Responses to Reviewer #1

Thanks to the reviewer for the critical and very constructive comments. We've done our best to find the advantages of our developed methods by comparing them with the previous methods.

In this manuscript, the authors compare the ability of several regression techniques to recover underlying long-term trends in daily SST (Sea Surface Temperature) records. They find that the ordinary least squares (OLS) method could be biased due to an imbalanced phase associated with seasonal cycles. They further propose a 7-step approach to account for imbalanced seasonal cycles to obtain a less-biased estimator. Despite the interesting technical discussion, the manuscript, in my view, does not provide sufficient methodological advances. The atmospheric and climate community already has other approaches to address this issue, potentially more efficiently (please refer to my second comment below).

Thanks for the comments. As suggested, we've compared our method with two commonly used methods in the atmospheric and climate community in detail, which will be illustrated below.

Moreover, I am concerned that the topic of this paper might not fall within the scope of this journal, which focuses on Natural Hazards. As a result, I would suggest that the authors compare their methods to the approach I suggest below and consider submitting this work to a more technical journal.
We had extensively searched for an appropriate journal for this study. After careful consideration, we have identified NHESS as a suitable option, as it welcomes manuscripts focusing on methodological advancements that contribute to addressing natural hazards, including issues related to global warming. Hence, we are motivated to submit our work to NHESS.

1. I appreciate the authors' careful introduction of OLS2, GMR, and OR, but this discussion may not be entirely relevant in this context because the timing of SST measurements should be well-known. The authors also point this out in line 115. On the other hand, the STL (Seasonal-Trend Decomposition procedure based on Loess) probably deserves a more detailed introduction, including the mathematics and equations.

The details of STL have been described in Cleveland et al. (1990). The conceptual *description has been provided in L102-109. As the methodology of STL involves numerous trivial and detailed operations, we don't think it is suitable to incorporate into the present manuscript. We've cited the paper of Cleveland et al. (1990) for readers interested in the details.*

2. In atmospheric, ocean, and climate research, the first step of an analysis typically involves removing the seasonal cycle. Long-term trends are then estimated from the anomalies. When using daily data, directly calculating daily climatology often results in noisy estimates. Hence, the community uses sine and cosine functions to fit the amplitude and shape of seasonal cycles. For a problem that also estimates long-term trends, the model would be:

$$T = \mu + k*yr + \sum_{i=1}^{N}(a_i * \sin(i*yr*2\pi)) + \sum_{i=1}^{N}(b_i * \cos(i*yr*2\pi)),$$

where the goal is to fit for μ, k, a_i, and b_i from the data, with $i = 1, 2, ..., N$. In practice, this is simply a multi-linear regression, and N can be determined if increasing N does not further improve the fitting (using, for example, an F-test). Fitting sine and cosine functions simultaneously captures different phases. Hence, unless the authors demonstrate that their method outperforms the community's common practice, I am not fully convinced that the method they propose would make a significant methodological improvement.

We must emphasize that we do not intend to replace the previous methods but rather propose another way to achieve the long-term trend estimate. After adding a sub-session to compare our method with the previous two, we found that our method is robust and has advantages over the other two in certain aspects. Please see session 5.2 and Figures 7 and 8.

"We have demonstrated our proposed evenization method as a feasible approach to estimate the long-term trend of SST. It is desired to compare our method with the commonly used methods in the climate community. The first method involves computing the daily climatological value of SST using the available SST data. The long-term trend can be estimated by applying OLSR1 to the SST anomalies derived by subtracting the climatological SST from the original SST data. This is expected to lower the bias resulting from seasonality. The second method models the SST data as a combination of linear and sinusoidal functions (e.g., Park et al., 2022):

$$SST(t) = b_0 + b_1 t + A\cos\left(\frac{2\pi t}{T}\right) + B\sin\left(\frac{2\pi t}{T}\right). \qquad (5)$$

The first two terms on the right-hand side are the linear function, representing the long-term trend. The third and fourth terms on the right-hand side represent the seasonal component, where T=365 days is the period. The amplitude of seasonal variations can be obtained as $\sqrt{A^2 + B^2}$*. Here,* $\widehat{b_o}$*,* $\widehat{b_1}$*,* $\hat{A}$*, and* $\hat{B}$ *are obtained using nonlinear least squares fitting the SST dataset. Adding other periodic components, such as interannual variations, may only sometimes be helpful due to the increased number of fitting parameters, which could lower the numerical accuracy and stability.*

