general comments). Moreover, the authors may need to emphasize the novelty and the uniqueness of the methods used to create the new data set.

Response: We would like to thank you for reviewing our manuscript. Your comments and good suggestions are very important for us to improve the quality of manuscript and dataset. We have carefully addressed all the issues raised by you and the response is presented below.

At present, there are three main methods for obtaining ocean temperature: The first is the traditional method, which obtains sea surface temperature through sea observation sites. The main advantage of this method is that it has continuity in time and is hardly affected by weather. The disadvantage is that the number of observation sites is limited and the space lacks continuity, especially in remote sea areas. The second method is to obtain sea surface temperature through remote sensing retrieval. Remote sensing has advantages in space, but lacks continuity in time. Remote sensing is divided into two inversion methods: thermal infrared remote sensing and passive microwave remote sensing. Thermal infrared inversion of sea surface temperature has a high accuracy and resolution, but it has a great influence on clouds. There are 60% of the area covered by clouds every day, so there are 60% of the area missing values. Although passive microwave is less affected by clouds, the resolution is relatively low. Passive microwave remote sensing is affected by the land near the coast, and the accuracy of sea surface temperature inversion is not high. The third is to output sea temperature products through the assimilation model. This method relies on the accuracy of the input parameters of the assimilation model.

Although different methods are used to obtain ocean surface temperature, they actually represent temperature information at different ocean depths, and the observation time is also inconsistent. The sea temperature observed by traditional sites is deeper than the temperature observed by remote sensing. Even if they are all the temperatures retrieved from remote sensing, the temperatures retrieved from thermal infrared and microwave are from different ocean depths. The sea temperature observed by thermal infrared is the skin temperature, and the sea temperature observed by microwave is a bit deeper than the depth observed by thermal infrared. The sea surface temperature obtained by the assimilation model should also be different.

Thermal infrared remote sensing is currently recognized as the most accurate method for obtaining sea surface temperature in a large area. Therefore, thermal infrared remote sensing is usually used to obtain sea surface temperature. For the ocean, the effective sea surface temperature value obtained through thermal infrared remote sensing every day is less than 40% of the total area (as shown in Figure 1), which means that 60% of the daily data have no value every day.

When calculating the monthly average data, some data sets use the average temperature value obtained by dividing the effective days of a month by the effective days. For example, if a certain pixel has only 25 effective temperature values in a certain month, then the average temperature of this month is calculated by using the average value of these 25 valid days. Although the monthly average temperature map calculated in this way has few missing values, the average temperature error of some pixels is relatively large (as shown in Figure 2).

The highlight of our work is to make full use of the research foundation of the predecessors to traverse the MODIS ocean temperature data set to find the pixels with low quality data, and then use high-quality daily data and other multi-source data (including sea surface temperature retrieved from passive microwave and observation site data, etc.) to improve the accuracy of the data, including pixels with low data quality and missing pixel data.

