Dear editor and reviewers,

Thank you very much for taking the time to provide constructive comments and helpful suggestions on our manuscript. We have addressed the reviewers' concerns, and carefully revised our manuscript. In follows, we would like to address and respond the issues the reviewers concerned and give their detailed descriptions of the changes made in the revised version. All responses are listed below in blue font.

Referee(s)' Comments to Author:

Reviewer: 1

Comments to the Author

The study addresses the dynamical complexity over the tropical Pacific by using the Dynamical System Instantaneous Dimension (DSID) derived from SSTA fields. The authors identified an evident mean shift of DSID around 2007 over the western tropical Pacific, and further explored its linkage with high-frequency variability based on both idealized Lorenz models and realistic observations. Several impact factors, such as noise intensity, persistence, and regional differences among Niño regions, were also examined.

The paper is interesting and well organized, and advance our understanding of the physical interpretation of DSID and offer valuable insights for the study of complex real-world systems. However, some revisions are necessary. It can be recommended for publication after the following issues have been properly addressed. Detailed comments are below.

Response: Thanks for your comments and valuable suggestions.

1. The authors note that dynamical systems (DS) methods can characterize the dynamical properties of systems. However, they should acknowledge that alternative approaches—such as the nonlinear local Lyapunov exponent (NLLE) and attractor radius (AR) methods—are equally effective and have been widely adopted in the literature. The introduction would benefit from a comprehensive comparison of these methodologies.

Reference:

Li, X., Ding, R. & Li, J, 2023: The BaSIC method: a new approach to quantitatively assessing the local predictability of extreme weather events. Clim Dyn. 60, 3561 3576.

Li, X., Ding, R., 2023: The backward nonlinear local Lyapunov exponent and its application to quantifying the local predictability of extreme high-temperature events. Clim Dyn. 60, 2767–2781.

Response : Thanks very much for your valuable suggestion. In the revised manuscript, we have included a discussion and comparison of the nonlinear local Lyapunov exponent (NLLE), attractor radius (AR), and dynamical systems (DS) methods in the introduction, as follows:

"Several methods have been proposed to characterize the dynamical properties of systems. Ding and Li (2007) proposed the nonlinear local LE (NLLE) method to study predictability of the dynamical system. This method accounts for nonlinear error growth of a finite size, and has been widely used in studies of atmospheric predictability (Li and Ding, 2011; Feng et al., 2014, 2018; Ding et al., 2015; Hou et al., 2018, 2022; Li et al., 2020a; He et al., 2021). Li et al. (2019) proposed the backward NLLE (BNLLE) method, based on the NLLE method, which is applicable to studies of the predictability of extreme events (Li and Ding, 2023). Li et al. (2017) introduced the attractor radius (AR) and global AR (GAR) to depict the geometric characteristics and average behavior of chaotic systems. These two statistics can quantify the predictability limits of chaotic systems from the geometric characteristics and average behavior of error growth, and have been applied to the predictability of extreme weather or climate events (Li et al., 2023). **However, these methods focus primarily on the predictability of dynamical systems and do not explore the dynamical properties of the system from other perspectives.**

Recently, Lucarini, Faranda, and their collaborators developed a method based on dynamical systems and extreme value theory to characterize the instantaneous dynamical properties of a given system, which is referred to as the dynamical systems (DS) method (Lucarini et al., 2012; Faranda et al., 2017a and 2017b).This method reconstructs the attractor of the dynamical system using the system trajectory $X(t)$, and computes two dynamical parameters—the instantaneous dimension $d(\zeta)$ and the instantaneous persistence parameter $\theta(\zeta)$ — to characterize the instantaneous dynamics of a given climate system. **The instantaneous dimension $d(\zeta)$ measures the degree of divergence of the attractor trajectories in the neighborhood of time $t = \zeta$ in phase space, and is related to the local number of degrees of freedom of the dynamical system. The instantaneous persistence parameter $\theta(\zeta)$, defined as the inverse of the persistence, quantifies the strength of persistence of the attractor in the vicinity of time $t = \zeta$. This method provides an alternative perspective for understanding the behavior of dynamical systems.** Faranda and his collaborators (Faranda et al., 2017a) have interpreted the physical meaning of the DS instantaneous dimension (DSID) and inverse of persistence in some idealized systems, such as Lorenz system (Lorenz, 1963). They found that the instantaneous dimension of the low-dimensional systems can provide a direct way to compute the dimensions of the attractor without embedding (Faranda et al., 2017a), and the calculated minima and maxima of the instantaneous dimension are able to track the extremes of the Lorenz attractor."

The relevant revisions have been highlighted in yellow in the revised manuscript.

Reference:

[1] Ding, R. and Li, J.: Nonlinear finite-time Lyapunov exponent and predictability, Physics Letters A, 364, 396－400, https://doi.org/10.1016/j.physleta.2006.11.094, 2007.

[2] Ding, R., Li, J., Zheng, F., Feng, J., and Liu, D.: Estimating the limit of decadal-scale climate predictability using observational data, Climate Dynamics, 46, 1563－1580, https://doi.org/10.1007/s00382-015-2662-6, 2015.

