

## Reply to Reviewer 2

We would like to thank the reviewer for his/her patience and time spent reviewing the manuscript. The methodology has significantly changed in this revision in response to reviewer 3. The results are not qualitatively changed, but we have reduced the number of rain-gauge clusters from 20 to 6. This was done by merging some clusters together using a larger minimum cluster size parameter.

1- Line 16-17: “The correlation patterns for the first two SST modes were somewhat similar between d4PDF and rain-gauges, but were biased spatially and magnitude-wise.”: The correlation patterns for the second mode don’t look similar between observation and d4PDF (Figs. 5b, 6b, 7b, and Table 2). So, I think the authors need to revise this sentence.

I may be biased psychologically. This sentence has been removed.

2- Line 20: “water vapour flux,.” -> “water vapour flux.”

The error has been corrected.

3- Line 243: I’m afraid that the threshold of CC is 0.254 in the case of a two-tailed test with the significance level of 0.05 and N=60. Please check it.

We have neglected to state that this was a one-tail test, which has a threshold of 0.22 for N=60 (0.217 for N=59). Since this is not the commonly understood meaning, and  $\alpha=0.10$  may be too large for a two-tailed test, we have adjusted the definitions of the correlation (R) strength. The new definitions are “no relation” for  $|R| < 0.25$ , “weak” for  $0.25 \leq |R| < 0.50$ , “moderate” for  $0.50 \leq |R| \leq 0.75$ , “strong” for  $0.75 \leq |R| \leq 1.0$ , and “perfect” for  $|R|=1.0$ . The old definition “very strong” was never used. The colormap of the plots have been accordingly adjusted. Figure R2.3.1 below shows how a plot changes from the original colormap (panel c), to one using the two-tailed significance value as the threshold for shading (panel b), to the new definition (panel a). The difference between original and new versions is mainly due to the colour threshold from 0.20 to 0.25. Figures R3.3.2A–R3.3.5B show comparisons of figures in the old manuscript and the new manuscript. New figures are on the left column and old figures are on the right column. Clusters are different from the d4PDF plots because of the revised clustering method, based on Reviewer 3’s comments. Also, the mean of clusters is now used instead of the max of clusters, based on Reviewer 3’s comments.

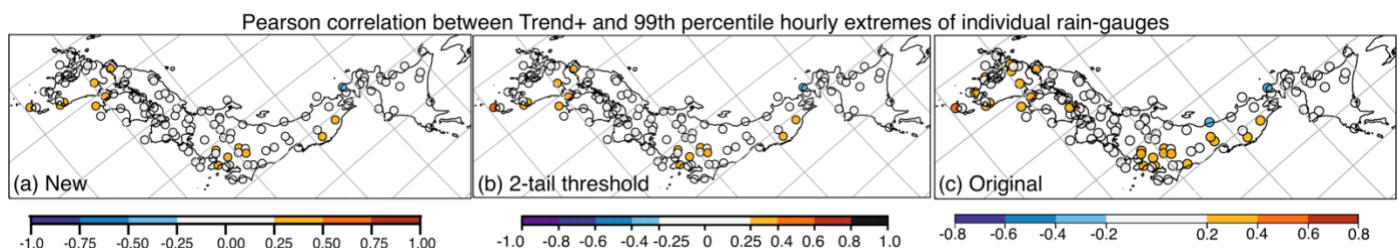
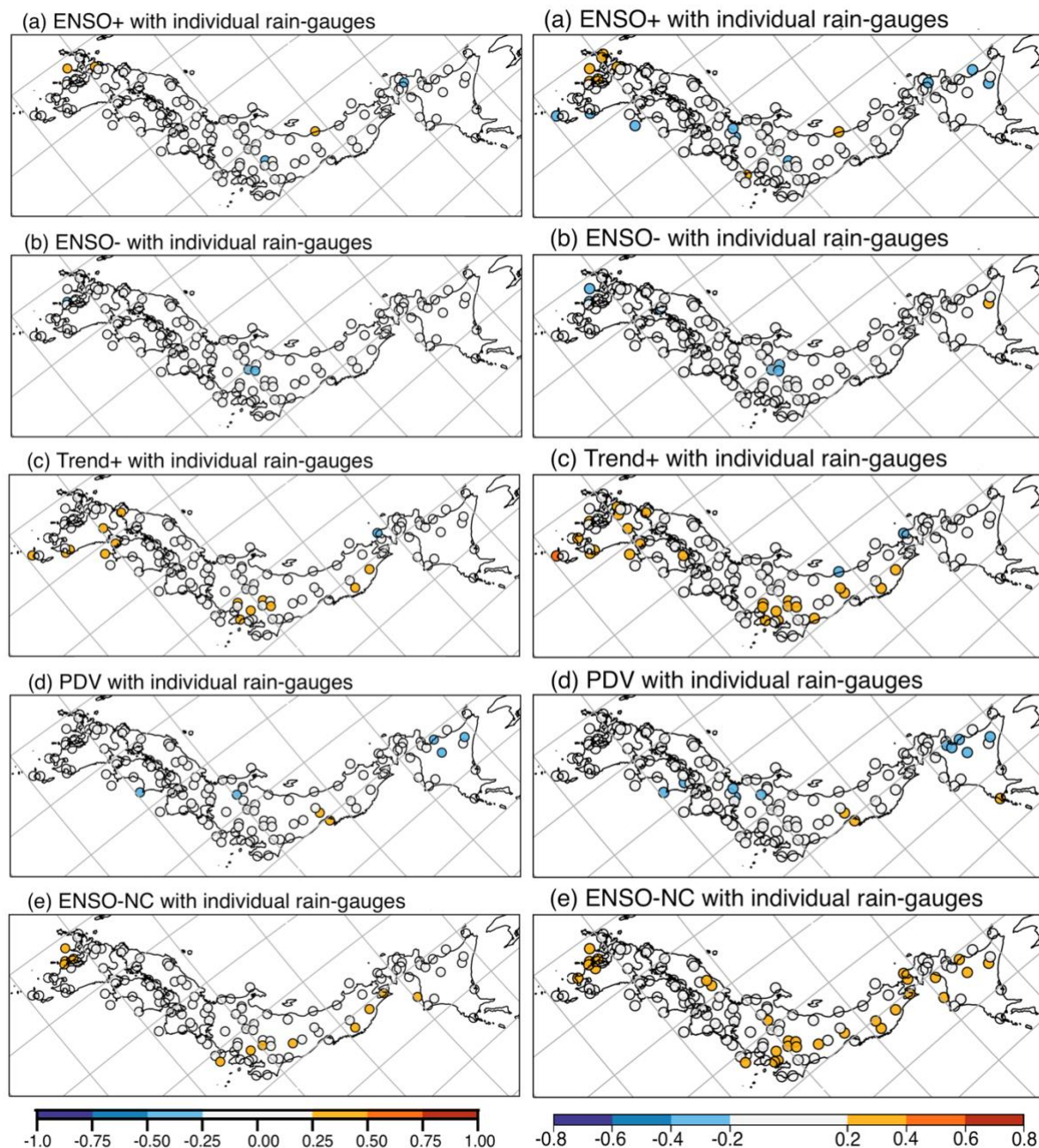


Figure R2.3.1.