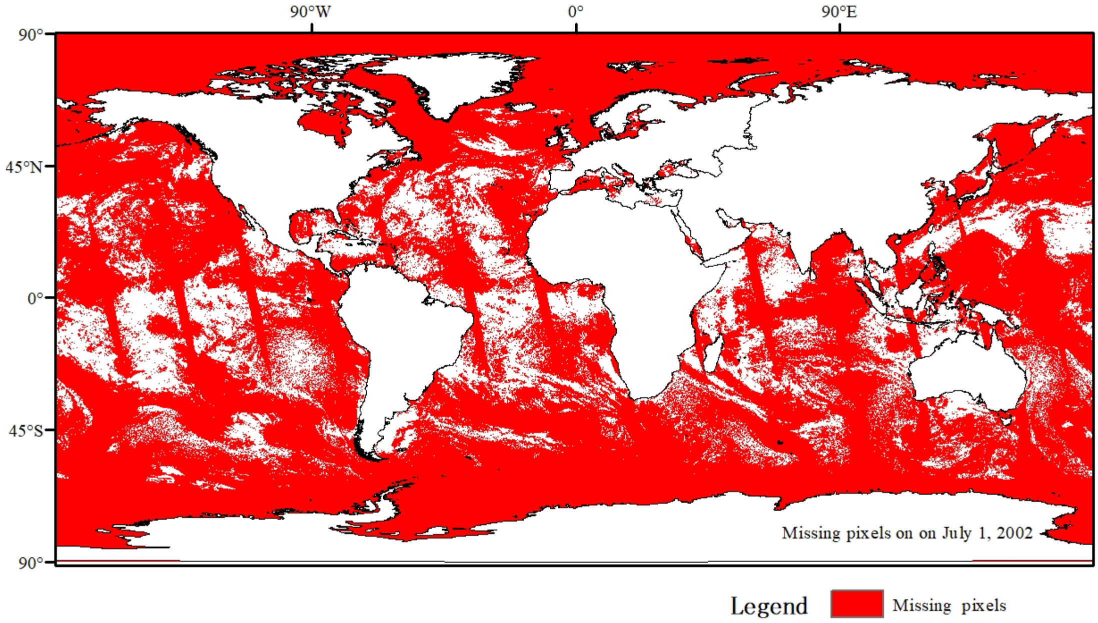

Figure 1. Distribution map of missing pixel values of thermal infrared remote sensing daily data (July 1, 2002).

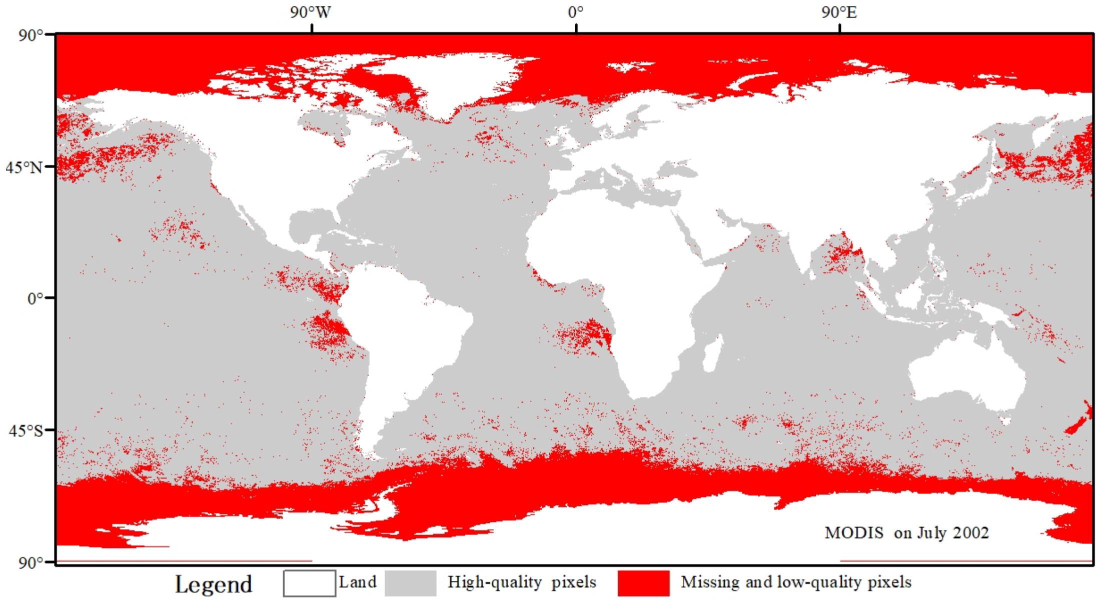

Figure 2. Distribution map of missing and low quality pixel of thermal infrared remote sensing monthly data (July, 2002).

