Response to Review # 2 for the manuscript: Cross-scale causal information flow from El Niño Southern Oscillation to precipitation in eastern China

Yasir Latif[1], Kaiyu Fan[2,3], Geli Wang[2,*], and Milan Paluš[1,*]

[1]Department of Complex Systems, Institute of Computer Science of the Czech Academy of Sciences, 182 00 Prague 8, Czech Republic
[2]Key Laboratory for Middle Atmosphere and Global Environment Observation, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China
[3]Dalian Meteorological Bureau, Dalian 116001, China
[*]Joint corresponding authors

Correspondence: Geli Wang (wgl@mail.iap.ac.cn) and Milan Paluš (mp@cs.cas.cz)

Dear Editors & Reviewers,

We would like to thank the reviewers for careful reading and insightful comments which give us the opportunity to improve the manuscript. There were many suggestions to change the figures, including the scale for the precipitation maps. Therefore we present relevant (sub)figures in two different scales, blue and red-yellow (the original) and ask the reviewers and editors to advice which one to use. In order to avoid confusion, we use the same numbers of figures (2–10) as in the original submission. In a future revised manuscript the numbers of figures will be decreased by one, since the former Fig. 1 will be presented as Table 1, as requested by a reviewer.

In the following there are our answers to the comments of the Reviewer # 2. Reviewer's comments are in *blue italics*, while our answers in black roman font. If we cite a text from the manuscript, the changes are in red color.

RC2: 'Comment on egusphere-2024-400', Anonymous Referee #2, 26 Mar 2024

Review of the manuscript "Cross-scale causal information flow from El Niño Southern Oscillation to precipitation in eastern China" by Yasir Latif , Kaiyu Fan, Geli Wang and Milan Paluš.

Authors present original research on the link between ENSO states and/or quasi-oscillatory (QO) ENSO components and QO Chinese-river-basins precipitation components obtained by a wavelet approach. Authors use an information-theoretic approach to establish causal links and optimal delays between ENSO and precipitation. They also estimate statistical significance causality thresholds using means over surrogates preserving the spectrum. The method and results are interesting and susceptible of development and generalization by using other possible atmospheric-oceanic indexes as drivers of precipitation. The material of the manuscript is acceptable for publication after the adjustment of several points that can improve much better the presented research.

We would like to thank the reviewer for the careful reading of the manuscript, its positive evaluation and valuable advices to improve the manuscript.

Major points:

Fig. 2 A color bar for the orography height must be included. Add in the maps the location of local precipitation stations used in manuscript.

Fig. 2 was modified as suggested, see below.

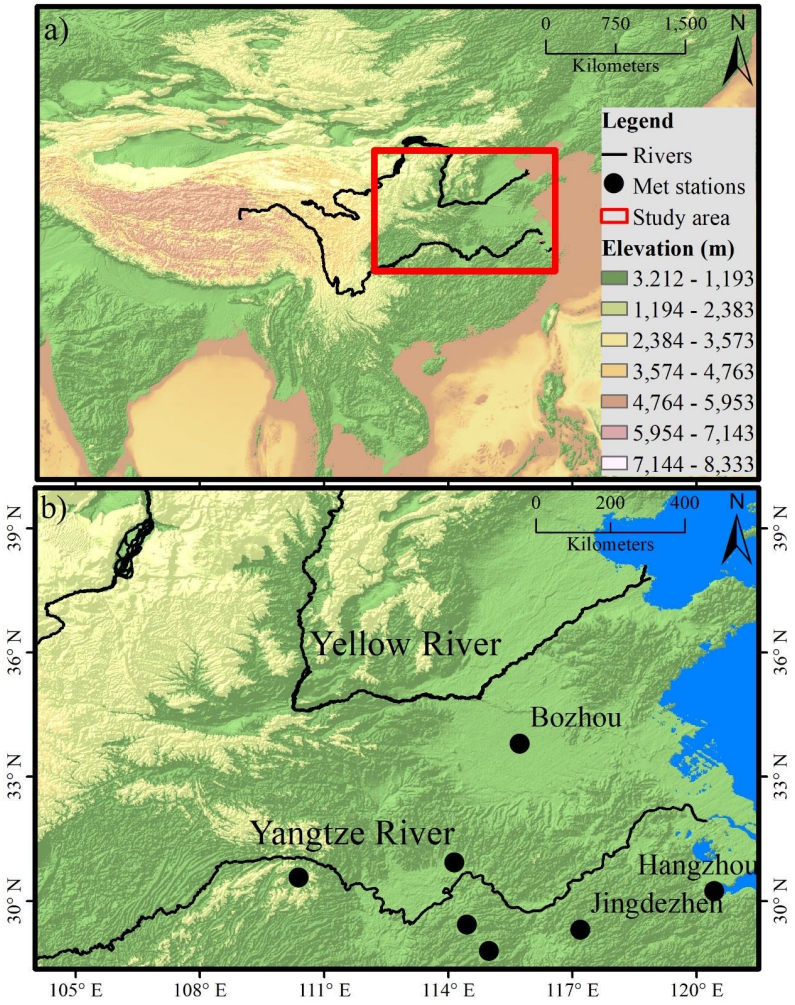

Figure 2. Study area. (Top) Localization of the selected region in Yangtze and Yellow River Basins. (Bottom) Study area in a detailed view, including the positions of selected stations.

Fig. 4a. The ENSO time-series (or a proxy of that) is plotted (black curve) with indication of positive, negative, and neutral ENSO states. However, it is the ONI (lines 117-120) that is used to discriminate ENSO states. From the graph the cross of +0.5 and -0.5 to determine ENSO phases is not well suited. Change the graph accordingly with the chosen criterium of ENSO phases.

The data presented in the original form of Fig. 4 was normalized "raw", i.e., not anomalized data marked as Nino3.4 at https://www.cpc.ncep.noaa.gov/data/indices/ersst5.nino.mth.91-20.ascii. For better understanding of the definition of the ENSO states, in the new version of the Figure 4a we present anomalized Nino3.4 data as well as the ONI index used for the states definition. We present the new version of Fig. 4 below.

