The authors present a statistical method to relate the NINO3.4 index to precipitation in Eastern China based on wavelet transform. The technique reveals a statistical relation to the phase of the quasi-biennial mode to the variability of the monsoon that suggest that there might be a physical mechanism linking the two. While the general idea of the paper seems clear to me, several details, especially related to figures and discussion are problematic and should be addressed before considering publication.

1) Fig 1. Is rather a table more than a figure, and it would be better if it was considered so. Also, graphically, it makes no sense to have it as a low-resolution figure. Also, the colors do not provide any information. I suggest to re-elaborate it as a simple table. The first reference Yang et al. 2004 has a typo and, in the reference list, it has no year associated. I suggest the authors to revise the other references as well.

2) The term quasi-oscillatory, which is used very often, is unclear. What does that mean? It gives the idea that these components "almost oscillate", however these components are wavelet projections so they oscillate by design. Do the authors mean quasi-periodic? This would be consistent with the use of wavelets, that have a narrow bandwidth in the Fourier space.

3) Data: the authors use ERA 5 reanalysis for the precipitation and analyze it to the single grid cell level. I think the authors should talk about the several limitations of reanalysis products for precipitation and why they believe that these limitations are not an obstacle for their study

4) Figure 4 is problematic because in NINO3.4, the 97-98 Niño should not be chopped. It has another, pointier shape. This should be true both using ERA5 and observations. That timeseries seems to be obtained with raw data rather than anomalies (and therefore should not be addressed as Nino3.4). What I mean is that that the time series seems to be obtained averaging raw SST in the Nino3.4 area, and afterwards normalized to 0 mean (and 1 variance?). It should be explained in the text and, if so, not referred to as NINO3.4 index. Moreover, it is not clear how the Niño and Niña events are defined, if using this timeseries or if they were taken from some other source. If they are taken from this timeseries, then most likely they are not 100% right.

5) It is unclear what the panel d) is representing, the main text reads "*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*". So the second part of the sentence is unclear, what is the "*precipitation amplitude….quasi-biennal scale*"? Why isn't the precipitation index used again as in the previous case? It would have been more consistent. Also, it is not clear here how the surrogates are computed, are they time-shift surrogates of the Nino3.4 or of the phases timeseries? I think the only reasonable choice is the first (if it does not produce abrupt discontinuities) because the latter for sure produces an abrupt discontinuity so wouldn't be a good surrogate at all for the phases time series.

6) Fig. 5 panels b) and e) are very confusing and misleading. The broken vertical axes hide the fact that the values are indeed very similar. The black and red bars to represent the differences make no sense to me, and perhaps those figures are not needed at all, and a table would be more readable. The label of panels c) and f) reports "histogram" which is rather odd, also because the figure does not look like an histogram. Regarding to this part, the value 0.264 is reported as significant in the text but the p-value is not given. Moreover, it should be probably be corrected (something like a Bonferroni correction) as here the authors are testing multiple differences, implicitly, 5 * 6 / 2 differences.

7) Fig. 8c is either warrying or I didn't understand what it represents. The space is supposed to be divided in a regular grid representing ERA5 grid cells. Then, in each grid cell, rainfall data is analyzed and the ENSO state when the peak precipitation occurs most often marks the cell color. If so, why the blue squares do not fit the grid? How have they been obtained? The situation in Fig. 9 is even worse.
8) Fig. 10. What I said for Fig. 5 panels b) and e) holds here too.
9) *"Although physical mechanisms explaining the observed cross-scale information transfers are yet to be established, the uncovered causal relations can already be used in machine learning tools for forecasting precipitation anomalies."* I think that the first sentence should be expanded. The authors should provide an idea for further investigations of what could be a possible mechanism that makes the low frequency of ENSO impact the monsoon at the 2ys scale, referring to the literature that surely has already investigated the link between ENSO and China monsoon from another angle. The second sentence is very vague and does not provide any valuable information.
10) If the QB mode is the one that impacts the dataset the most why, to make all figures up to Fig 5, it was used the 6ys mode instead?

Minor:
The sentence *"other recent follow-up studies were particularly concerned with the predictability and future of ENSO projections attributable to global to regional scale interconnections, including the combined influence of ENSO and PDO"* is unclear.

"The Yangtze River is China's longest river and the world's third largest, contributes considerably to China's equitable economic and ecological growth", I understand economic growth but what does ecological growth mean?

"The YZRB is predominantly controlled by Siberian northwest winter and southeast summer monsoon." Written like this it suggests the existence of a Siberian monsoon. Perhaps the authors are mentioning the Siberian High?

Fig. 2 carries almost no information. In the text, a lot of geographical details are given *"hydrological station (YHS) separates the Yangtze River into upper and lower sections and is renowned as the 'Gateway to the Three Gorges'. The Three Gorges Dam (TGD) lies just approximately 40 kilometers above (Xu et al., 2007). The territory above Yichang station is commonly regarded as the upper sub-basin of the YZRB; the region from Yichang station and Hukou station is the middle sub-basin; and the region under Hukou station is the lower sub-basin of the YZRB (Fang et al., 2018). The YZRB lies in subtropical and temperate climate zone dominated by monsoonal winds; the southern region exhibits subtropical climate while northern region presents temperate zone. Major flooding in YZRB is linked with warm ENSO and strong summer monsoons typically occur after El Niño conditions in the winter, while weak winter"* which are impossible to locate in the maps, making both the map and this paragraph useless.

Fig 3. To use red for precipitation is a semiotically unfortunate choice. Moreover, one cannot see the rivers.

Eq.1 Exp parenthesis is too small

The sentence *"The causal delay can be found in the causality analysis"* in that part of the text is particularly unclear.

Nino3.4 sometimes is written with ~ (often in the text) and sometimes not (e.g. label of fig.4). The authors should be consistent: either with or without. Sometimes, it is written with a coma ("Nino3,4") the point should be employed instead.

In Fig5. The panels would benefit of some labels.

Fig. 6 Time unit of measure should be years, plural, not year. The colorbar needs a label

Fig 7. "a)", "b)", "c)" and "d)" are barely visible as the rivers and their names. The crosses hide the information about the difference patterns because they are too heavy.

Fig. 8. And 9. Using the same color scale for maximum and minimum precipitation makes very hard to see any differences between the two maps (especially in 8), aside the mean level. Again, rivers colors are unfortunate. "c)" and "d)" labels are barely visible. In fig 9 the color scale of the legend of c and d should be periodic, as the numbers from 1 to 6 represent phases of the 6ys mode, so close numbers should have "close" colors and 1 and 6 should be regarded as close among them.

Fig. 10 Labels are oversized compared to panels, a line break or a smaller character dimension should be used.

In the sentence *"for example, in neuroscience, where cross-frequency phase–amplitude coupling has recently been observed in electrophysiological signals reflecting the brain dynamics"*, I think that mentioning neurosciences here is really out of context given the journal.