[3] Feng, J., Ding, R., Liu, D., and Li, J.: The application of nonlinear local Lyapunov vectors to ensemble predictions in Lorenz systems, Journal of the Atmospheric Sciences, 71, 3554－3567, https://doi.org/10.1175/JAS-D-13-0270.1, 2014.

[4] Feng, J., Li, J., Ding, R., and Toth, Z.: Comparison of nonlinear local Lyapunov vectors and bred vectors in estimating the spatial distribution of error growth, Journal of the Atmospheric Sciences, 75, 1073－1087, https://doi.org/10.1175/JAS-D-17-0266.1, 2018.

[5] He, W., Xie, X., Mei, Y., Wan, S., and Zhao, S.: Decreasing predictability as a precursor indicator for abrupt climate change, Climate Dynamics, 56, 3899 – 3908, https://doi.org/10.1007/s00382-021-05676-1, 2021.
[6] Hou, Z., Li, J., Ding, R., Karamperidou, C., Duan, W., Liu, T., and Feng, J.: Asymmetry of the predictability limit of the warm ENSO phase, Geophysical Research Letters, 45, 7646 – 7653, https://doi.org/10.1029/2018GL077880, 2018.
[7] Hou, Z., Li, J., Ding, R., and Feng, J.: Investigating decadal variations of the seasonal predictability limit of sea surface temperature in the tropical Pacific, Climate Dynamics, 2022, 1 – 18, https://wgmktpgxscrjl-s.p.lib.tju.edu.cn/10.1007/s00382-022-06179-3, 2022.
[8] Li, J. and Ding, R.: Temporal – Spatial Distribution of Atmospheric Predictability Limit by Local Dynamical Analogs, Monthly Weather Review, 139, 3265 – 3283, https://doi.org/10.1175/MWR-D-10-05020.1, 2011.
[9] Li, J., Feng, J., and Ding, R.: Attractor radius and global attractor radius and their application to the quantification of predictability limits, Climate Dynamics, 51, 2359 – 2374, https://doi.org/10.1007/s00382-017-4017-y, 2017.
[10] Li, X., Ding, R., and Li, J.: Determination of the backward predictability limit and its relationship with the forward predictability limit, Advances in Atmospheric Sciences, 36, 669 – 677, https://doi.org/10.1007/s00376-019-8205-z, 2019.
[11] Li, X., Ding, R., and Li, J.: Quantitative study of the relative effects of initial condition and model uncertainties on local predictability in a nonlinear dynamical system, Chaos, Solitons & Fractals, 139, 110094, https://doi.org/10.1016/j.chaos, 2020.
[12] Li, X. and Ding, R.: The backward nonlinear local Lyapunov exponent and its application to quantifying the local predictability of extreme high-temperature events, Climate Dynamics, 60, 2767 – 2781, https://doi.org/10.1007/s00382-022-06469-w, 2023.
[13] Li, X., Ding, R., and Li, J.: The BaSIC method: a new approach to quantitatively assessing the local predictability of extreme weather events, Climate Dynamics, 60, 3561 – 3576, https://doi.org/10.1007/s00382-022-06526-4, 2023.

2. The authors perturbed the true states with noise and found that mean DSID values approach saturation as noise intensity increases. However, the noise amplitudes were relatively weak (<0.5). If larger noise magnitudes were superimposed on the true states, e.g., 10%–20% of the natural variability of the Lorenz models, would the mean DSID values still saturate, and would the saturation value remain at 3.0?

Response : Thanks very much for your valuable question.

Firstly, In our original noise addition procedure, the Lorenz63 model was first standardized, such that the x, y, and z sequences of the model were adjusted to have a mean of 0 and a variance of 1. On this basis, Gaussian noise with zero mean and unit variance was added, with α denoting the noise intensity (α ranging from 0 to 0.5). Thus, the noise intensity has actually reached 50% of the variability of standardized Lorenz63 model. When drafting the manuscript, we increased the noise intensity α from 0 to 1, meaning the noise amplitude reached 100% of the variability of the normalized Lorenz63 model. The results are shown in Fig.R1(a). The mean DSID values still saturate, and the saturation value remains around 3.0. This is because the Lorenz63 system is inherently three-dimensional, even when the noise completely dominates the system, its dimension remains 3.

Secondly, to better address your concern, we conducted another experiment. We did not standardize the Lorenz63 model and retained its natural variability. After computing its standard deviation, we added the Gaussian noise ranging from 0 to 1 times the standard deviation on it. In this experiment, the noise intensity has reached 100% of the natural variability of the Lorenz63 model. The results are shown in Fig.R1(b). The mean DSID values still saturate and remain around 3.0. Owing to the larger noise intensity with natural variability, the mean DSID values in this second experiment approach the saturation value earlier (at $\alpha = 0.15$) compared to the original experiment.