General comments:

1. The conclusion of the better performance of the new dataset is drawn based on the smaller differences between the new data set and the in-situ observations (iQuam) than between the original datasets and iQuam. Should it simply because the iQuam data is used to create the new dataset? If iQuam is supposed to be the closest to the truth (as it is used as reference to validate other datasets), then why researcher do not use iQuam but use the new dataset created by the authors? I think author needs to better justify it. Missing values generally do not have a significant impact on the statistical analysis of

climate. Moreover, the missing values can be obtained by interpolating other data using simple interpolation methods without losing much accuracy.

Response: Thank you for the valuable comment. It may be that we did not express it clearly in the manuscript.

We don't just use the in-situ observations (iQuam) data. In order to improve the low-quality and missing pixel values, we mainly use the sea surface temperature retrieved from microwave data, the sea surface temperature retrieved from AVHRR, and in-situ observations (iQuam) data, to fill in the relevant data by correcting consistency for time and space, and then part of in-situ observations (iQuam) data are used for verification. The purpose of the work is to build a long-term series of global major meteorological disaster remote sensing data sets with high spatio-temporal and consistency based on the current global multi-source remote sensing data and ground observation site data, and to provide key ocean temperature parameters (such as sea surface temperature) for marine meteorological disaster forecasting models, especially rapid forecasts of marine disasters such as typhoons, and provide early warning services for global fishing vessels and merchant ships.

It seems that there are not many data with low data quality in Figure 2 due to computer display. In fact, there are still many missing values for local areas. If we make up for vacancies or low-quality pixels through interpolation, the local accuracy needs to be improved. Although the data obtained by interpolation may have little impact on global research, it is not enough for local time-space analysis, local weather forecasting and disaster prediction with high accuracy.

2. I wonder how accurate the new dataset is compared with global reanalysis such as ERA5. The global reanalysis may not have such a high-resolution as the new dataset. But is it possible to compare the new dataset with some regional high-resolution reanalysis? The reanalysis is created using both information from observations and model simulations. Therefore, should one expect a reanalysis product to be more accurate than the new dataset generated using only observations?

Response: Thank you for your comments.
ERA5 provides hourly estimates of a large number of atmospheric, land and oceanic climate variables, which cover the Earth on a 30km grid. Although ERA5 claims that uncertainties for all variables will be included in the data information, the uncertainties for some variables cannot be downloaded yet, so it is difficult for us to make a direct evaluation. However, we have known from our research team's analysis for reanalysis product in some regions of China, such as the accuracy of air temperature, soil moisture and other products in the western region is not very well, and some products need to be further improved. We think the main reason is that the observation sites in these places are relatively sparse, resulting in limited data provided to the assimilation model. As a result, the accuracy of assimilation products in some regions with insufficient observation data is not high.

The assimilation model is not a panacea. The first one requires a lot of input parameters. When some input parameters are missing or the accuracy is insufficient in some regions, the accuracy of some output parameters may be not well. In addition, the estimation of each parameter in the assimilation model uses some parameter estimation calculation models. These estimation models are not suitable for all situations. Different regions in the world are different in terrain and climate conditions. Especially when some abnormal situations occur, such as typhoons or heavy rains, the results estimated by the assimilation model may have a large deviation. Therefore, many different assimilation models have been developed. These models have their own advantages and disadvantages, and the accuracy is different in different places. In the case of missing input parameters or insufficient precision, such as thick clouds and heavy precipitation, the error will be relatively large. Therefore, many researchers have been constantly improving these models, including ERA5, which is also constantly evolving and improving (https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5).

US NOAA also produces ocean temperature products with a resolution of 30 km, which should be one of the most authoritative products, and they are updated and released every day. We use this product as a reference comparison. CAAS SST with a resolution of 4 km made by us are resampled based on the resolution of ERA5. Figure 3 is the distribution map of OIS SST obtained from NOAA in July, 2002. Figure 4 is distribution map of CAAS SST (July, 2002). Figure 5 is distribution map of ERA5 SST with a resolution of 30 km obtained from ECMWF (July, 2002). Figure 6 is a distribution map of the difference between CAAS SST and NOAA SST. Figure 7 is the difference between ERA5 and NOAA.