**Panels from Figure 6 of new manuscript**

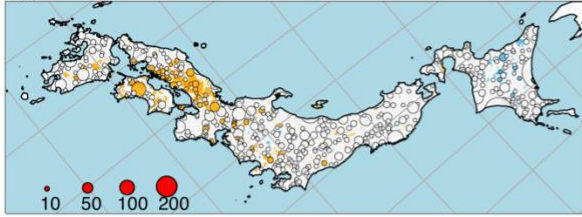
**Panels from Figure 5 of old manuscript**



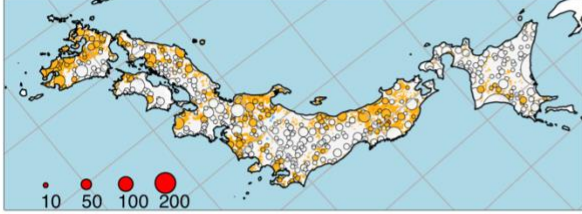
**Figure R2.3.1A**

**Panels from Figure 6 of new manuscript**

(f) ENSO+ with d4PDF



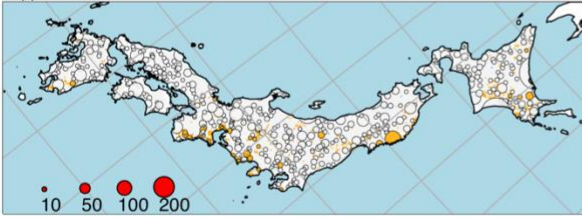
(g) ENSO- with d4PDF



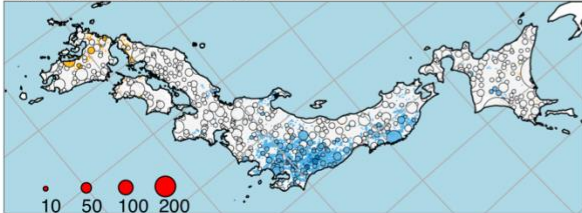
(h) Trend+ with d4PDF



(i) PDV with d4PDF

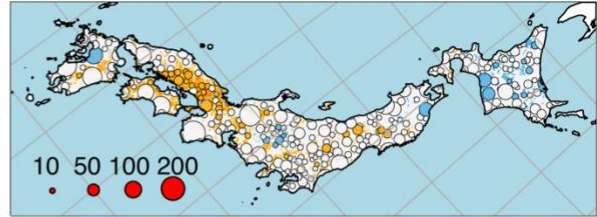


(j) ENSO-NC with d4PDF

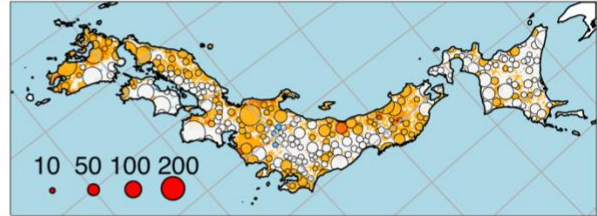


**Panels from Figure 5 of old manuscript**

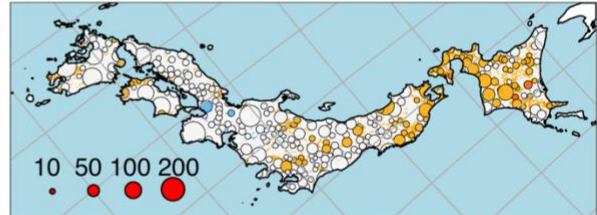
(f) ENSO+ with d4PDF



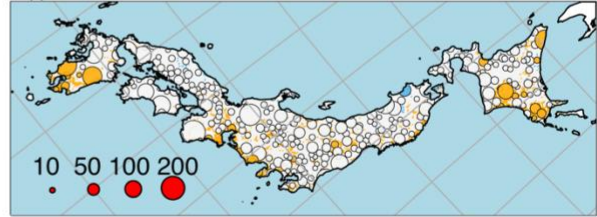
(g) ENSO- with d4PDF



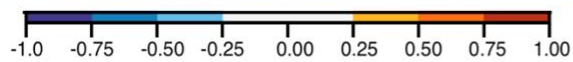
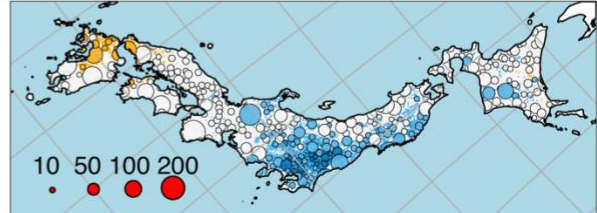
(h) Trend+ with d4PDF