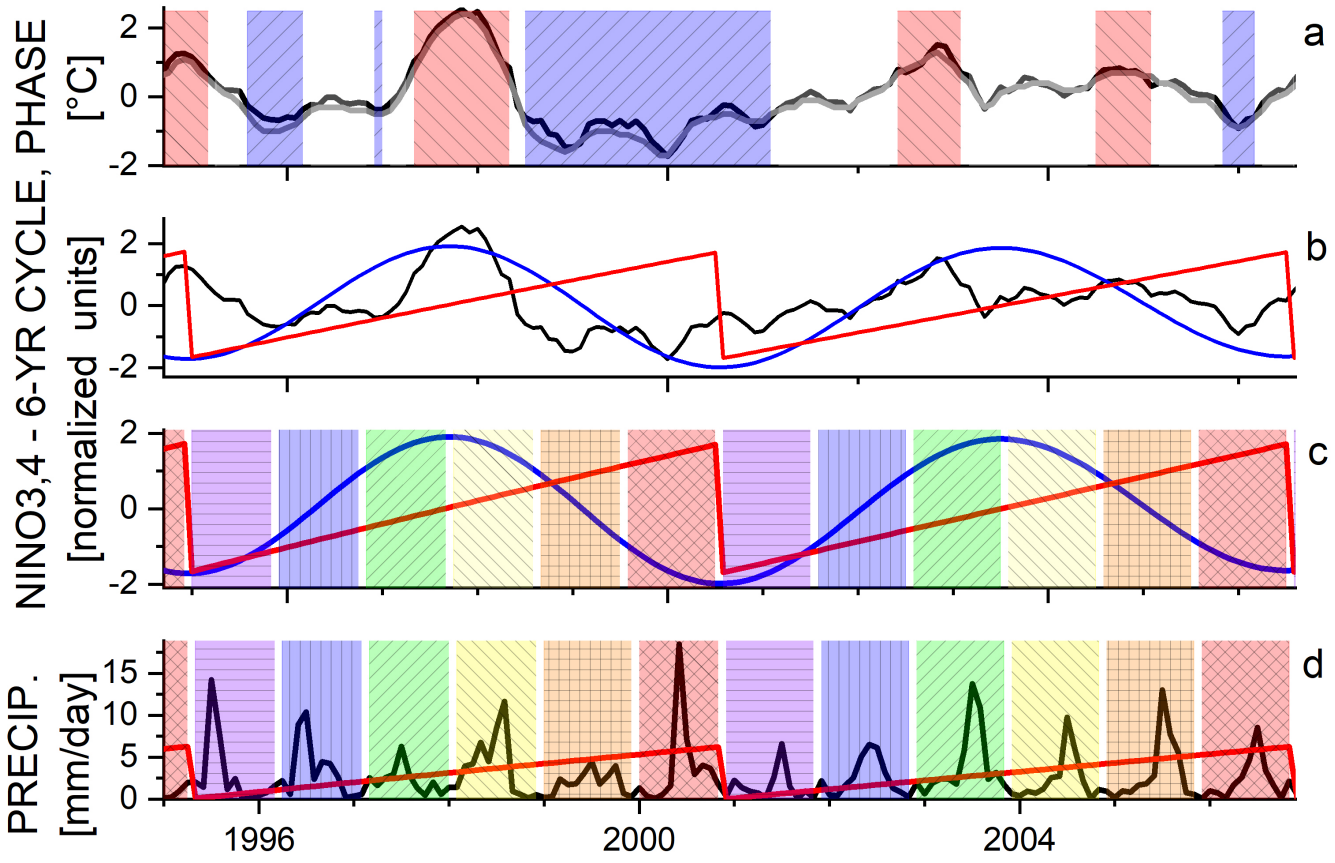

Figure 4. ENSO states and binning of the low-frequency cycle. From top to bottom: (a) A segment of anomalized Niño3.4 (black) and ONI (gray) time series with marked ENSO states: warm episodes ENSO+ (light red), cold episodes ENSO- (light blue) and neutral ENSO0 state (white). (b) The same segment of anomalized Niño3.4 time series (black) with its CCWT-extracted 6-yr component (blue) and the instantaneous phase (red) of the latter. (c) The 6-yr Niño3.4 component (blue) and its instantaneous phase (red). The bars of different colors and patterns mark the 6 phase bins into which each 6-yr cycle is divided. (d) A segment of reanalysis precipitation data from the gridpoint 33.75°N 115.75°E (black) and the 4 months lagged phase (red) of the 6-yr Nino3.4 cycle and related phase bins (bars of different colors and patterns) in which the precipitation conditional means are computed.

Line 171. The formula of the transfer-entropy (TE) (eq. 9) is proposed by Wibral et al. (2013) as a CMI in the case of SPO (Self Prediction Optimality) of Y states prior to the forecast delay tau. This is a very conservative estimate of TE since the SPO may be never reached with TE of eq.9 being underestimated. Authors shall comment on this.

Thank you very much for pointing to this issue. We can add an explanation why the Wibral formula is used for the causal delay estimation, while the original formulation 7, 8 for the statistical testing of the presence of causality:

The Wibral et al. (2013) formula 9 is used in order to establish the causal delay, while the formulas 7 and 8 are used for testing the statistical significance of uncovered causal relations. Wibral et al. (2013) formula 9 was proposed as a CMI in the case of Self Prediction Optimality (SPO) of y states prior to the forecast delay τ. This is a very conservative estimate of CMI/TE since the SPO may be never reached with CMI/TE of eq. 9 being underestimated. CMI estimated according to eq. 7 or 8 is more sensitive with respect to detection of causality.

Line 188. The Z statistics use Id and Is. Id is the CMI value estimated from the studied data, Is is the mean for 100 realizations of the surrogate data. What formula is used for Id and Is? Are formulas 8 or 9 used? Give an example used in the manuscript.

As stated above, formula 7/8 is used for testing, i.e. the Z-scores in Fig. 6 were obtained using eqs. 7 or 8. Eq. 9 was used only for the estimation of causal delay in Figs. 5a,d. More details will appear in the revised manuscript as:

Computing the conditional means, the precipitation time series is not exactly aligned in time with the ENSO states or the ENSO phase bins, since the causal effect of ENSO can occur with some time delay. The causal delay can be found in the causality analysis as follows: In Fig. 5a the conditional mutual information represents the causal influence of ENSO states on the precipitation (EASMI-ZQY index). It was computed using the Wibral et al. (2013) formula 9 in which $x(t)$ is a discrete 3-valued function of the ENSO states, $y(t)$ is the precipitation EASMI-ZQY index discretized into four bins using the equiquantal binning algorithm (Paluš and Vejmelka, 2007), $d_2 = 1$.