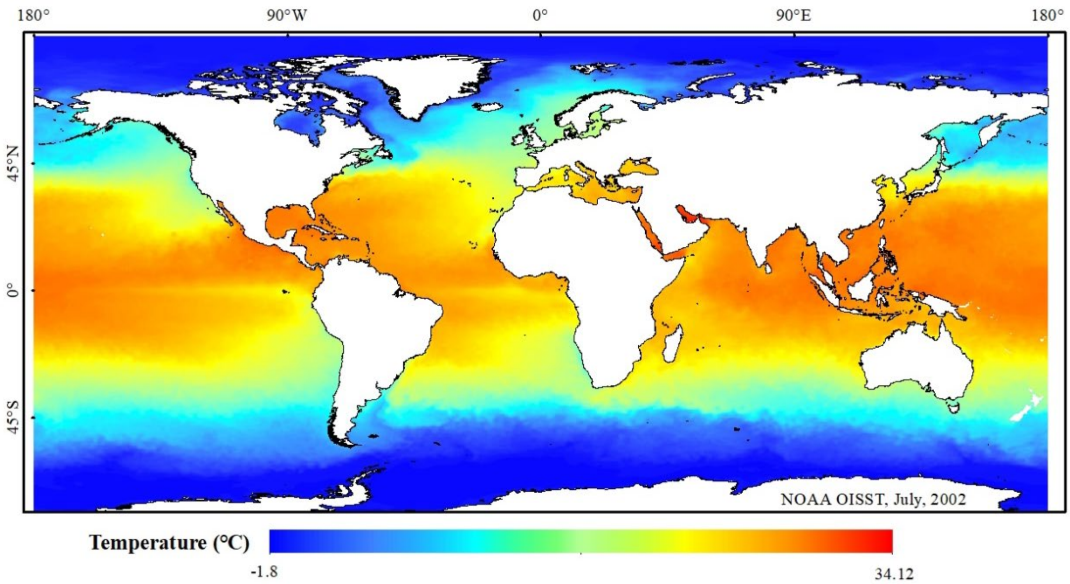

Figure 3. Distribution map of OIS SST obtained from NOAA (July, 2002).

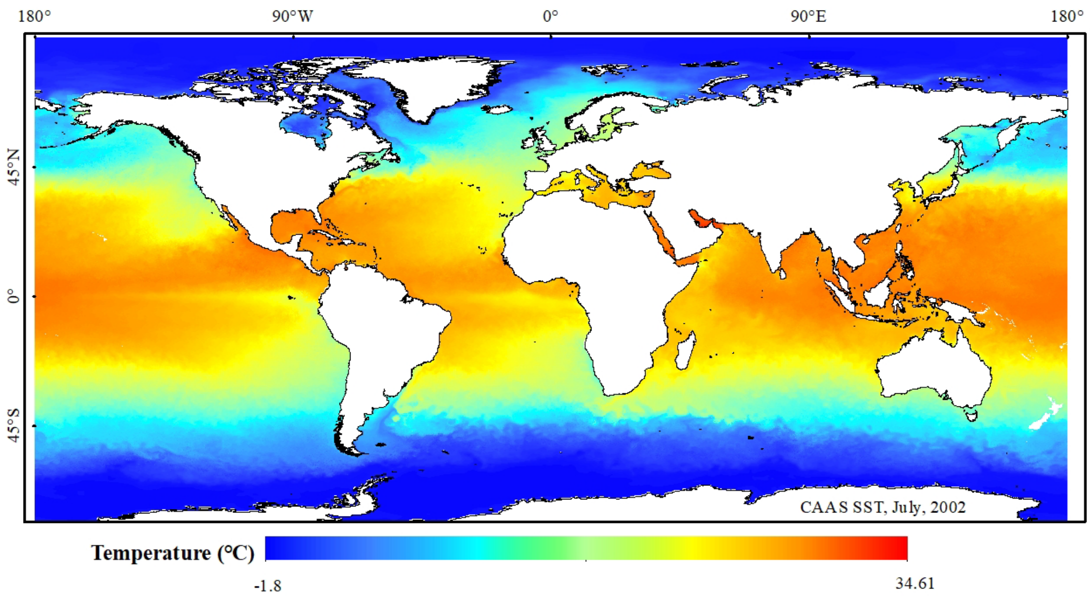

Figure 4. Distribution map of CAAS SST made by us (July, 2002).