The performance of the three methods is evaluated using the 14-year time series, as depicted in Fig. 5a, which is generated using equation (3). The semidiurnal tidal amplitude is increased from 0.2 °C to 0.3 °C to better investigate the impacts of small fluctuations. Fig. 8 shows how the estimated slope changes with different data lengths (3-14 years) used for estimation, allowing for the evaluation of the uncertainty of each method. Overall, as the data length increases, there is a reduction in $\widehat{b_1}$ *uncertainty for both the linear trends of SST anomalies (cyan curve in Fig. 8a) and evenized SST (red curve). In both methods, the uncertainty of* $\widehat{b_1}$ *is significant when the data length is less than seven years, and the deviation could reach 0.01 °C/yr. The deviation is less than 0.003 °C/yr for data length > 7 years.* $\widehat{b_1}$ *obtained from the evenized SST method closely aligns with the correct value (represented by the black dashed line), whereas* $\widehat{b_1}$ *obtained from the SST anomalies method tends to be consistently lower than the correct value. This can be clearly found in the probability density function (PDF) shown in Fig. 8b. The PDF of* $\widehat{b_1}$ *estimated using the evenized SST is unbiased because it concentrates around the correct value (0.1643 °C/yr; the vertical dashed line). In contrast, the SST anomalies estimate (cyan line) is biased because its PDF deviates from the correct value. This suggests our method could be a better estimator.*

When using the combined linear and sinusoidal fitting method, there is no clear relationship between the uncertainty and data length, as depicted in Fig. 8a by the blue line. The PDF shows a more concentrated distribution at $\widehat{b_1}$*=0.1643 °C/yr (blue line in Fig. 8b), suggesting better performance than our method. However, unexpected fitting failure can cause large deviation (blue line in Fig. 8a). We re-examine the PDF using* $\widehat{b_1}$ *with data length > 7 years, which is generally applicable for long-term trend estimates (Fig. 8c). The method of SST anomalies remains biased. The methods of linear and sinusoidal fitting are unbiased, and the peak value of PDF slightly increases from 0.79 to 0.81. Similarly, the methods of evenized SST are unbiased, but the peak value of PDF significantly increases from 0.32 to 0.5. To summarize, our proposed method is unbiased and better than the conventional SST anomalies method. While our method may have a more significant degree of* *uncertainty than linear and sinusoidal fitting, this uncertainty remains within an acceptable range. Furthermore, linear and sinusoidal fittings can be unstable when applied to natural data containing significant noise.*

The same examination using the CWA's data for SST anomalies (cyan lines in Fig.7) and combined linear and sinusoidal fitting (blue lines in Fig. 7) methods is carried out. Again, we focus on the comparison in the stable period, six months trimmed time. $\widehat{b_1}$ *obtained using SST anomalies is 0.004-0.015 °C/yr lower than obtained using evenized SST (cyan lines in Fig. 7 and Table 1). The obtained result agrees with the expected outcome based on the simulated data, as shown in Fig. 8. Using the combined linear and sinusoidal fitting (blue lines), the obtained* $\widehat{b_1}$ *roughly aligns with SST anomalies. The result differs from the simulated data. It is anticipated that complex and diverse natural signals could have interfered with the fitting results, often considered noise. Indeed, unexpected peaks related to the failed fitting occur for the data in the Chenggong station (Fig. 7c), where the tidal signal (Fig. 2a) is strongest among the three stations. Finally, the slope obtained from linear fitting to the STL nonlinear curve (magenta dashed lines in Fig. 7 and Table 1) is close to the result obtained from evenized SST."*