.....

Fig. 5d shows the conditional mutual information showing the causal influence of the phase of the 6-year component obtained from the Niño3,4 time series, on the precipitation amplitude for the variability in the quasi-biennial scale (blue). It was again computed using the Wibral et al. (2013) formula 9, however, since the cross-scale causality is evaluated, before applying Eq. 9, the Niño3.4 and precipitation data underwent CCWT and now $x(t)$ is the ENSO phase $\phi_f(t)$ for the frequency f related to the period 6 years and $y(t)$ is the precipitation amplitude $A_f(t)$ for the frequency f related to the period 2 years. $d_2 = 3$ and the Gaussian estimator (Paluš, 2014) is used.

Line 202-205. Conditional means of precipitation, given the phase of ENSO CCWT component (or the ENSO state) are computed on a grid basis. However on Figs. 8c,d and Fig 9c,d the values are overlapped over the grid cells. Figures must be redone precisely.

The grids which appeared in the first versions od figures were used to avoid interpolating of the mapped values. While mapping precipitation values, the interpolation has a physical sense, in mapping extrema occurrence in bins 1–6 we need to have only integer values, while interpolation does not have any sense. In the new version of figures we made the maps in a different way and the misleading grids did not appear. See the new versions of figures bellow.

Line 220. What authors conclude from the histogram of Fig.5c?

It was stated in the lines 220–221 of the original manuscript:

"It can be observed that the red bar lies inside the surrogate histogram which means the difference between two ENSO states is not statistically significant in this grid point."

In other words, the difference of this value can occur by chance, as the surrogate results suggest.

Fig. 5 Authors compute: The conditional mutual information measuring the causal influence of ENSO states on precipitation characterized by the EASMI-ZQY index (solid blue line) and causality in the opposite direction (dashed black line). What exact formulas authors use? (Eqs 8,9?). What are the embedding dimensions used? Be precise about X and Y in this case. By opposite direction, what authors mean? X, Y are swapped? Give a mathematical expression in the method section. Authors use certain precipitation stations. What was the criterium of choice? Explain. The values of significance values (red thresholds9 on panels a) and d) are not very well explained how they are computed. Explain it in detail in the method section.

The formulas and embedding dimensions are specified as follows:

The causal delay can be found in the causality analysis as follows: In Fig. 5a the conditional mutual information represents the causal influence of ENSO states on the precipitation (EASMI-ZQY index). It was computed using the Wibral et al. (2013) formula 9 in which $x(t)$ is a discrete 3-valued function of the ENSO states, $y(t)$ is the precipitation EASMI-ZQY index discretized into four bins using the equiquantal binning algorithm (Paluš and Vejmelka, 2007), $d_2 = 1$.

.....

Fig. 5d shows the conditional mutual information showing the causal influence of the phase of the 6-year component obtained from the Niño3,4 time series, on the precipitation amplitude for the variability in the quasi-biennial scale (blue). It was again computed using the Wibral et al. (2013) formula 9, however, since the cross-scale causality is evaluated, before applying Eq.

9, the Niño3.4 and precipitation data underwent CCWT and now $x(t)$ is the ENSO phase $\phi_f(t)$ for the frequency f related to the period 6 years and $y(t)$ is the precipitation amplitude $A_f(t)$ for the frequency f related to the period 2 years. $d_2 = 3$ and the Gaussian estimator (Paluš, 2014) is used.

The opposite direction indeed means swapped variables, as was explained in equations 7 and 8.

Figure 5 is a part of the Data and Methods section, the particular gridpoint was chosen as an example. The full grid of data is reported in Figure 7.

For the signidicance criterion we will add to the Fig. 5 caption:

The red line is the significance threshold given as the mean+2SD for the surrogate data.

Fig. 5 In this figure the authors choose a point in the North where the ENSO states does not discriminate significantly the precipitation (see crosses in Figs. 7b,d). However in the South region (Yangtze River basin) it is apparent the existence of locations where ENSO states have an important role on precipitation. Authors should add the equivalent of Figs. 5a,b,c for a particular significant location in the South.

Figure 5 is a part of the Data and Methods section, the particular gridpoint was chosen as an example. The full grid of data is reported in Figure 7.

Figure 6 presents the Z score of a certain CMI (eq. 7) between the instantaneous amplitudes of the CCWT of El-Niño and of the precipitation. Which is the value of lag tau used? Clarify. Values presented in Fig. 6 depend on tau. Explain.

The end of the subsection **2.6 Conditional mutual information as a causality measure** will be completed as follows:

... while the formulas 7 and 8 are used for testing the statistical significance of uncovered causal relations. For testing the cross-scale causality, before applying Eq. 8, the Niño3.4 and precipitation data underwent CCWT and now $x(t)$ is the ENSO phase $\phi_{f_i}(t)$ for a particular frequency f_i, and $y(t)$ is the precipitation amplitude $A_{f_j}(t)$ for a frequency f_j. The Gaussian estimator was used and $d_2 = 3$ was chosen as in (Paluš, 2014) based on "saturation of the results", i.e., obtaining unchanged results for $d_2 = 4$ in comparison with $d_2 = 3$. The tested value is the CMI average for time lags τ = 1 to 6 months, according to the recommendation in (Paluš and Vejmelka, 2007).

Fig. 6 uses values for 6 stations but the geographical coordinates are not given. Provide them in a Table in the data section and point them in the first maps.

The table will be added, see also the new version of Fig. 2. The six averaged stations include also two individually presented stations.

Table 1. Geographical coordinates of seven local precipitation stations used in the combined region of Yellow and Yangtze River basins

Station ID	Station name	Province	Longitude	Latitude
57355	Huangjiawan	Hubei	110.4 E	31 N
57494	Wuhan	Hubei	114.1 E	30.6 N
57598	Hejiadian	Jianxi	114.6 E	29 N
57799	Yankeng	Jiangxi	114.9 E	27.1 N
58102	Bozhou	Anhui	115.7 E	33.7 N
58457	Hangzhou	Zhejiang	120 E	30.2 N
58527	Jingdezhen	Jiangxi	117.2 E	29.3 N

Figure 7. Authors present the effects of two causal mechanisms. Concerning the effect of oscillatory components of ENSO, only results for the 6-year component are presented in a map. Since the QB component is also relevant in the annual cycle of precipitation, authors shall conceive a map (similar to Fig. 7a,c) giving the representativeness of such link.