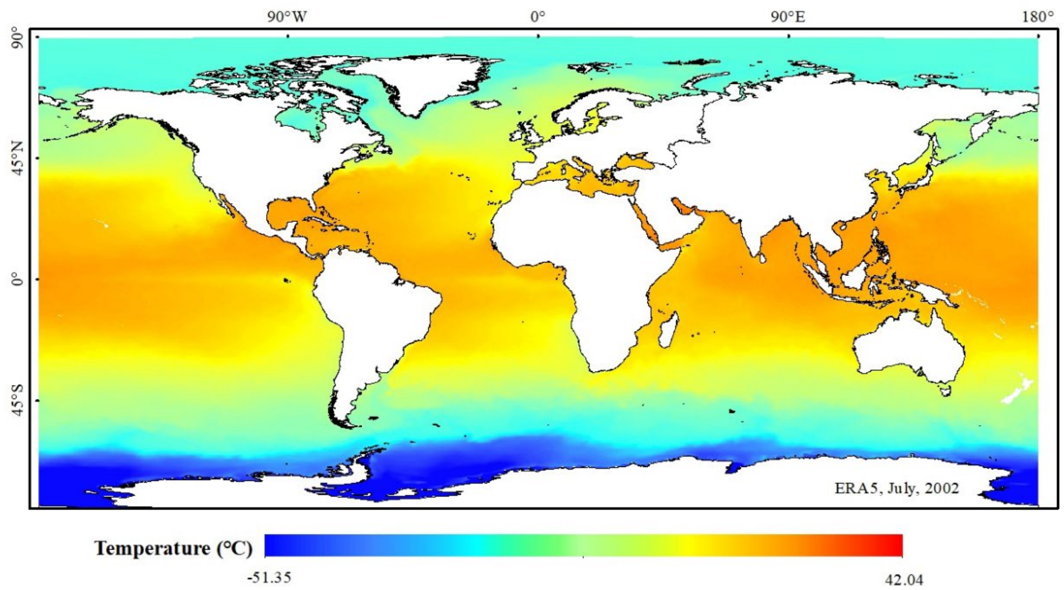

Figure 5. Distribution map of ERA5 SST obtained from ECMWF(July, 2002).