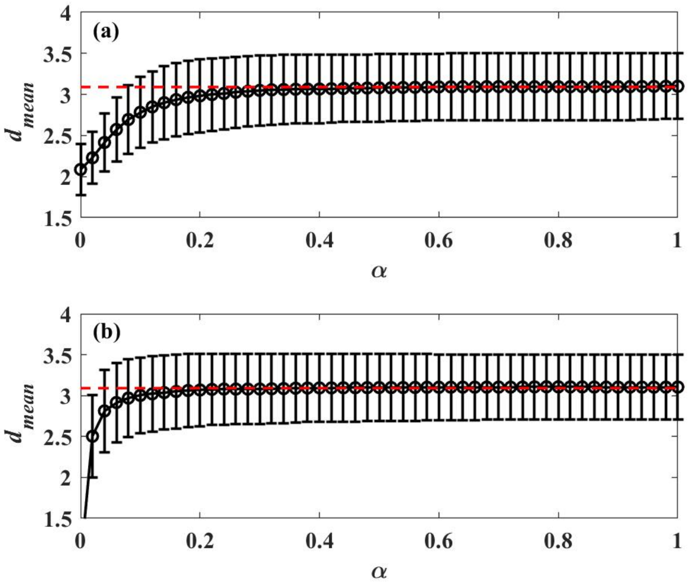

Figure R1 The variation of mean (dots) and standard deviation (bars) of d with the noise intensity α from noise-corrupted Lorenz-63 system. (a) The noise-corrupted Lorenz-63 system, normalized first. (b) The noise-corrupted Lorenz-63 system, not normalized and retaining its natural variability. The red line for the asymptotic mean of d (3.09).

3. In the present study, was only a fixed threshold at the 2% quantile adopted for selecting the analogues? Has a sensitivity test been conducted? Would different thresholds (e.g., 1% and 5%) significantly alter the DSID series and its detected mean shift around 2007?

Response : Thanks for your valuable question. The explanation for this question is as follows:

Previous studies have shown that the performance of the dynamical systems (DS) method is not sensitive to the choice of the threshold. In the literature, Faranda et al.(2017) applied the DS method to investigate the predictability and extremes of atmospheric flow over the North Atlantic[1].They initially set the threshold p to 0.98 (i.e., the top 2% of values), and subsequently examined the robustness of the results under various threshold configurations. It was found that the results remained stable when the threshold p ranged from 0.975 to 0.99. In fact, if too high a threshold is selected, the number of values exceeding it is insufficient to successfully fit the generalized Pareto distribution. Datseris et al.(2023) indicated that the choice of quantile threshold p should satisfy the condition that $N(1 - p) \geq 100$-1000 (N is the data length), to ensure that the number of exceedances is sufficient for a reliable fit**[2]**. And the quantile threshold p should satisfy the NRMSE test **[2]**. In this study, the quantile threshold p=0.95, 0.98, and 0.99 all satisfy this requirement.

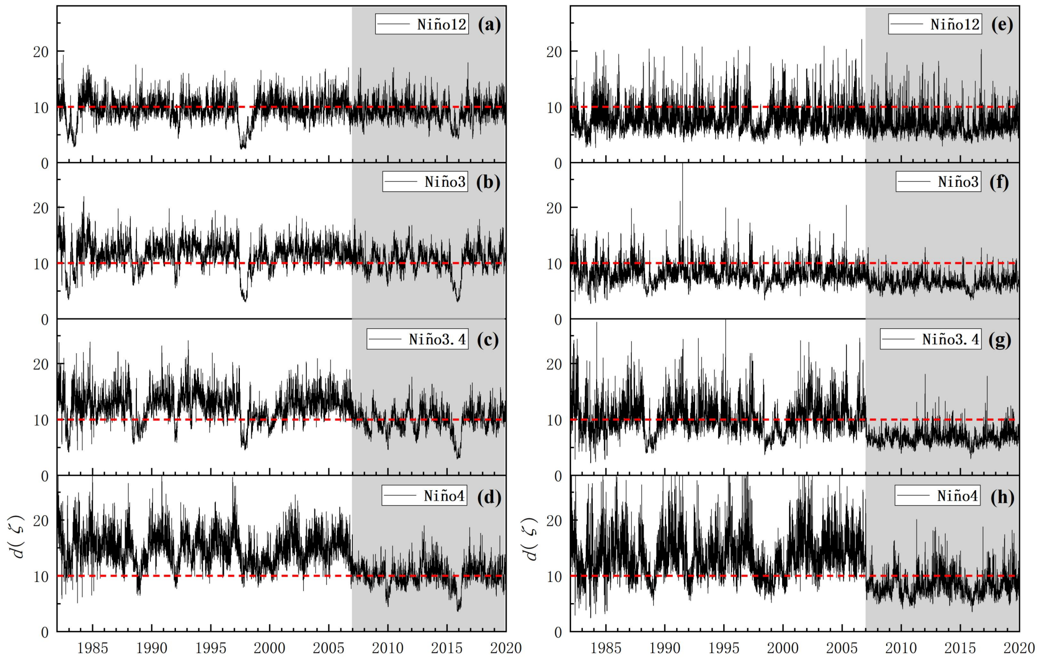

Figure R2. Evolution of the instantaneous dimension d of SST at different thresholds over four typical Niño regions.The threshold is 5% quantile for (a)-(d) and 1% quantile for (e)-(h). Panels (a) and (e) correspond to Niño1+2 area, (b) and (f) to Niño3 area, (c) and (g) to Niño3.4 area, and (d) and (h) to Niño4 area. The gray shaded area represents the years after 2007. Red dashed lines for the eye-guided variation of DSID mean value.