In causality analysis, we report the effects of both QB and LF modes. In the conditional means analysis, using the six bins of the LF phase, both effects are apparent, as described in the paper: The effect of the QB mode results in alternating strong and weak years (bins), while the LF affects the precipitation-QB amplitude, i.e. the difference between the adjacent bins of the 6yr phase. Focusing on the QB mode alone, one would effectively average all even and odd bins and the LF mode would attenuate the effect of the QB mode. The present methodology estimates the combined effect of both modes. In principle we could distinguish, in each grid point, which mode has a stronger effect. We are afraid, however, that the manuscript is already very complex to understand, so that we leave such distinction for further research.

Conclusion: Authors say: 'physical mechanisms explaining the observed cross-scale information transfers are yet to be established, the uncovered causal relations'. Can authors elaborate a little bit more here,

We will extend the concluding paragraph as follows:

The fact that complex evolution of climate, atmosphere, or circulation regimes is influenced by interactions of dynamics on multiple time scales is known (Muñoz et al., 2017; Zhang et al., 2023). For instance, Muñoz et al. (2015) suggest that cross-time scale interactions between different climate drivers improves the predictive skill of extreme precipitation. Hsu et al. (2023) show that multiscale interactions, in particular, scale interactions between the monsoon mean field, two modes of intraseasonal oscillation, and synoptic disturbances, were driving the devastating floods in Henan Province, China during July 2021. Liu et al. (2023) used the multiscale window transform (MWT) and MWT-based energy and vorticity analysis (MS-EVA), to identify

three scales fields: basic-flow fields (>64 days), intraseasonal oscillation fields (8–64 days) and synoptic-scale-eddy fields (<8 days), responsible for the torrential rainfall event, which hit Zhengzhou on July 20, 2021. Ungerovich et al. (2023) emphasize the role of the large scale circulation anomalies associated with ENSO teleconnections in simulation of extreme rainfall events in Uruguay, while Pineda et al. (2023) suggest that the early onset of heavy rainfall on the northern coast of Ecuador in the aftermath of El Niño 2015/2016 was favored by the convective environment in late January due to cross-time-scale interference of the very strong El Niño event and a strong and persistent Madden-Julian oscillation. The presented research, however, is a first step in developing a methodology able to establish a solid statistical evidence for existence of cross-scale causal interactions and to estimate their effect in measurable, physical quantities. In particular, the results presented here can open a new direction in understanding and predicting precipitation anomalies in eastern Asia. Although physical mechanisms explaining the observed cross-scale information transfers are yet to be established, the uncovered causal relations can already be used in statistical or machine learning tools for forecasting precipitation anomalies. In related considerations, Muñoz et al. (2023) propose to find "Windows of Opportunity" in forecasts across timescales by combining wavelet spectral analysis and a non-stationary time-frequency causality analysis. Materia et al. (2024) try to understand the causal factors behind these windows of opportunity using Liang-Kleeman information flow (Liang, 2013). This study demonstrates the ability to identify sources of cross-scale predictability by using complex continuous wavelet transform and information-theoretic approach to causality (Paluš, 2014).

Minor points:

Line 53. After the work of Jajcay et al. (2018), it is worth citing the work of Pires et al. (2021) which also studies the interactions between quasi-oscillatory ENSO components and the role of triadic resonances and synchronization in the explanation of super El-Niños.

Pires C.A. and Hannachi A. (2021) Bispectral analysis of nonlinear interaction, predictability and stochastic modelling with application to ENSO, Tellus A: Dynamic Meteorology and Oceanography, 73:1, 1-30, DOI: 10.1080/16000870.2020.1866393

Thank you very much for bringing this reference which will be added in the revised manuscript.

Eq. (1) time t is missing in the complex exponential.

Thank you very much for pointing this omission, we will correct as follows:

$$\psi(t) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{t^2}{2\sigma_t^2}\right) \exp\left(2\pi i f t\right), \qquad 1$$

Line 127: Delete 'sigma t ='

There is $\sigma_f = 1/\pi\sigma_t$.

The use of red color for the rivers is not ideal since the color bar includes red in several figures. Suggestion: use black.

All the maps were modified and two version in two different color scales are given for assessment of the referees which is more appropriate.

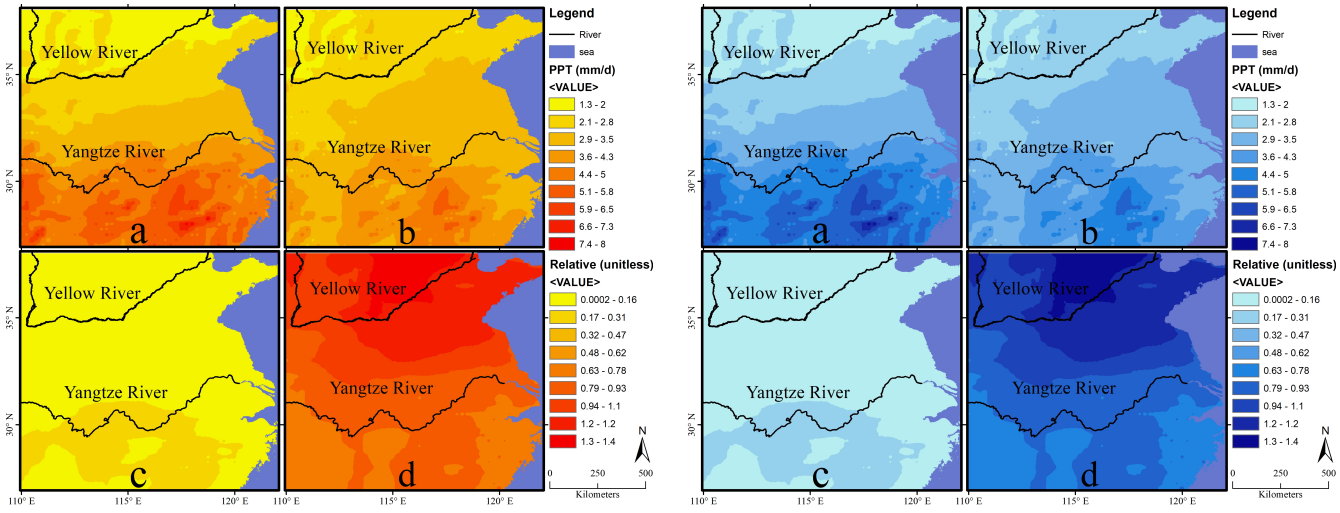

Figure 3. Precipitation and its variability in the study area. Spatial distribution of precipitation and its variability during 1951-2020; (a) mean precipitation, (b) precipitation standard deviation (SD), (c) relative difference between ENSO positive and ENSO neutral state, (d) relative precipitation SD (SD/mean precipitation).