(i) PDV with d4PDF



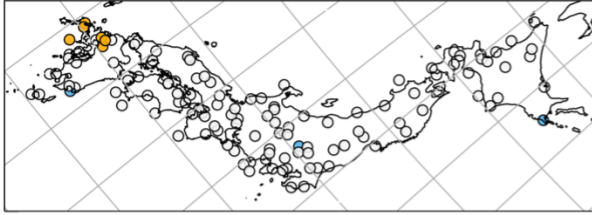
(j) ENSO-NC with d4PDF



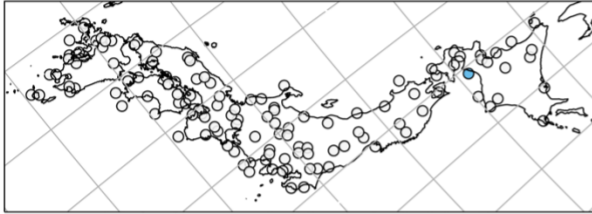
**Figure R2.3.1B**

**Panels from Figure 7 of new manuscript**

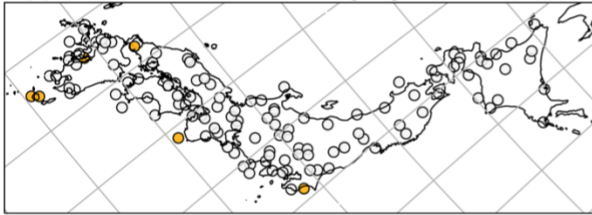
(a) ENSO+ with individual rain-gauges



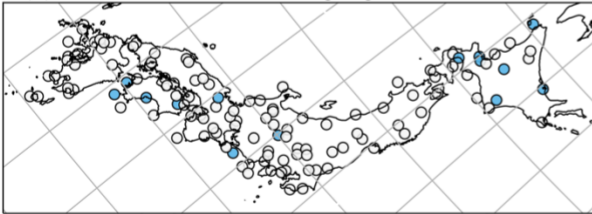
(b) ENSO- with individual rain-gauges



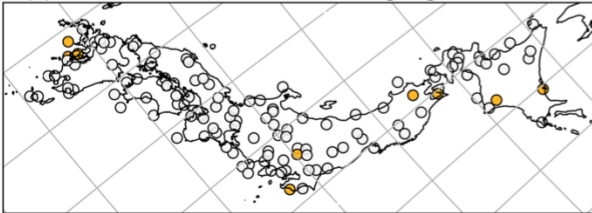
(c) Trend+ with individual rain-gauges



(d) PDV with individual rain-gauges

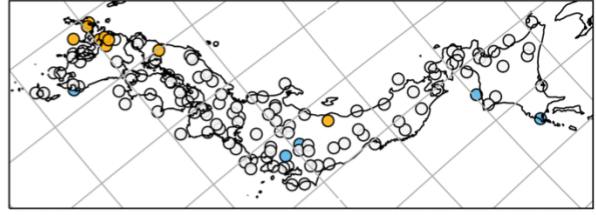


(e) ENSO-NC with individual rain-gauges

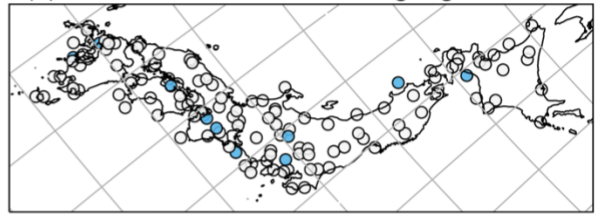


**Panels from Figure 6 of old manuscript**

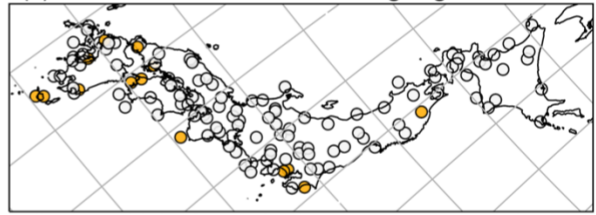
(a) ENSO+ with individual rain-gauges



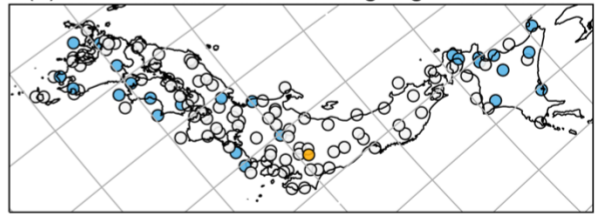
(b) ENSO- with individual rain-gauges



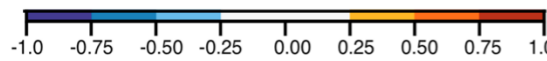
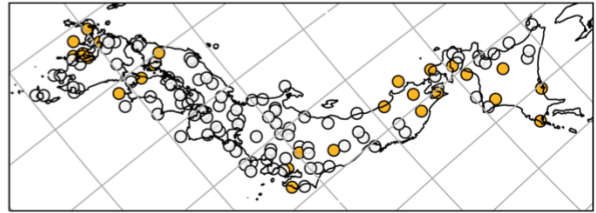
(c) Trend+ with individual rain-gauges