To assess the robustness of our results, we performed a sensitivity analysis with respect to the threshold choice. The 5% (p = 0.95) and 1% (p = 0.99) quantiles were adopted as thresholds for identifying analogous states, and the DSID of SST was computed for the four Niño areas. The results are presented in Fig.R2, TabR1, and Table R2. When the stricter threshold (1%) is applied, only the nearest analogous states in the phase space are selected for each reference point ζ. As a result, the chosen phase-space trajectories become more convergent, leading to a lower mean of the computed DSID series (see $\bar{d}$ in Tables R1 and R2). Nonetheless, although the DSID series obtained with the two thresholds (5% and 1%) exhibit minor differences, both series clearly capture a significant mean shift around 2007. This confirms that the choice of threshold does not compromise the ability of the DSID to detect changes in dynamical states. We therefore conclude that the selection of the quantile threshold does not affect the findings of this study.

The results of the sensitivity test are discussed in the manuscript (with the revised parts highlighted in yellow) and provided in the Supporting Information.

Table R1. Statistics of the mean and variance of the DSID before and after 2007 at the 5% quantile threshold. Bold for significant at the level of 0.001 from moving block bootstrap (MBB) test

	$\bar{d}$			σ^2		
	Before 2007	After 2007	$\bar{d}_{Diff}$	Before 2007	After 2007	σ^2_{Diff}
Niño 1+2	9.41	9.07	**-0.34**	5.09	3.39	**-1.70**
Niño 3	11.49	10.30	**-1.19**	6.15	4.89	**-1.25**
Niño 3.4	12.30	9.73	**-2.58**	7.40	4.67	**-2.73**
Niño 4	14.73	10.00	**-4.72**	8.90	4.13	**-4.77**

Table R2. Statistics of the mean and variance of the DSID before and after 2007 at the 1% quantile threshold. Bold for significant at the level of 0.001 from moving block bootstrap (MBB) test

	$\bar{d}$			σ^2		
	Before 2007	After 2007	$\bar{d}_{Diff}$	Before 2007	After 2007	σ^2_{Diff}
Niño 1+2	7.75	6.77	**-0.99**	5.68	3.77	**-1.90**
Niño 3	8.12	6.49	**-1.63**	3.59	1.41	**-2.18**
Niño 3.4	10.11	6.70	**-3.40**	8.43	1.74	**-6.69**
Niño 4	13.35	8.12	**-5.23**	14.60	3.92	**-10.68**

Reference:

[1] Faranda, D., Messori, G., and Yiou, P.: Dynamical proxies of North Atlantic predictability and extremes, Sci Rep, 7, 41278, https://doi.org/10.1038/srep41278, 2017.

[2] Datseris, G., Kottlarz, I., Braun, A. P., and Parlitz, U.: Estimating fractal dimensions: A comparative review and open source implementations, Chaos: An Interdisciplinary Journal of Nonlinear Science, 33(10), https://doi.org/10.1063/5.0160394, 2023.

4. Lines 220, “define” should be corrected to “defining”.

Response： Thanks for your careful correction. We modified it in the revised manuscript.

Reviewer: 2

Comments to the authors

The paper has some clear strengths. It asks a meaningful question, namely what DS instantaneous dimension is actually sensitive to and how it should be interpreted in tropical Pacific SSTA fields. The main result, especially the marked shift in DSID over Niño 3.4 and Niño 4 around 2007 and its weakening after low-pass filtering, is interesting and worth further investigation.

At the same time, I am not yet fully convinced by the paper in its current form. The attribution of the DSID shift mainly to high-frequency variability is plausible, but the evidence is still somewhat indirect. The choice of 2007 as a breakpoint feels rather empirical, the statistical testing is not fully convincing for autocorrelated time series, and the physical discussion remains limited. In addition, the manuscript also seems to overlook some highly relevant literature, which weakens its connection to the broader existing work on this topic. Please find more detailed comments below.

Response: Thanks for your comments and valuable suggestions.

Detailed Comments:

1. There are two references that are highly relevant to this work are not cited by this manuscript. I believe adding some discussion about them could help this work better connect with existing work in the community. In [1], they explored dynamical properties of a similar region with exactly the same dynamical systems theory indices. Although the two study have different perspectives, I'm still quite surprised to see that their study did not find such shift in mean value of local dimension. I wonder if the authors could provide some comments on the link and differences with [1]. The second reference I would recommend is [2]. This paper also examines the sensitivity of instantaneous dimension, and in my view its objective overlaps with yours to some extent. At the same time, it includes some discussion about the caveats of this methodology that worth mentioning in your work.

References

[1] Fabrizio Falasca and Annalisa Bracco. Exploring the tropical pacific manifold in models and observations. Physical Review X, 12(2):021054, 2022.

[2] George Datseris, Inga Kottlarz, Anton P Braun, and Ulrich Parlitz. Estimating fractal dimensions: A comparative review and open source implementations. Chaos: An Interdisciplinary Journal of Nonlinear Science, 33(10), 2023.

Response : Thanks for your valuable suggestion.

In reference [1], the authors proposed a general data-driven framework to characterize the climate attractor and showcased it in the tropical Pacific Ocean using a reanalysis as observational proxy and two state-of-the-art models[1]. In this study, they quantified the local geometry and local stability of the high-dimensional, multivariable climate attractor through the local dimension and persistence metrics, but the shift in the mean value of local dimension was not detected.