We would like to thank again the reviewer for the suggestions which helped us to improve the manuscript.

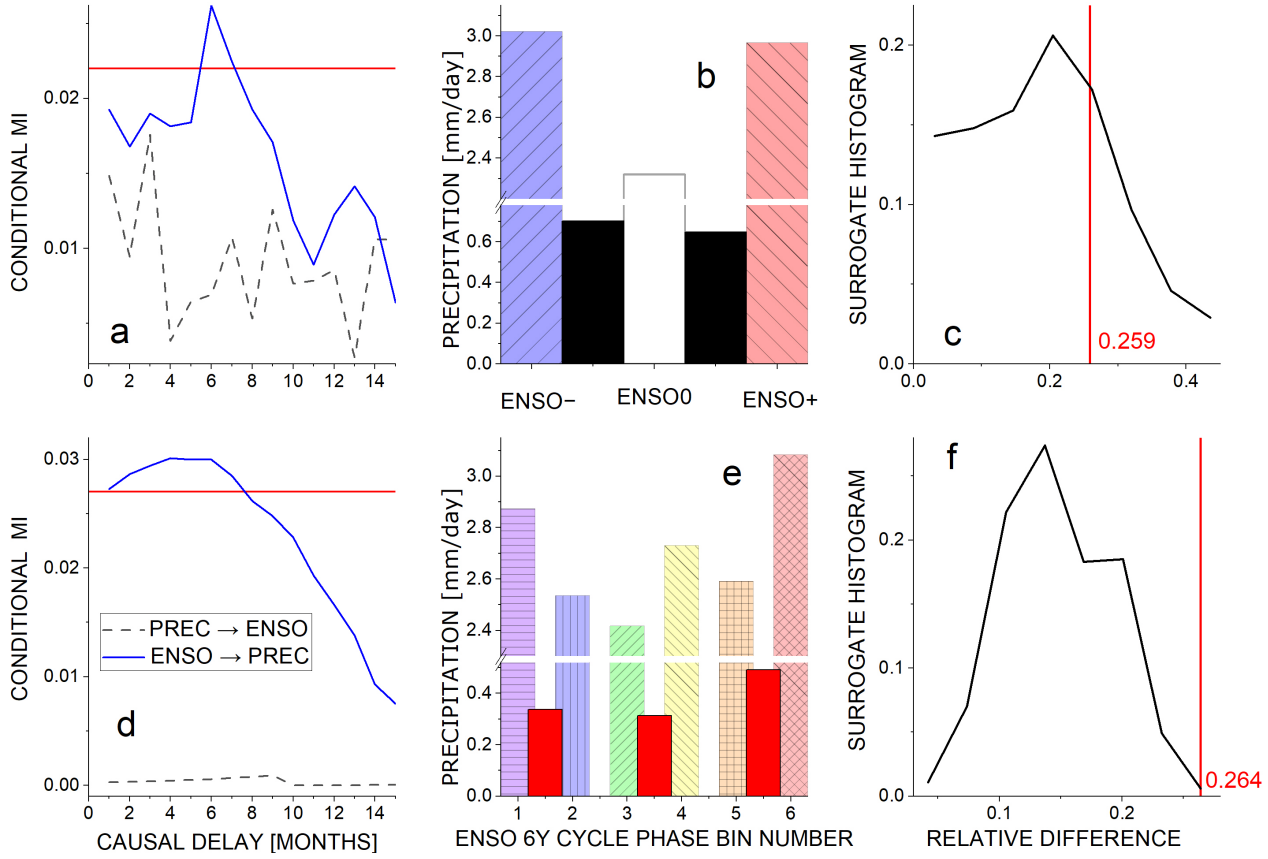

Figure 5. Causal mechanisms and their effects. (a) Conditional mutual information measuring the causal influence of ENSO states on precipitation characterized by the EASMI-ZQY index (solid blue line) and causality in the opposite direction (dashed black line). The red line is the significance threshold given as the mean+2SD for the surrogate data. (b) Conditional means for the precipitation data from the gridpoint 33.75°N 115.75°E for different ENSO states (ENSO- light-blue, ENSO0 white, ENSO+ light-red) computed for the lag of 6 months. Differences of the adjacent states in black. (c) Evaluation of statistical significance of the maximum relative difference between states, here ENSO- and ENSO0 (red vertical line) using the histogram for the surrogate data (black), (d) Conditional mutual information measuring the causal influence of ENSO 6yr cycle phase on 2yr cycle amplitude for precipitation characterized by the EASMI-ZQY index (solid blue line) and causality in the opposite direction (dashed black line). The red line is the significance threshold given as the mean+2SD for the surrogate data. (e) Conditional means for the precipitation data from the gridpoint 33.75°N 115.75°E for the 6 phase bins within the ENSO 6yr cycle (various colors). Differences of adjacent bins (red) considered as the amplitude of the precipitation quasibiennial cycle. The effect of the 6yr cycle phase is estimated as the maximum difference of the bin values – here the difference between the values of the 6th (orange) and the 3rd (yellow) bins. This value relative to the total precipitation mean is 0.264 and is marked by red vertical line in (f) and found statistically significant in comparison with the surrogate histogram (black).

References

Hsu, P.-C., Xie, J., Lee, J.-Y., Zhu, Z., Li, Y., Chen, B., and Zhang, S.: Multiscale interactions driving the devastating floods in Henan Province, China during July 2021, Weather and Climate Extremes, 39, 100 541, https://doi.org/https://doi.org/10.1016/j.wace.2022.100541, 2023.