(d) PDV with individual rain-gauges



(e) ENSO-NC with individual rain-gauges



**Figure R2.3.2A**

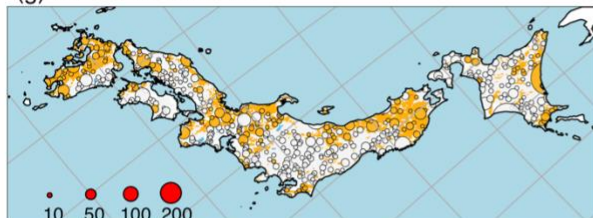
Panels from Figure 7 of new manuscript

Panels from Figure 6 of old manuscript

(f) ENSO+ with d4PDF



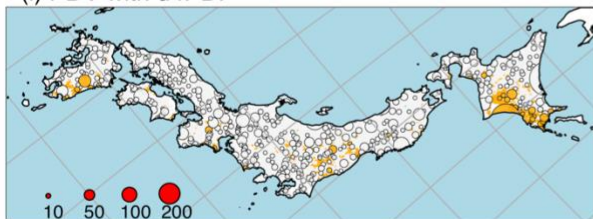
(g) ENSO- with d4PDF



(h) Trend+ with d4PDF



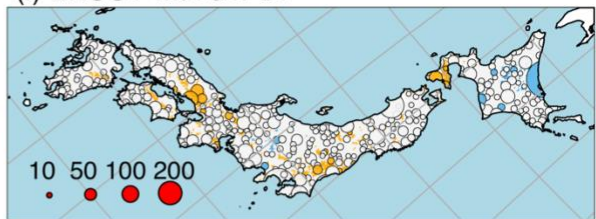
(i) PDV with d4PDF