In our view, **this is because their study treated the entire tropical Pacific**

region (20°S - 20°N, 120°E - 70°W) as a unified analysis area, thus failing to capture the significant differences in dynamical properties among different sub-regions. In reality, El Niño events exhibit considerable spatial diversity. Based on the location of the center of SST anomaly, it can be classified into Eastern-Pacific El Niño (EP El Niño), Central-Pacific El Niño (CP El Niño), and mixed El Niño (see Fig.R3)[2]. The climatic impacts of different El Niño types also differ significantly[3-5]. Studies have shown that over the past two decades, the SST anomalies associated with El Niño events have shifted westward toward the central Pacific[6]. Correspondingly, CP El Niño events have become more frequent, while EP El Niño events have decreased[7]. Therefore, as the dynamical properties of different regions in the tropical Pacific change in distinct ways, it is necessary to conduct refined, region-specific studies. In this study, we apply the DS method to four Niño regions and indeed identify a region-dependent decline in the mean value of DSID, which is more pronounced in the central Pacific than in the eastern Pacific.

Furthermore, they employed a multivariate approach to investigate the evolution of a highly dimensional climate system, including the surface temperature (T), zonal and meridional velocities (u, v) at the surface, and outgoing longwave radiation (OLR). **The multivariate analysis may have weakened the contribution of the surface temperature (T) as a single variable to the local dimension.** In addition, the variables used as 'surface temperature (T)' are temperature at 2m in ERA5 and temperature at the surface in the models in their study, which essentially represents the air temperature just above the sea surface, rather than the actual sea surface temperature (SST). **This difference in the variables examined may partly account for the discrepancy between the two studies.**

The discussion concerning this reference has been added to 'Section 4. Discussion and Conclusions' and is highlighted in yellow in the revised manuscript, as follows:

'Falasca and Bracco(2022) applied the same methodology over the entire tropical Pacific domain (20°S–20°N, 120°E–70°W) and failed to detect a significant mean decrease of DSID. This may be attributable to the choice of a larger research region and the multivariate approach, which together mask the structural change in the high-frequency variability of SSTA over the tropical central–western Pacific. This indicates that, when applying DSID to diagnose climate data, a refined regional analysis is more effective than a global aggregation in capturing regionally distinct variations of dynamical properties.'

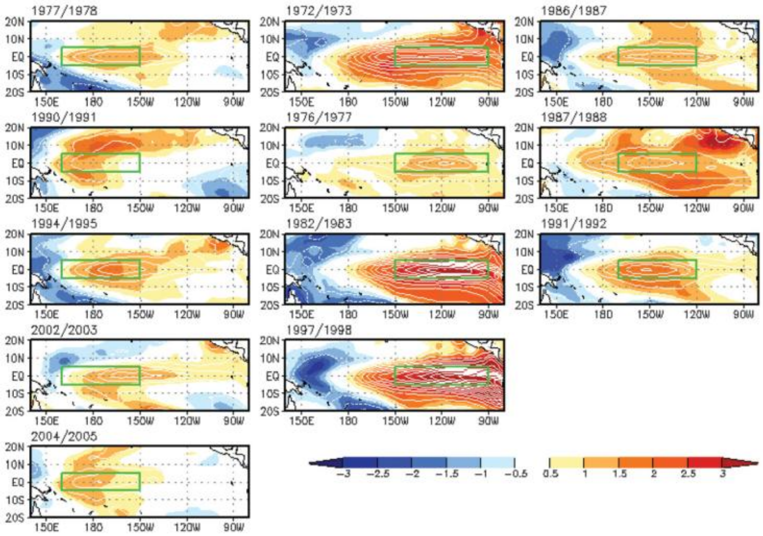

Figure R3: The SST anomalies of El Niño events during 1970–2005. The anomalies are averaged from September to the following February. Shading indicates normalized anomalies; contour interval is 0.3 K. The El Niño events are classified into (left) CP El Niño, (middle) EP El Niño, and (right) mixed El Niño. The green boxes indicate (left) Niño-4, (middle) Niño-3, and (right) Niño-3.4 regions. **(Source: reference [2])**

The second reference you recommended used the instantaneous dimension to definie and estimate fractal dimensions and discussed the caveats of this methodology[8]. **In the revised manuscript, we have discussed and cited the details regarding the use of this method as described in the reference.**

In addition, according to the method described in the reference, we examined the choice of the quantile threshold p. We adopted the 2 % quantile (i.e., p=0.98) as the threshold in our study, which satisfies the condition that N(1-p) falls within the range of 100-1000 (N = 13870 and N(1 − p) = 277.4). This value of p also satisfies the NRMSE test proposed in the reference, in the sense of most NRMSE values being less than 0.5. It indicates that the quantile threshold (i.e., 2 %) used in the manuscript is reasonable. To assess the robustness of our results, we added a sensitivity analysis with respect to the threshold choice. The 1% (p = 0.99) and 5% (p = 0.95) quantiles were adopted as thresholds for identifying analogous states, and the mean shift in DSID can be detected in both cases. **This confirms that the selection of the quantile threshold does not affect the findings of this paper.** The results of the sensitivity analysis are discussed in the manuscript (with the revised parts highlighted in yellow) and are provided in the Supporting Information.