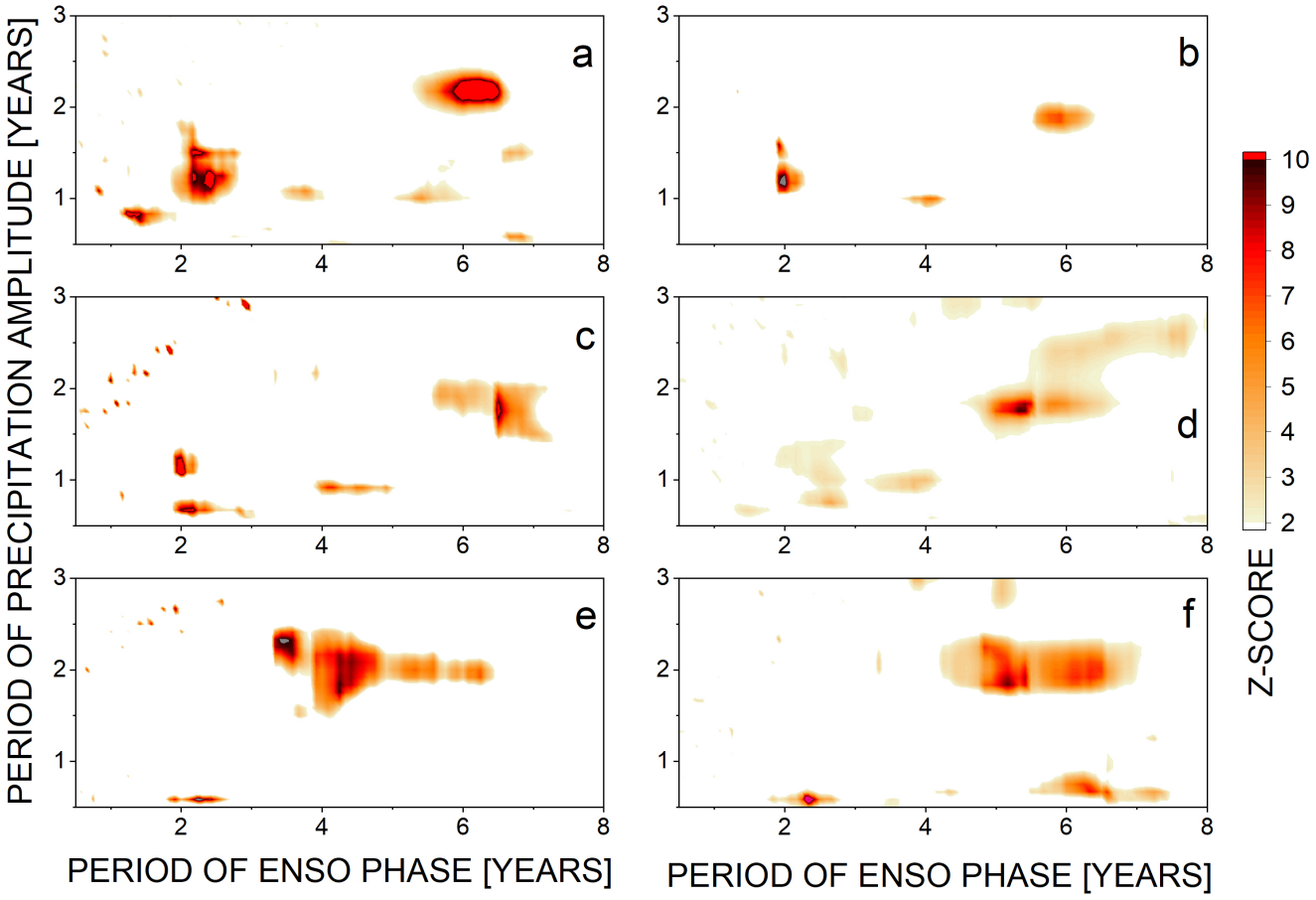

Figure 6. Cross-scale ENSO influence on precipitation in eastern China. Cross-scale phase-amplitude information transfer characterizing the causal influence of the phase of ENSO quasioscillatory components, with periods given on the abscissa, on the amplitude of precipitation quasioscillatory components with periods given on the ordinate. Significant causal influence of ENSO detected in (a) EASMI-ZQY index, (b) precipitation data from 6 stations from Hu Bei, Jiang Xi and Zhejiang provinces (averaged results), (c) precipitation data from station 58457 Hangzhou from Zhejiang province, (d) precipitation data from station 58527 Jingdezhen from Jiangxi province, (e) precipitation data from 58102 Bozhou station from An Hui province, (f) ERA 5 reanalysis precipitation data from the gridpoint 33.75°N 115.75°E. The colour codes present the conditional mutual information Z-score for $Z > 2$, obtained in the test using 100 realizations of surrogate data.

Liang, X. S.: The Liang-Kleeman Information Flow: Theory and Applications, Entropy, 15, 327–360, https://doi.org/10.3390/e15010327, 2013.

Liu, J., Tao, L., and Yang, Y.: Dynamical analysis of multi-scale interaction during the "21·7" persistent rainstorm in Henan, Atmospheric Research, 292, 106 857, https://doi.org/https://doi.org/10.1016/j.atmosres.2023.106857, 2023.

Materia, S., Ardilouze, C., and Muñoz, Á. G.: Deciphering Prediction Windows of Opportunity: A Cross Time-Scale Causality Framework, in: EGU General Assembly Conference Abstracts, pp. EGU24–18 766, 2024.

Muñoz, A. G., Goddard, L., Robertson, A. W., Kushnir, Y., and Baethgen, W.: Cross–Time Scale Interactions and Rainfall Extreme Events in Southeastern South America for the Austral Summer. Part I: Potential Predictors, Journal of Climate, 28, 7894 – 7913, https://doi.org/10.1175/JCLI-D-14-00693.1, 2015.