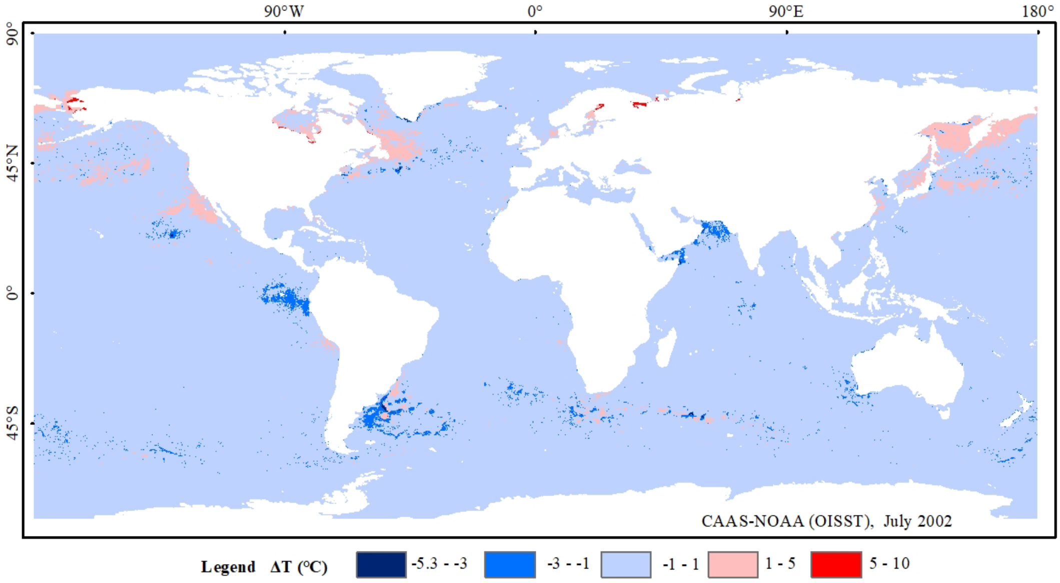

Figure 6. Distribution map of the difference between CAAS SST and NOAA SST

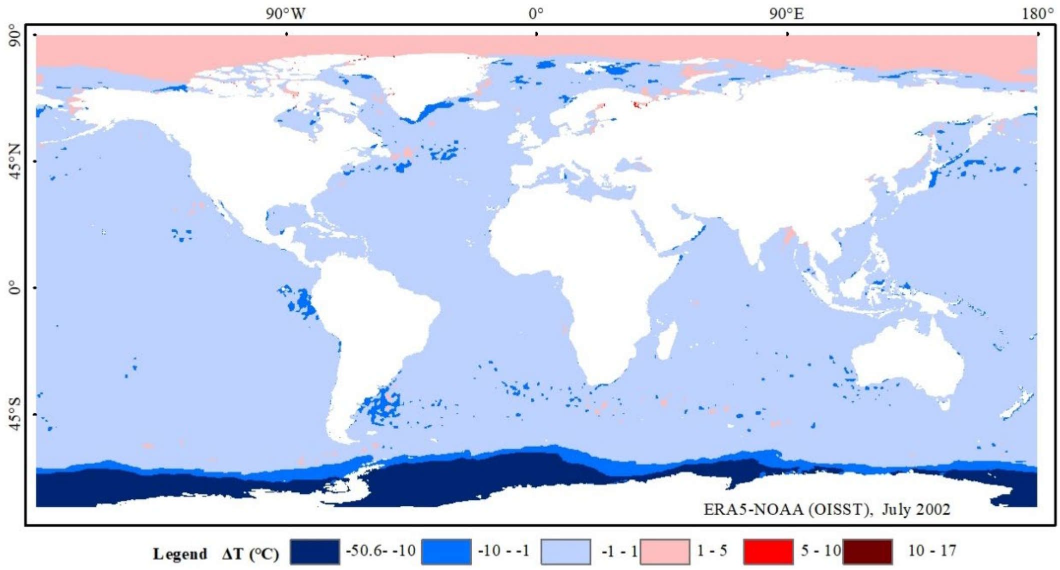

Figure 7. Distribution map of the difference between ERA5 SST and NOAA SST

Shown from figure 3-5, on the whole, the distribution trend of CAAS SST and NOAA SST ocean temperature products is not much different, and the distribution range of ERA5 temperature products is larger than that of CAAS SST and NOAA SST. We took NOAA SST ocean temperature products as a reference and made a distribution map of the difference between ERA5 and CASS SST and NOAA SST. Shown from figure 6 and 7. The regions with the largest differences are mainly distributed in the Antarctic and Arctic regions, as well as the regions with low MODIS product quality and no pixel values, which can be seen in Figure 2.

The difference between ERA5 sea temperature products and NOAA products in the Antarctic and Arctic regions is relatively large. The main reason is that there are less observation data in the polar regions, especially the low resolution of ERA5, and the polar region itself is not large. Due to the influence of mixed pixels, in places where the sea and land meet, especially where ice and water are mixed, the temperature of the polar coast or the ice surface is used as the pixel temperature as the pixel temperature value. We know that if the temperature of the ocean is too low, it will freeze. The resolution of remote sensing data is relatively high, and the ice along the coast of the polar regions can be distinguished. Therefore, at this point, high-resolution remote sensing observation data still has advantages. For other areas with large differences, we used microwave and AVHRR and site data to improve low-quality and invalid pixels. So theoretically, our accuracy should be slightly higher.