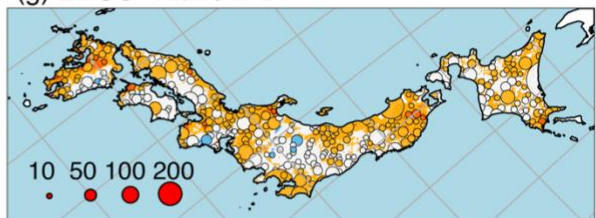
(j) ENSO-NC with d4PDF



(f) ENSO+ with d4PDF



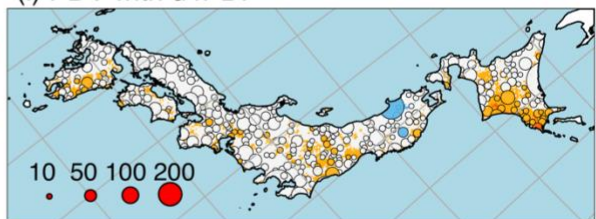
(g) ENSO- with d4PDF



(h) Trend+ with d4PDF



(i) PDV with d4PDF



(j) ENSO-NC with d4PDF

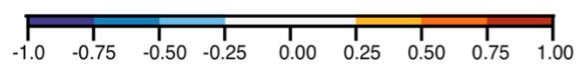
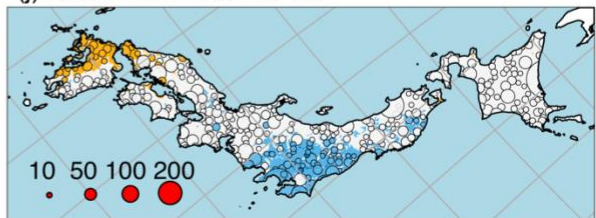
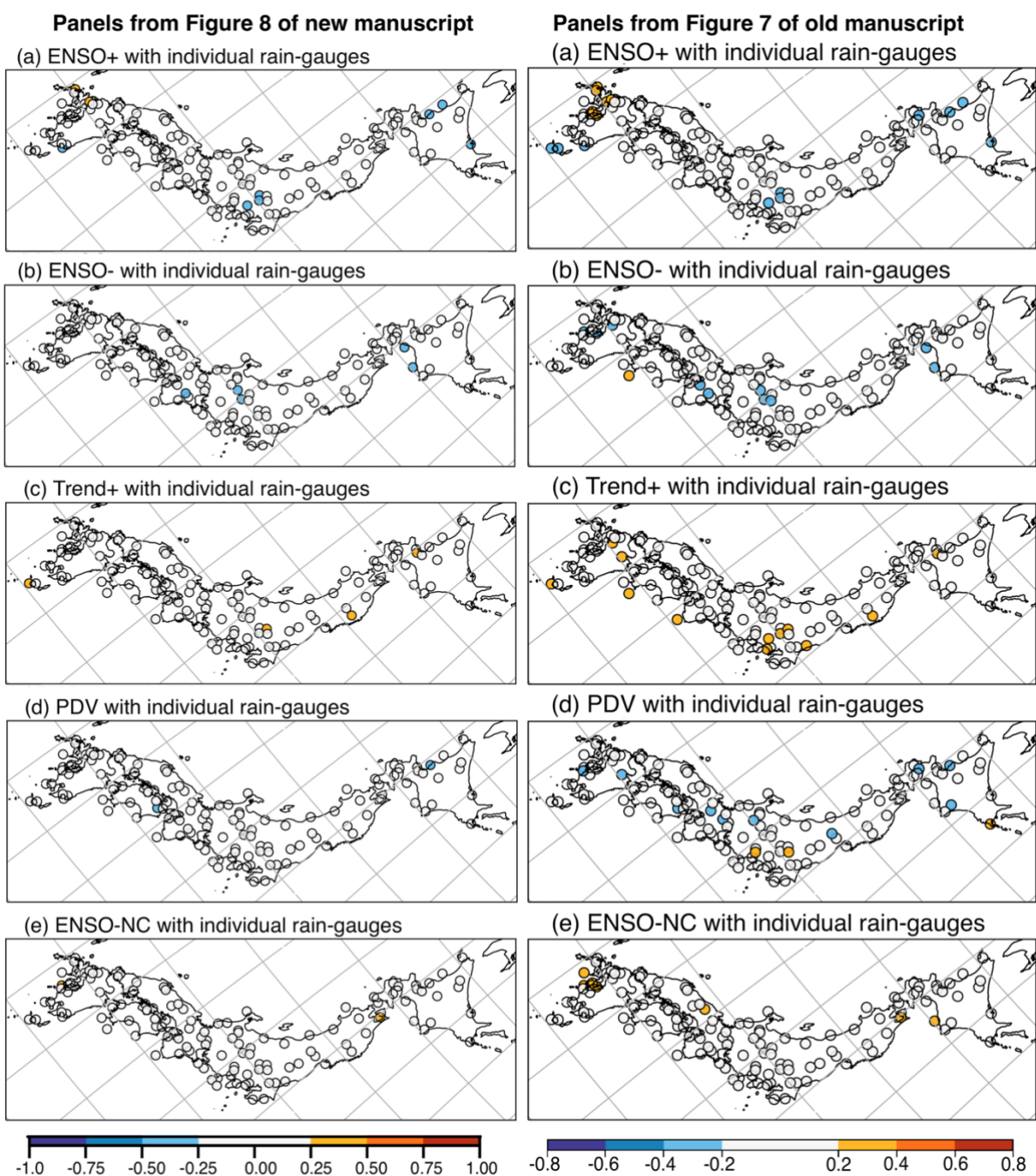
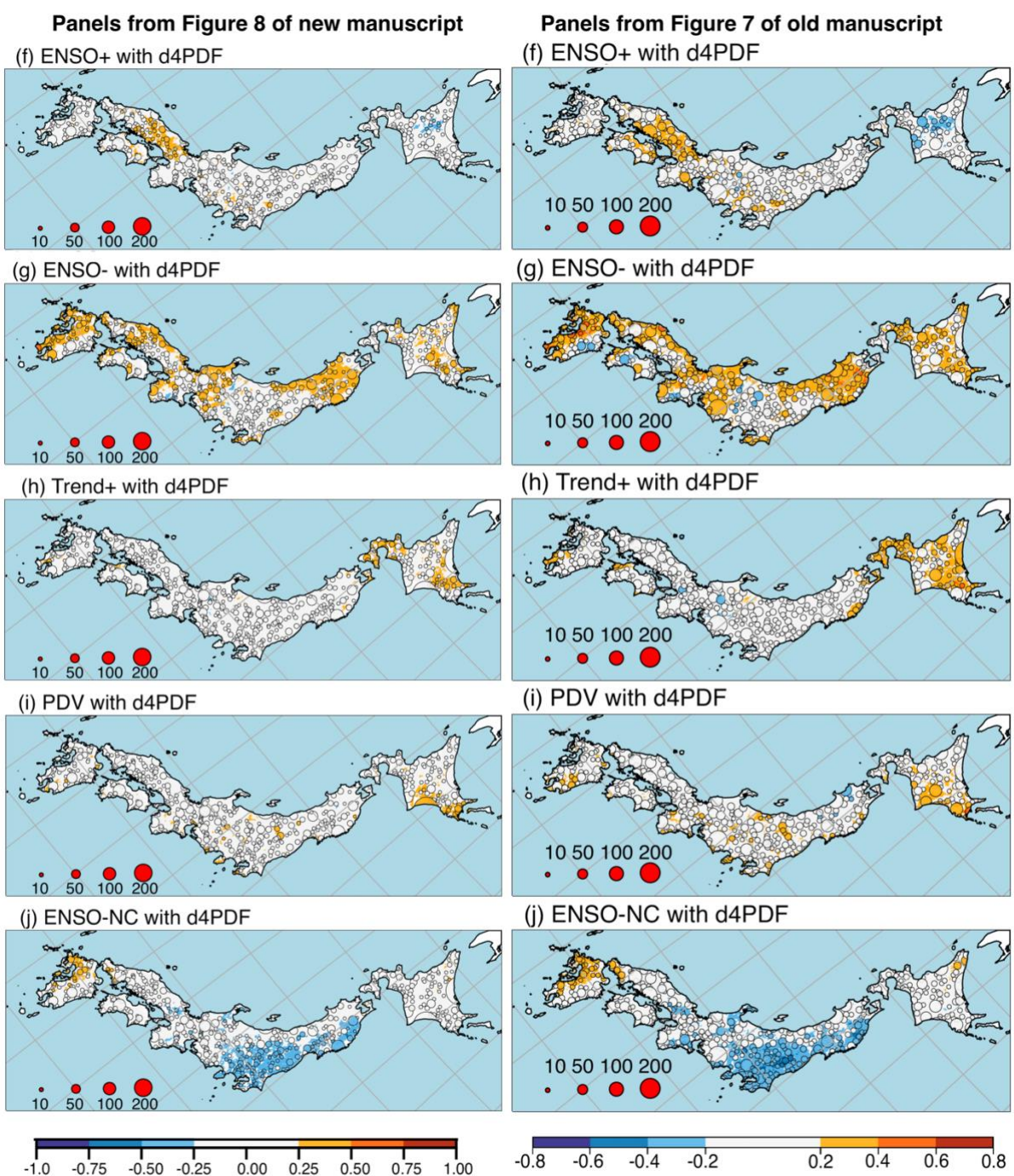