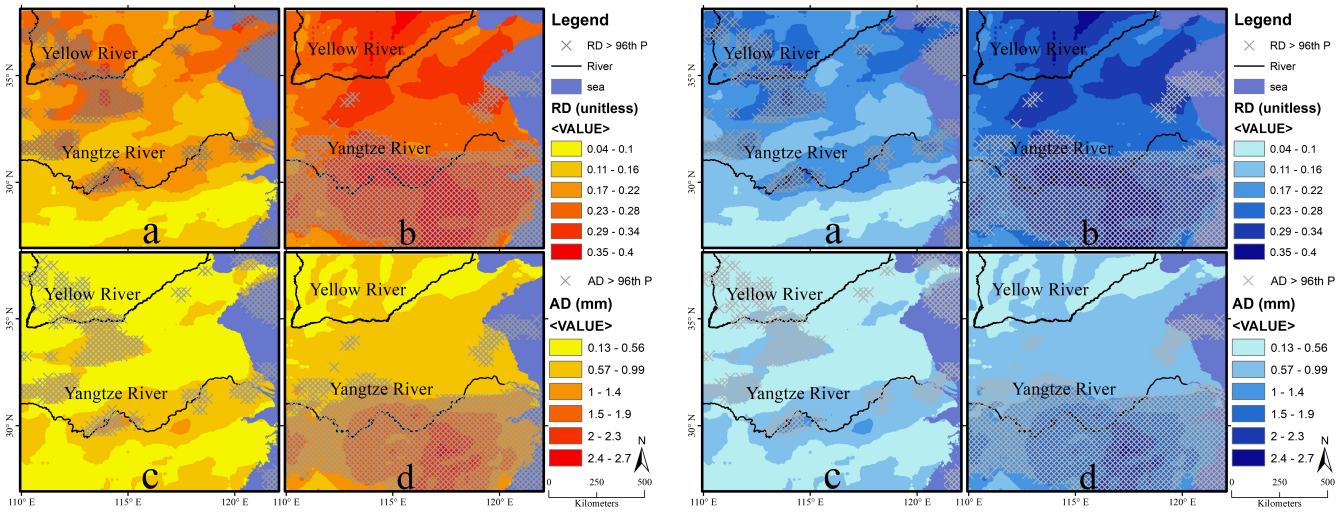

Figure 7. Quantification of the effects of two causal mechanisms. Relative (a, b) and absolute (c, d) maximum differences of precipitation conditional means: (a, c) conditioning on the six phase bins, i.e., the effect of the phase of the low-frequency ENSO component on precipitation; and (b, d) conditioning on the three ENSO states, i.e., the effect of the ENSO amplitude on precipitation. Statistically significant differences marked by X.

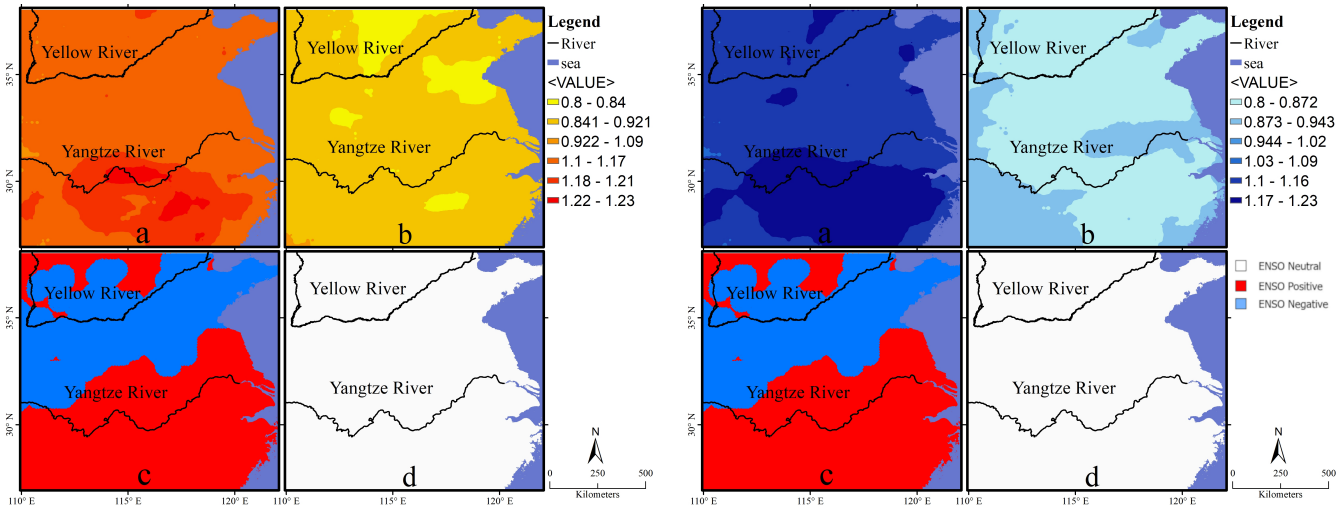

Figure 8. Geografy of ENSO amplitude influence. Maximum and minimum precipitation values in ENSO states; (a) maximum precipitation value, (b) minimum precipitation value, (c) ENSO state in which maximum precipitation occurs, (d) ENSO state in which minimum precipitation occurs.

Muñoz, A. G., Yang, X., Vecchi, G. A., Robertson, A. W., and Cooke, W. F.: A Weather-Type-Based Cross-Time-Scale Diagnostic Framework for Coupled Circulation Models, J. Climate, 30, 8951 – 8972, https://doi.org/10.1175/JCLI-D-17-0115.1, 2017.