Re-analyzed data such as ERA5 is a data set (about 271 global products) produced by the assimilation model considering the global multi-factors, which has an overall advantage. However, when a single parameter or special parameter has a high spatiotemporal observation, the accuracy of the ERA5 product is not necessarily higher than that of the inversion data product. Especially in the local area when there is not enough input parameter information or high-precision parameter to input assimilation model, the accuracy of the observation data inversion should be higher than that of the assimilation model.

In addition, the advantage of our data set can be used for local research and analysis. The 30-kilometer resolution of assimilated data products cannot meet the requirements of high spatial resolution in local areas. Especially in offshore weather forecasting and analysis of near-shore agricultural disasters, we need to further improve the spatial resolution. This is also one of the main reasons why many countries have been continuously increasing ocean ground observation sites and improving the spatial and temporal resolution of satellite observations. Different data sets are not contradictory, and we can absorb each other's advantages, and further improve the accuracy of monitoring and forecasting.

 # Specific comments:

1. Line 109-110: Why need to correct the observation time difference if the new dataset is monthly data. Isn’t the difference measured in days?

Response: Thank you for the valuable comment. The observation time of different sensors is different, and the calculation must be corrected to the same time to be comparable. For example, when the satellite transits at 10 o'clock, the observed sea temperature data is the temperature at 10 o'clock. The data of the sea surface observation station is 10:30, so we must calibrate the observed sea temperature at 10:30 to 10 o'clock and then calculate or compare. The calculated average must also be consistent in time to be comparable.

2. Section 2.4: Suggest separating the section of data and method.

Response: Thank you for your good suggestions. We have tried to make revisions.

3. Line 224: Figure caption in wrong place.

Response: Thank you very much for your advice. We have modified it in the manuscript.

4. Figure 6: The scatter points for B looks strange (different behavior than the pdf and the box plot). Are the box plots already enough to demonstrate what you want show?

Response: Thank you very much for your careful review. When we made this figure, there were actually more points than shown on the graph (as shown in Figure 8).Take H with relatively few scatter points as an example. In fact, the figure exported is shown in the lower left corner of the figure below. Because the figure in this form is not beautiful enough, our manuscript directly uses the unexported figure in the software, which may cause some differences. We have modified the figure in the manuscript.