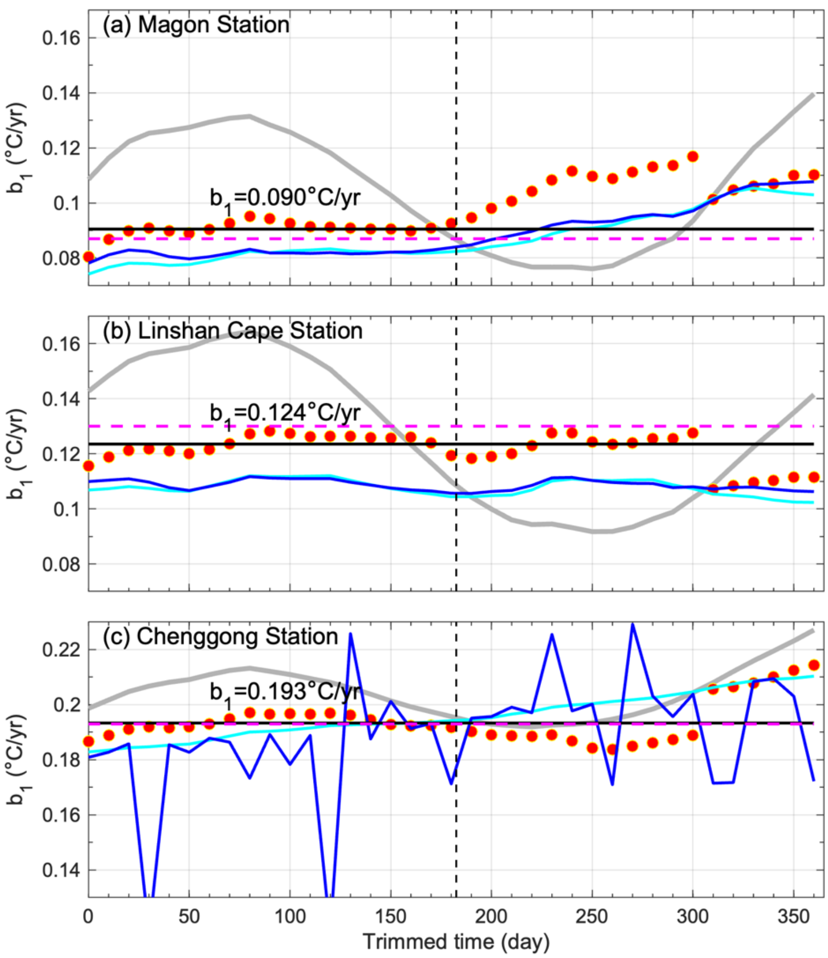

Figure 7: $\widehat{b_1}$ ***as function of trimmed time using the SST data collected in (a) Magong, (b) Linshan Cape, and (c) Chenggong stations. The OSLR1 method and our proposed method (steps 1-7) are represented respectively by the grey curves and red dots. The black lines depict the averaged*** $\widehat{b_1}$ ***values based on corrected data within a trimmed time shorter than*** $T_{sv}/2$***, indicated by the vertical dashed line. The cyan and blue curves are the slope estimated using SST anomalies and combined linear and sinusoidal fitting methods, respectively. The magenta dashed lines are the slope derived from linear fitting to the STL nonlinear curve.***

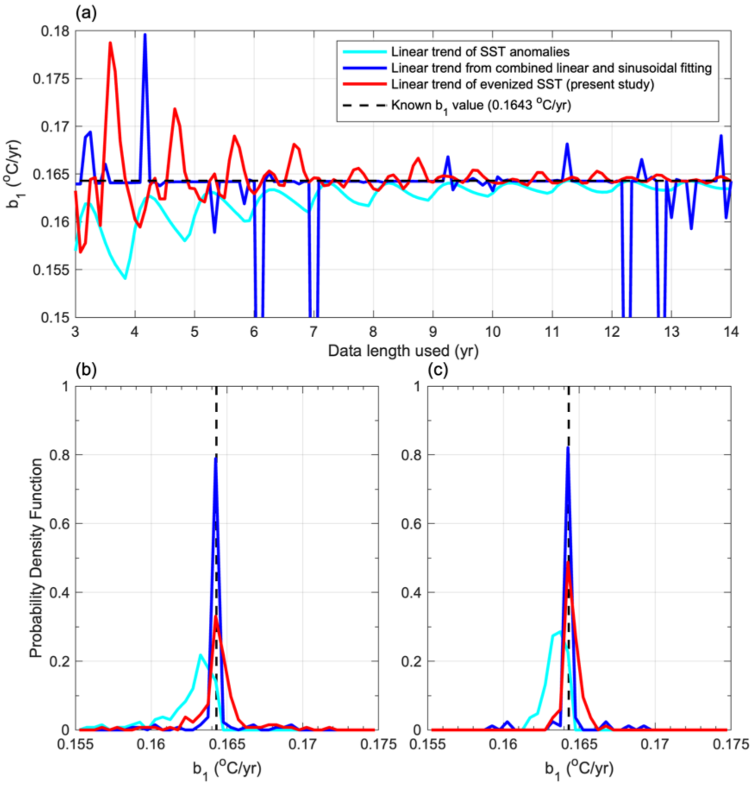

Figure 8: (a) $\widehat{b_1}$ ***as function of data length using the simulated SST data and probability density function of*** $\widehat{b_1}$ ***using data length of (b) 3-14 years and (c) 7-14 years by applying the methods of SST anomalies (cyan lines), combined linear and sinusoidal fitting (blue lines), and evenized SST (red lines). The black dashed lines denote the known*** b_1 ***value.***