References

[1] Falasca, F. and Bracco, A.: Exploring the tropical Pacific manifold in models and observations, Physical Review X, 12, 021054, https://doi.org/10.1103/PhysRevX.12.021054, 2022.

[2] Kug, J.-S., Jin, F.-F., and An, S.-I.: Two Types of El Niño Events: Cold Tongue El Niño and Warm Pool El Niño, Journal of Climate, 22, 1499–1515, https://doi.org/10.1175/2008JCLI2624.1, 2009.

[3] Ashok, K., Behera, S. K., Rao, S. A., Weng, H., and Yamagata, T.: El Niño Modoki and its possible teleconnection, Journal of Geophysical Research: Oceans, 112, C11007, https://doi.org/10.1029/2006JC003798, 2007.

[4] Yu, J.-Y. and Kim, S. T.: Identifying the types of major El Niño events since 1870, International Journal of Climatology, 33, 2105–2112, https://doi.org/10.1002/joc.3575, 2013.

[5] Capotondi, A., Wittenberg, A. T., Newman, M., Di Lorenzo, E., Yu, J.-Y., Braconnot, P., Cole, J., Dewitte, B., Giese, B., Guilyardi, E., Jin, F.-F., Karnauskas, K., Kirtman, B., Lee, T., Schneider, N., Xue, Y., and Yeh, S.-W.: Understanding ENSO diversity, Bulletin of the American Meteorological Society, 96, 921–938, https://doi.org/10.1175/BAMS-D-13-00117.1, 2015.

[6] Hu, S. and Fedorov, A. V.: Cross-equatorial winds control El Niño diversity and change, Nature Climate Change, 8, 798–802, https://doi.org/10.1038/s41558-018-0248-0, 2018.

[7] Capotondi, A., Wittenberg, A. T., Newman, M., Di Lorenzo, E., Yu, J.-Y., Braconnot, P., Cole, J., Dewitte, B., Giese, B., Guilyardi, E., Jin, F.-F., Karnauskas, K., Kirtman, B., Lee, T., Schneider, N., Xue, Y., and Yeh, S.-W.: Understanding ENSO Diversity, Bulletin of the American Meteorological Society, 96, 921–938, https://doi.org/10.1175/BAMS-D-13-00117.1, 2015.

[8] Datseris, G., Kottlarz, I., Braun, A. P., and Parlitz, U.: Estimating fractal dimensions: A comparative review and open source implementations, Chaos: An Interdisciplinary Journal of Nonlinear Science, 33, 103101, https://doi.org/10.1063/5.0163888, 2023.

2. I find the first half of the title a bit awkward. I would suggest something like: ‘On the Sensitivity of the Instantaneous Dimension of Dynamical Systems ...’

Response : Thanks for your valuable suggestion. We have revised the title of the manuscript to “On the sensitivity of the instantaneous dimension of dynamical systems and its insights on sea surface temperature anomaly field over the tropical Pacific”.

3. For Figure 1, one could add a subfigure showing the bounding boxes of these regions to make it more clear for the readers. I would also suggest to add a time series of the El Nino index considering the strong modulation of El Nino to the dynamical properties of the studied region.

Response : Thanks for your valuable suggestion.

We have added a subfigure to Figure 1 showing the geographical locations of the four Nino regions (see Fig. R4(e)). In addition, considering the strong modulation of ENSO on the dynamical properties of the tropical Pacific sea surface temperature (SST), we have added the time series of Niño3.4 index (see Fig. R4(c)), which is the most widely used index for defining El Niño and La Niña events.

As analyzed in the manuscript, the DSID value drops sharply during typical ENSO events. This indicates that the spatial pattern of SST is relatively uniform at these periods and can be described with fewer degrees of freedom in phase space. In contrast, when the SST is in neutral state, the DSID is higher, reflecting a more complex and irregular spatial pattern of SST. This phenomenon is consistent with observations. However, the modulation of DSID by ENSO is not the primary focus of this study. This paper mainly investigates the sensitivity of the DSID metric and its region-dependent mean decrease after 2007.

All relevant revisions have been highlighted in yellow in the revised manuscript.