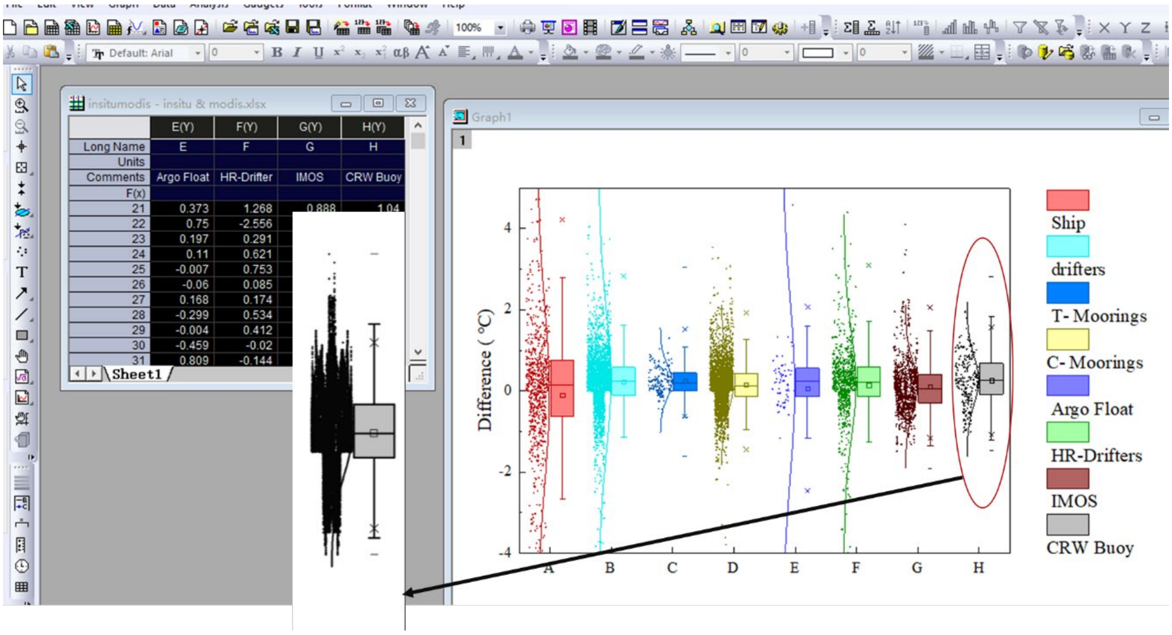

Figure 8.

5. Eq. (8): Suggest using bold font for vector and matrix.

Response: Thanks for your good suggestion. We have modified it in the manuscript.

6. Eq. (9): X_{(t)} looks a bit uncommon, suggest X_{t} or X(t). But it is the author's choice.

Response: Thanks for your good suggestion. We have modified it in the manuscript.

7. Eq. (10): I don’t really understand this equation, if Eq. (10) holds, then Eq. (11) shouldn’t be correct.

Response: Thanks for your good suggestion. We have modified it in the manuscript.

8. Eq. (12): Should it be (P+R). Additionally, they are matrices and should be write in such a way P(P+R)^{-1}.

Response: Thanks a lot for pointing these out. We are sorry for our unclear expression and carelessness. We have modified it in the manuscript.

9. Fig. 15: the original SST data seems quite interesting, as it has a lower bound of around 9 Celsius.

Response: Thank you for the valuable comment. Most of the global ocean observing station data are distributed in low and middle latitudes (shown as figure 9), and relatively few in high latitudes. Data is very valuable. When doing verification, we mainly choose the site data in the low and middle latitudes, so there are not many low values. The biggest advantage of our data set is the high spatial resolution.

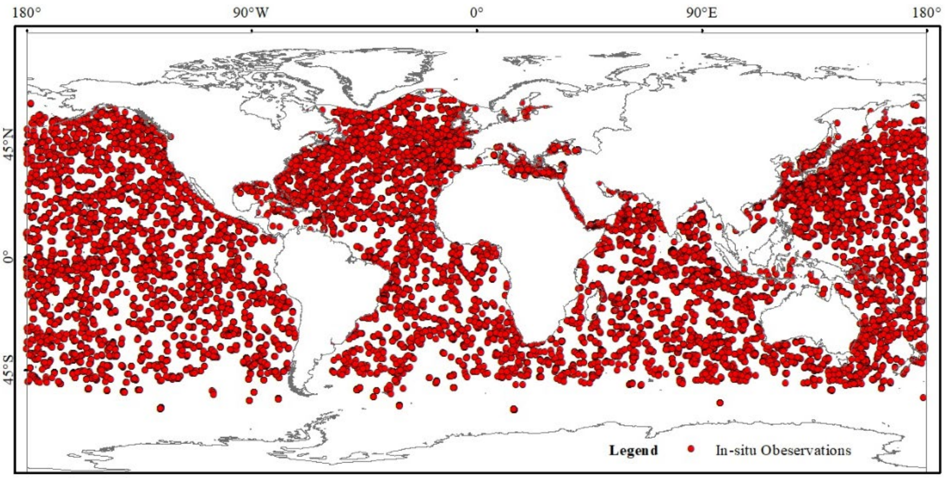

Figure 9. Distribution map of in-situ observation (iQuam) stations.