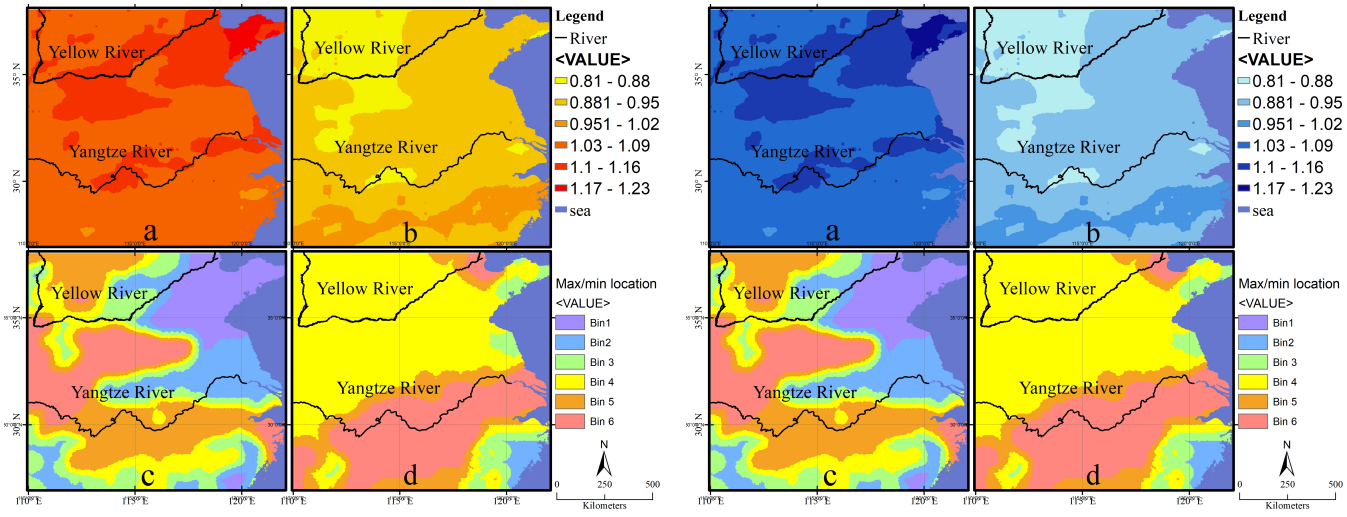

Figure 9. Geografy of ENSO phase influence. Maximum and minimum precipitation values in the six phase bins, given by the low-frequency ENSO phase; (a) maximum precipitation value, (b) minimum precipitation value, (c) ENSO phase bin in which maximum precipitation occurs, (d) ENSO phase bin in which minimum precipitation occurs.

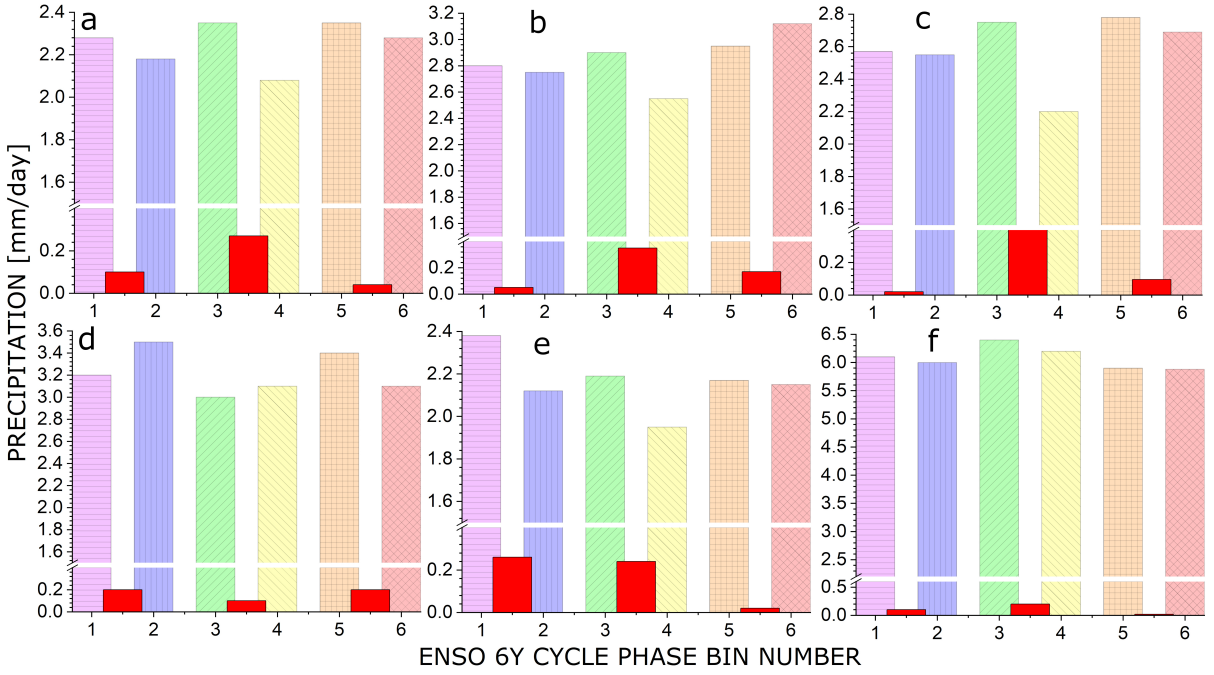

Figure 10. Local ENSO phase effects. Precipitation conditional means in the 6 ENSO phase bins at various coordinates: (a) 36°N 110°E, (b) 33°N 114°E, (c) 36°N 111°E , (d) 33°N 120°E, (e) 36°N 117°E, (f) 28°N 117°E. Red color represents difference between two adjacent bins.

Muñoz, Á. G., Doblas-Reyes, F., DiSera, L., Donat, M., González-Reviriego, N., Soret, A., Terrado, M., and Torralba, V.: Hunting for "Windows of Opportunity" in Forecasts Across Timescales? Cross it, in: EGU General Assembly Conference Abstracts, pp. EGU23–15 594, 2023.

Paluš, M.: Cross-scale interactions and information transfer, Entropy, 16, 5263–5289, https://doi.org/10.3390/e16105263, 2014.

Pineda, L. E., Changoluisa, J. A., and Muñoz, A. G.: Early onset of heavy rainfall on the northern coast of Ecuador in the aftermath of El Niño 2015/2016, Frontiers in Earth Science, 11, https://doi.org/10.3389/feart.2023.1027609, 2023.

Ungerovich, M., Barreiro, M., and Kalemkerian, J.: Simulation of extreme rainfall events in Uruguay: Role of initialization and large scale dynamics, Atmospheric Research, 292, 106 842, https://doi.org/https://doi.org/10.1016/j.atmosres.2023.106842, 2023.

Wibral, M., Pampu, N., Priesemann, V., Siebenhühner, F., Seiwert, H., Lindner, M., Lizier, J. T., and Vicente, R.: Measuring Information-Transfer Delays, PLOS ONE, 8, 1–19, https://doi.org/10.1371/journal.pone.0055809, 2013.

Zhang, Q., Zhang, Y., and Wu, Z.: Multiple time scales of the southern annular mode, Climate Dynamics, 61, 1–18, https://doi.org/10.1007/s00382-022-06476-x, 2023.