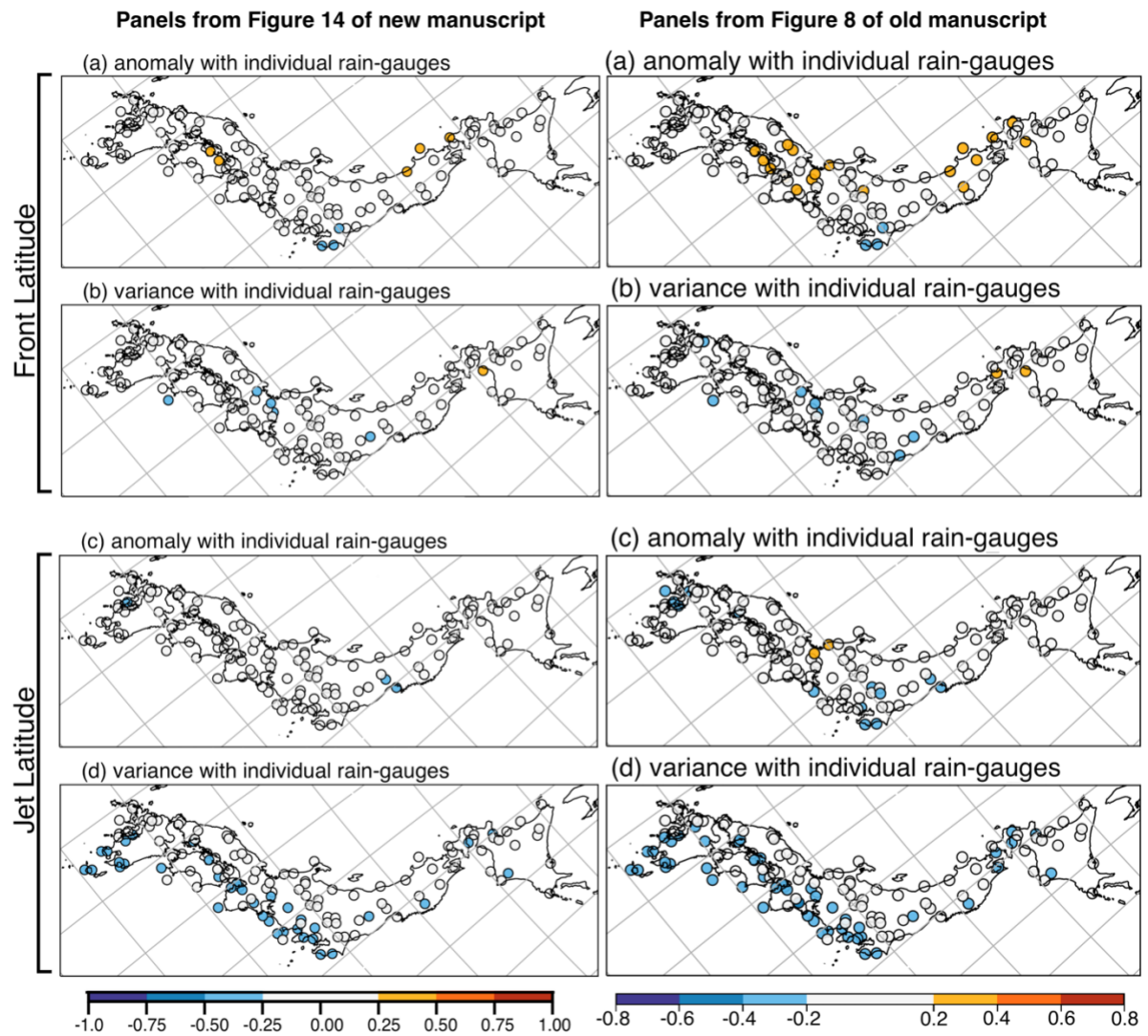
Figure R2.3.2B



**Figure R2.3.3A**



**Figure R2.3.3B**



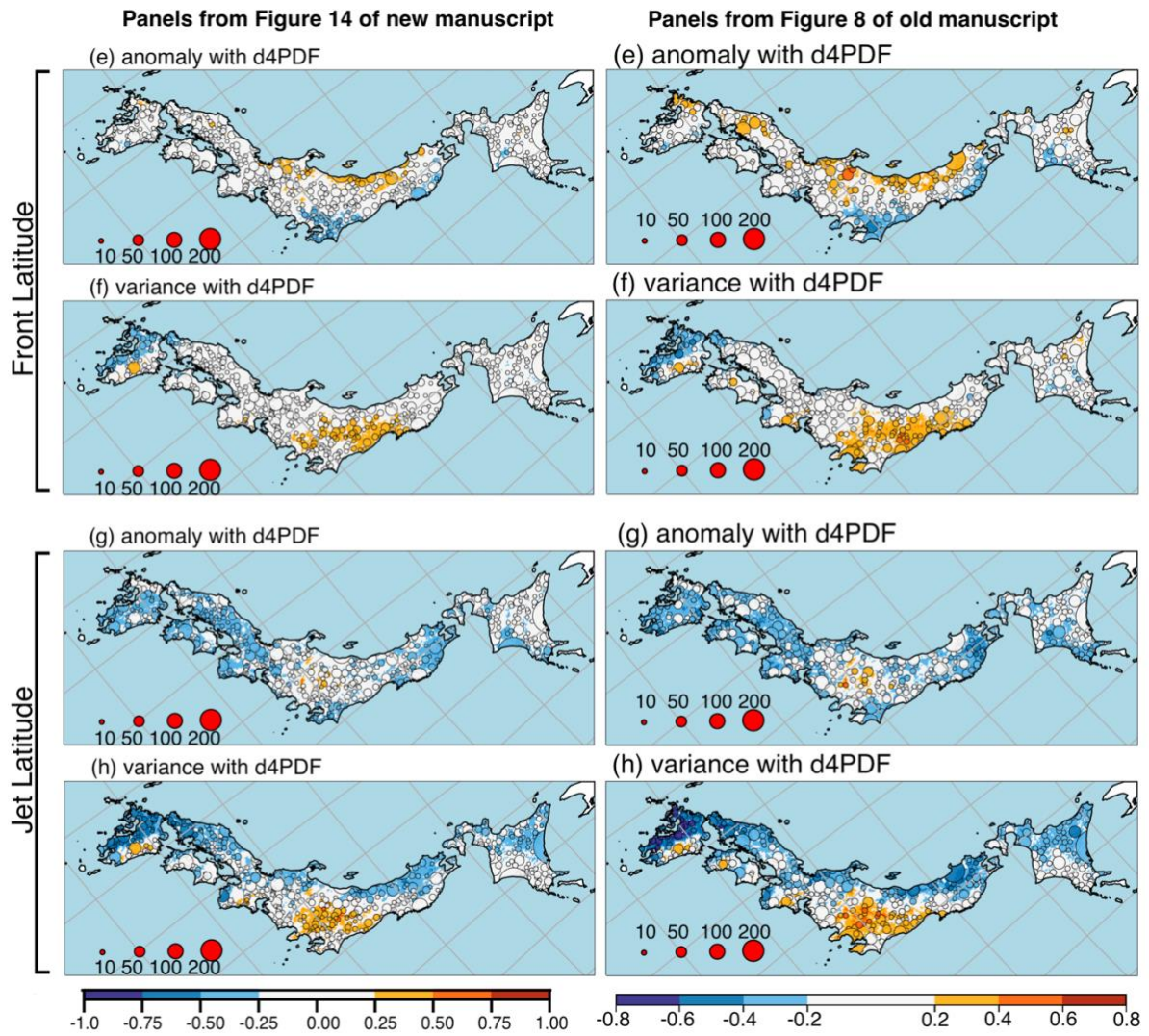


Figure R2.3.4B

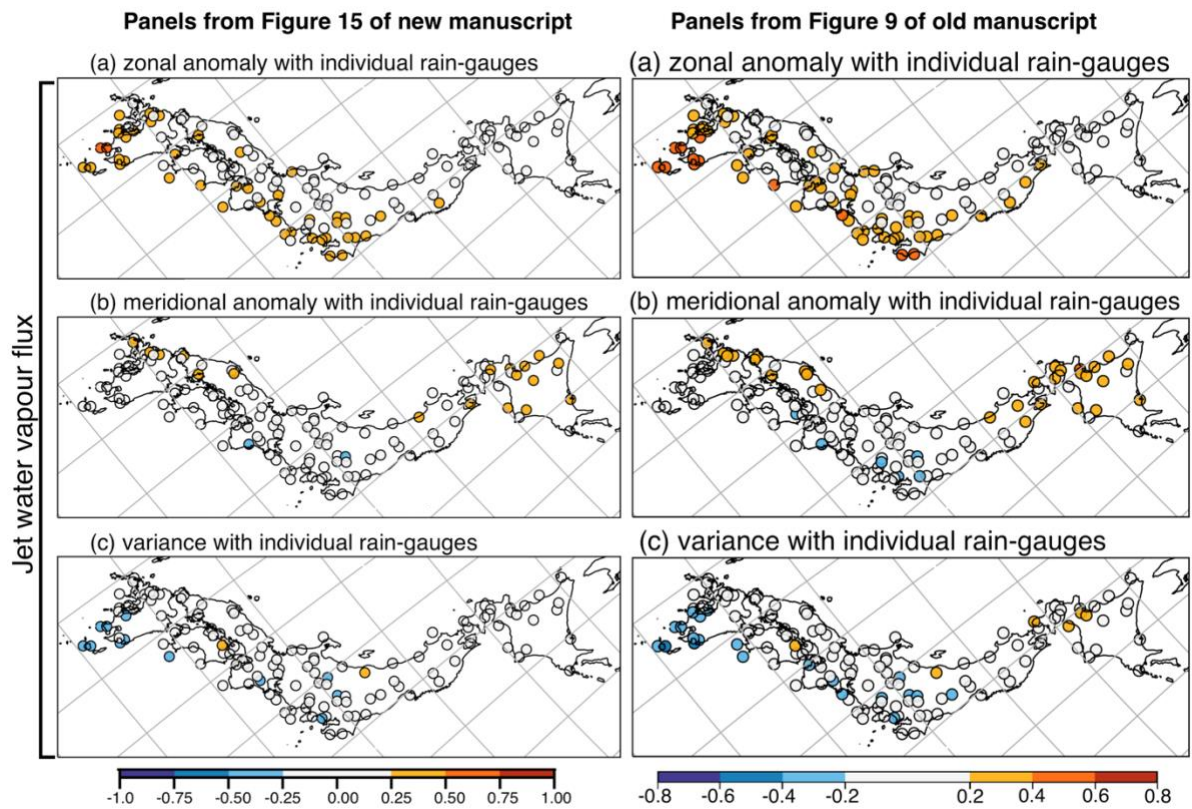


Figure S2.3.5A

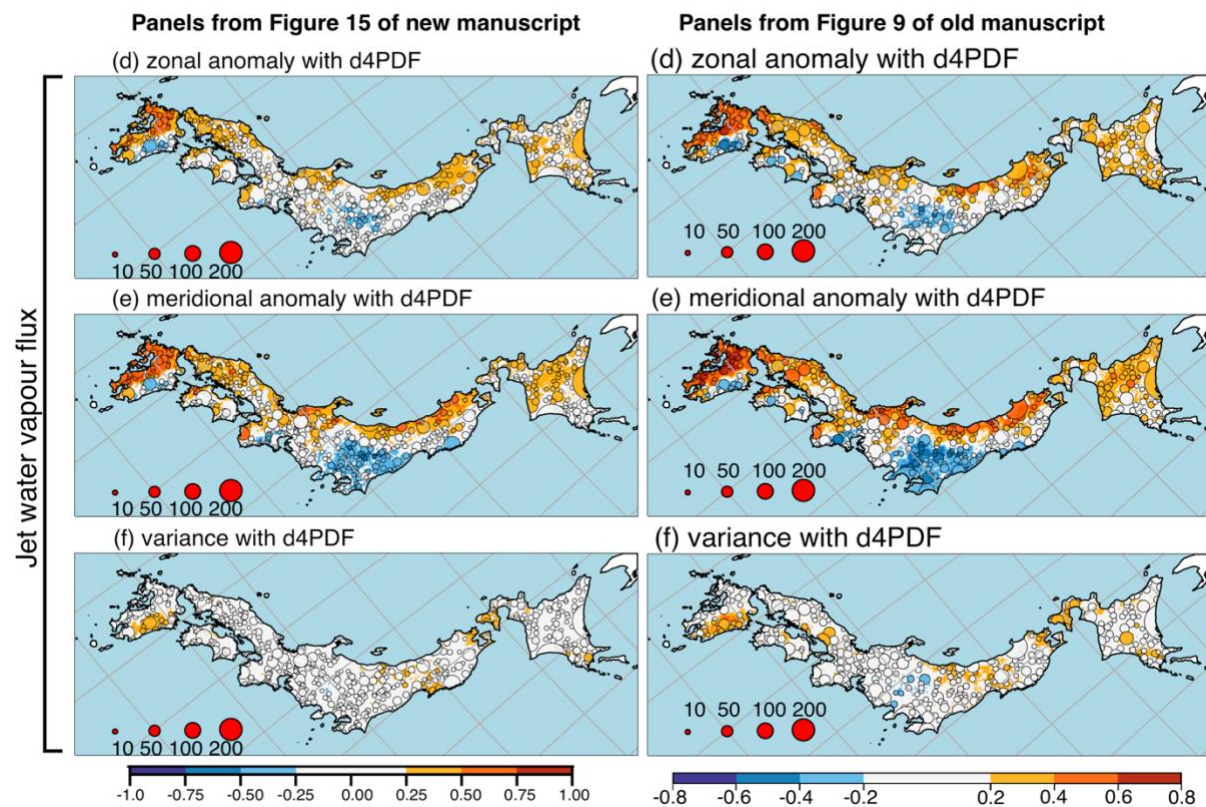


Figure R2.3.5B

4 - Line 268: For the denominator of the WVflux formula, “pi - pi+1” instead of “pi-l - pi”?

The error has been corrected.

5- Line 279: “For each year, ---” It would be better to insert a line break before this sentence.

A line break has been added.

6- Around Line 505: I think description about the relationship between ENSO+ and monsoon indices should be added here.

Lines 697–716 have been added to discuss ENSO+. (The Discussion section has been reorganised and expanded.)

7- Table 2: There are too many numbers listed, making it difficult to understand the feature intuitively. How about presenting them as a bar graph?

The clustering method has been changed but the clusters produced by the new method are similar to the original manuscript. The value of  $c_{min}$  was then increased until smaller clusters merged into 7 clusters. One more cluster was added from  $c_{min}=5$  labelled cluster 8. The clusters are plotted in Fig. S2.7.1. The correspondence between the original clusters and the new clusters is shown in Table S2.7.1. The correlation values are now represented as box and whisker plots, Figs 10, 12 and 19 of the new manuscript.