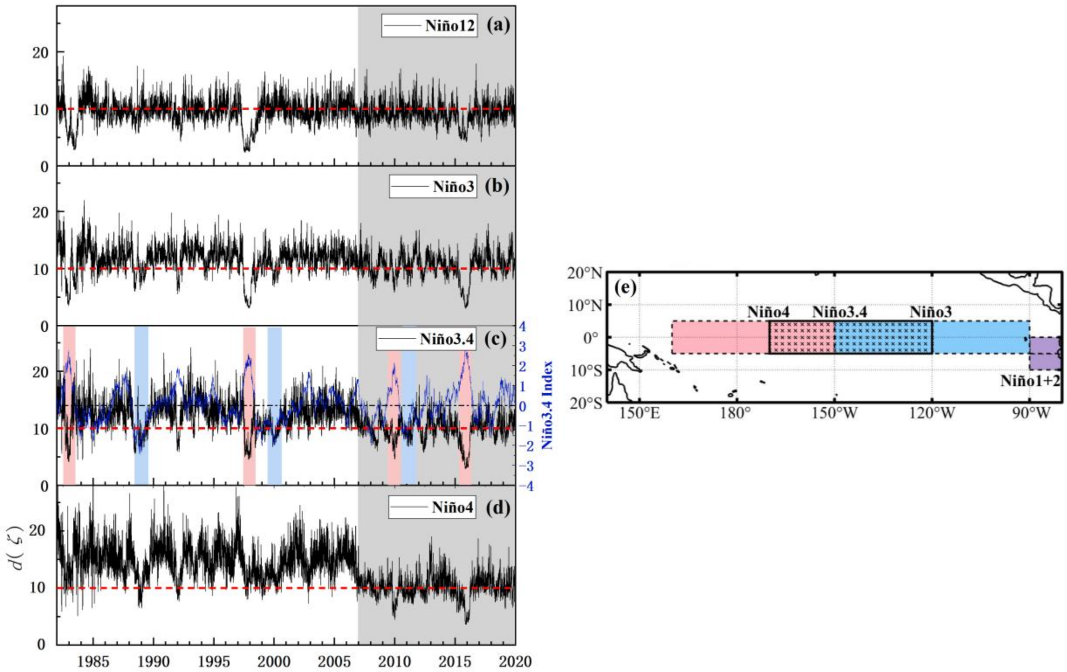

Figure R4: The evolution of calculated instantaneous dimension *d* of daily sea surface temperature anomalies over four typical Niño regions. (a) Niño 1+2, (b) Niño 3, (c) Niño 3.4, (d) Niño 4. The blue curve in (c) represents the Niño3.4 index. Shading denotes the typical El Niño events (red) and La Niña events (blue). The gray shaded area represents the period after 2007. Red dashed lines are eye guides for the variation of the DSID mean value. (e) The geographical locations of the four Niño regions : Niño 1+2 region (0°-10°S, 90°-80°W), Niño 3 region (5°N-5°S, 150°W-90°W), Niño 3.4 region (5°N-5°S, 170°W-120°W) and Niño 4 region (5°N-5°S, 160°E-150°W).

4. Line 59: The authors could consider to cite a more recent work on predictability: [3]. This reference [4] could also be added as it also works on tropical Pacific.

References

[3] Chenyu Dong, Davide Faranda, Adriano Gualandi, Valerio Lucarini, and Gianmarco Mengaldo. Time-lagged recurrence: A data-driven method to estimate the predictability of dynamical systems. Proceedings of the National Academy of Sciences, 122(20):e2420252122, 2025.

[4] Davide Faranda, Yuzuru Sato, Chenyu Dong, Adriano Gualandi, Robin Noyelle, Tommaso Alberti, Berengere Dubrulle, Lucas Fery, Gabriele Messori, Mathieu Vrac, et al. El niño and droughts in southeast asia: A stochastic-chaotic modeling approach. Physical Review E, 111(6):064209, 2025.

Response : Thanks for your valuable suggestion. We have added citations to two relevant references. The corresponding revisions have been highlighted in yellow in the manuscript.

5. Line 130: If I understand correctly, the authors would like to perform a statistical test to see if there is an actual drop in the mean value. I would recommend bootstrapping than the surrogate data approach since it could be more clearly stated with a null hypothesis.

Response : Thanks very much for your valuable suggestion.

As indicated in the literature, the shuffling surrogate is the most basic type of surrogate data method. It is designed to test the null hypothesis that the data are generated by an independent and identically distributed (i.i.d.) random process, that is, without any temporal dependence or deterministic structure **[1-3]**. Therefore, the shuffling surrogate method is not suitable for statistical testing of autocorrelated time series.

Given the strong autocorrelation in the DSID time series, we applied the **block bootstrap method** to test whether the shift in the mean value (SMV) of DSID after 2007 is statistically significant. The bootstrap method, a resampling technique, was proposed by Efron in 1979 **[4]**. Its principle is to conduct repeated sampling with replacement from the original data, thereby increasing the sample size and facilitating the estimation of parameters and confidence intervals.The classical bootstrap method assumes that the data are independent and identically distributed (i.i.d.) **[4]**. However, for data with autocorrelation, directly resampling individual points would disrupt their inherent structure, leading biased estimations. The block bootstrap method can address this problem.The procedure is as follows: the data are divided into blocks, the blocks are resampled with replacement, and the resampled blocks are then concatenated to form a new sequence. The dependency characteristics of the original sequence can be completely preserved within each block in this way. Depending on the different block partitioning methods, it can be classified into the moving block bootstrap (MBB) **[5]**, the non-overlapping block bootstrap (NBB) **[6-7]**, and the stationary bootstrap (SB) **[8]**.

We employ the most commonly used moving block bootstrap (MBB) method for statistical testing and determine the block length l using the Hall–Horowitz–Jing method **[9]**. The specific steps are as follows. We generated 1000 pseudo-sequences using the MBB method to estimate the DSID. Then we have tested the significance of all results by repeatedly calculating the mean values or standard deviation of the estimated DSID over two periods separated by a break found in Fig. 1 or other figures. From 1000 calculated DSID differences, we can obtain the confidence intervals at a given significance level. The results of the significance test using MBB indicate that the conclusions regarding statistical significance in the original manuscript remain essentially unchanged. The corresponding revisions have been highlighted in yellow in the manuscript.

Reference

[1] Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Farmer, J.D.: Testing for nonlinearity in time series: the method of surrogate data, Physica D, 58, 77–94, https://doi.org/10.1016/0167-2789(92)90102-S, 1992.

[2] Schreiber, T., and Schmitz, A.: Surrogate time series, Physica D, 142, 346–382, https://doi.org/10.1016/S0167-2789(00)00043-9, 2000.

[3] Small, M., and Tse, C.K.: Applying the method of surrogate data to cyclic time series, Physica D, 164, 187–201, https://doi.org/10.1016/S0167-2789(02)00382-2, 2002.

[4] Efron, B.: Bootstrap methods: another look at the jackknife, in Breakthroughs in Statistics: Methodology and Distribution, 569－593, https://doi.org/10.1007/978-1-4612-4380-9_41, 1992.

[5] Künsch, H. R.: The jackknife and the bootstrap for general stationary observations, Ann. Statist., 17, 1217－1241, https://doi.org/10.1214/aos/1176347265, 1989.

[6] Hall, P.: Resampling a coverage pattern, Stochastic Process. Appl., 20, 231–246,

https://doi.org/10.1016/0304-4149(85)90217-6, 1985.
[7] Carlstein, E.: The use of subseries values for estimating the variance of a general statistic from a stationary sequence, Ann. Statist., 14, 1171–1179, https://doi.org/10.1214/aos/1176350057, 1986.
[8] Politis, D. N., and Romano, J. P.: The stationary bootstrap, J. Amer. Statist. Assoc., 89, 1303–1313, https://doi.org/10.1080/01621459.1994.10476870, 1994.
[9] Hall, P., Horowitz, J. L., and Jing, B.-Y.: On blocking rules for the bootstrap with dependent data, Biometrika, 82, 561–574, https://doi.org/10.1093/biomet/82.3.561, 1995.

6. Line 155: It looks like a SDE instead of observational noise.

Response: Thank you for your question. The confusion may stem from an inappropriate use of the symbol $\boldsymbol{X}'(t)$ in our original manuscript.

$$\boldsymbol{X}'(t)=\boldsymbol{X}(t)+\alpha\boldsymbol{\varepsilon}(t) \tag{4}$$

In Equation (4), $\boldsymbol{X}'(t)$ does not denote the time derivative $\boldsymbol{dX/dt}$; rather, it represents the state corrupted by additive observational noise of the ideal model. To remove this ambiguity, we have reformulated Equation (4) as $\boldsymbol{X}_{\mathbf{noise}}(t)=\boldsymbol{X}(t)+\alpha\boldsymbol{\varepsilon}(t)$ in the revised manuscript and explicitly stated that $\boldsymbol{X}_{\mathbf{noise}}$ corresponds to measured data contaminated by observational noise. This revision clarifies its distinction from the underlying deterministic evolution equations of the system.

We fully agree with the reviewer that the dynamical noise described by stochastic differential equation (SDE) is fundamentally distinct from the additive noise employed in our study. The additive noise we used is intended to represent sensor measurement errors or observational uncertainties during data acquisition. In this context, the noise is directly superimposed onto the observations without affecting the underlying deterministic evolution equation of the system. By contrast, the SDE mentioned by the reviewer typically involves dynamical (or process) noise, where the noise constitutes a term in the system's evolution equation and directly alters the true physical state at the next time step.

7. The explanation based on high-frequency variability is interesting, but I wonder whether the current evidence is still somewhat indirect. In particular, could the authors clarify more carefully why the change around 2007 should be interpreted primarily as a change in high-frequency variability, rather than a broader shift in the background state or low-frequency variability of the system?

Response: Thank you for your insightful question. In the following text, we will explain why the shift in the mean value (SMV) of DSID after 2007 is mainly attributed to changes in high-frequency variability.

Firstly, our filtering experiment provides the most direct evidence. If the SMV in DSID originated from a broader shift in the background state or low-frequency variability of the climate system, then applying a 7-day low-pass filter — which removes high-frequency components while preserving low-frequency signals — should preserve this SMV. However, we observed that the SMV almost completely disappeared after filtering (Fig. 4(b) in the manuscript).This clearly demonstrates that the energy causing the SMV is located in the high-frequency band, rather than the low-frequency background variations. The conclusion is further supported by idealized system experiments (Lorenz-63 and Lorenz-96 systems), where we reproduced the same phenomenon, confirming the universality of this mechanism.

Secondly, if the SMV originated from large-scale climate transitions, such as aa regime shift of the Pacific Decadal Oscillation (PDO), its impact should be spatially consistent. However, we found that SMV is more pronounced in the Niño 3.4 and Niño 4 regions, i.e., the central-western tropical Pacific. This regional dependence characteristic is incompatible with the attribution of large-scale climate transitions.

We acknowledge the reviewers' point that the evidence is still somewhat indirect. However, we believe that the direct filtering test and the identical results obtained in the ideal Lorenz system, as well as the phenomenon of region dependence in the SMV, can to some extent support this view. We have added these explanations (highlighted in yellow) in the revised manuscript to make our reasoning clearer and more explicit.