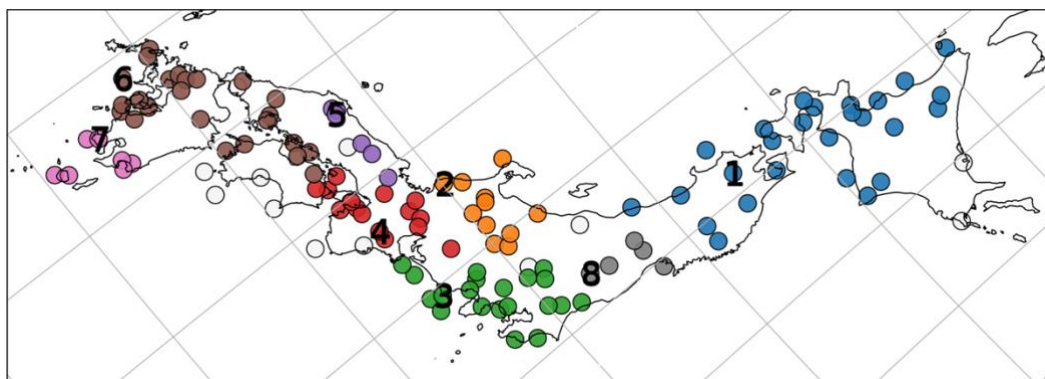


Figure S2.7.1

Table S2.7.1

Old cluster	Original	New	New cluster
1	Asahikawa, Haboro, Kitami-esashi, Omu, Rumoi, Wakkanai	Asahikawa, Haboro, Kitami-esashi, Omu, Rumoi, Wakkanai,	1
2	Iwamizawa, Kutchan, Otaru, Sapporo, Suttsu	Iwamizawa, Kutchan, Otaru, Sapporo, Suttsu,	
3	Tomakomai, Urakawa	Tomakomai, Urakawa, Muroran, Obihiro,	
4	Esashi, Hakodate, Mutsu	Esashi, Hakodate, Mutsu, Aomori, Fukaura,	
5	Akita, Morioka, Sakata	Akita, Morioka, Sakata, Kushiro, Miyako, Hachinohe	
6	Fukushima, Ishinomaki, Sendai	Fukushima, Ishinomaki, Sendai, Yamagata, Shirakawa	8
7	Choshi, Katsuura, Mito, Tateno	Choshi, Katsuura, Mito, Tateno,	3
8	Tokyo, Yokohama, Kumagaya	Tokyo, Yokohama, Ajiro, Mishima, Shizuoka, Hamamatsu, Omaezaki, Irozaki, Kumagaya, Okunikko, Chichibu, Kofu, Lake Kawaguchi,	
9	Chichibu, Kofu, Lake Kawaguchi		
10	Fushigi, Kanazawa, Takata, Toyama	Fushigi, Kanazawa, Takata, Toyama,	2
11	Fukui, Takayama, Tsuruga	Fukui, Takayama, Wajima, Karuizawa, Suwa, Matsumoto, Nagano	
12	Gifu, Hikone, Kyoto, Nagoya	Tsuruga, Iida, Gifu, Hikone, Kyoto, Nagoya,	4
13	Himeji, Kobe, Nara, Osaka, Tsu, Ueno	Himeji, Kobe, Nara, Osaka, Tsu, Ueno,	
14	Sumoto, Tokushima, Wakayama	Sumoto, Tokushima, Wakayama	
16	Matsue, Sakai, Yonago	Matsue, Sakai, Yonago, Tottori, Toyooka, Maizuru	5
15	Fukuyama, Hiroshima, Kure, Okayama	Fukuyama, Hiroshima, Kure, Okayama, Tadotsu, Takamatsu,	6
17	Hagi, Shimonoseki	Hagi, Shimonoseki, Fukuoka,	
28	Hirado, Hita, Iizuka, Saga, Sasebo	Hirado, Hita, Iizuka, Saga, Sasebo,	
29	Kumamoto, Mount Unzen, Nagasaki	Kumamoto, Mount Unzen, Nagasaki, Akune, Hitoyoshi, Ushibuka, Hamada, Matsuyama, Oita, Uwajima	
20	Aburatsu, Kagoshima, Makurazaki, Miyazaki, Miyakonojo	Aburatsu, Kagoshima, Makurazaki, Miyazaki, Miyakonojo, Tanegashima, Yakushima	7

## Reply to Reviewer 3

We would like to thank the reviewer for his/her insightful comments on clustering. Most significantly the metric used for clustering has been revised according to their suggestion, and almost all of the results have been recalculated. The Discussion has also been expanded on.

### Major points

- 1) Rewrite some parts of the main manuscript.

The manuscript has been mostly rewritten based on the reviewer's comments.

- 2) Explore further the results linking extremes with monsoon and Pacific indices providing physical reasons.

The Discussion has been rewritten, to describe how the extremes-monsoon and monsoon-Pacific relationships combine to produce the extremes-Pacific relationships. An explanation is now provided for the difference between model and observation, which is that the model showed a the two week delay in northward movement of the monsoon.

### Corrections

- 1) Line 15. Are the modes sorted by decreasing explained variance? Clarify.

The abstract now lists the modes in order of explained variance for the 1952–2010 period of analysis with ENSO-NC listed first then PDV. This is now explicitly stated in L15–17. The order of the 4th and 5th modes differed depending on the analysis period. For longer time periods, the PDV mode accounted for more of the variance than ENSO-NC, but for the 1952–2010 period ENSO-NC accounted for more variance than PDV.

In line 18 'higher modes' refer to the previous order of modes?

The abstract now states the modes explicitly.

- 2) Line 17-20. Rewrite the phrase of lines 17-20 in a much clear way by splitting it into two sentences.

The sentence has been changed to "correlations between the monsoon indices and standardised principal components of the Pacific SST modes were calculated."

- 3) Line 95 Clarify the simulation period of the '100-member historical-warming (HPB) climate ensemble' as well as the temporal and spatial resolution of the referred GCM.

The spatial and temporal resolutions of the GCM and RCM are now both stated in L98–99.

- 4) Lines 99-100. No systematic comparison is made between the distribution of observed extremes and the distribution of simulated extremes by the 4DPDF dataset. A synthetic study of the representativeness of 4DPDF extremes shall be made.

A comparison for 6 analysis regions has been added in Section 3.3, L487–499 and Figure 6. Verification of the d4PDF rainfall extremes was indeed done first before calculating relationships with SST modes. This was inside an earlier version of the manuscript, but then removed at the request of an earlier reviewer. A detailed comparison for 83 regions using radar is currently under review.

- 5) Line 142 and wherever needed: Change 'Principle Component Analysis' to 'Principal Component Analysis', throughout all the manuscript.

The error has been corrected.

- 6) Line 144. The linear long trend of SST was not removed to get the anomalous SST. It was obtained farther as a mixed mode (TREND+) or pseudo-trend. Authors could be more direct by correlating the linear SST trend (for instance averaged throughout all Japan) with extremes. Comment on that.

A comparison has been made with a smoothed Trend+ mode in Section 3.3 L550–569 and Figs 11–12. A 11-year moving average had a long enough window (by eye) to sufficiently smooth the scores.

As for why the SSTs were not detrended first, our viewpoint is that results should emerge naturally from the PCA of the data without any assumption of warming (even though we know there is warming). Also, if a linear trend was first calculated, the trend might change depending on the time period chosen. This would make the comparison of different time periods in Section 3.1 more complicated. I think Trend+ reflects the difficulty in distinguishing the warming trend from decadal/multi-decadal variability.

We did assume that the effect of warming on the seasonal cycle was small, when the long-term mean seasonal cycle was removed. The first few attempts at ePCA tried to include the seasonal cycle, but the results were too complicated to work with.

Figures R3.6.1A and R3.6.1B below compare plots of correlation done with smoothed and original Trend+. Left columns show smoothed and right columns show original Trend+. Rows show the temporal resolution of the extremes.

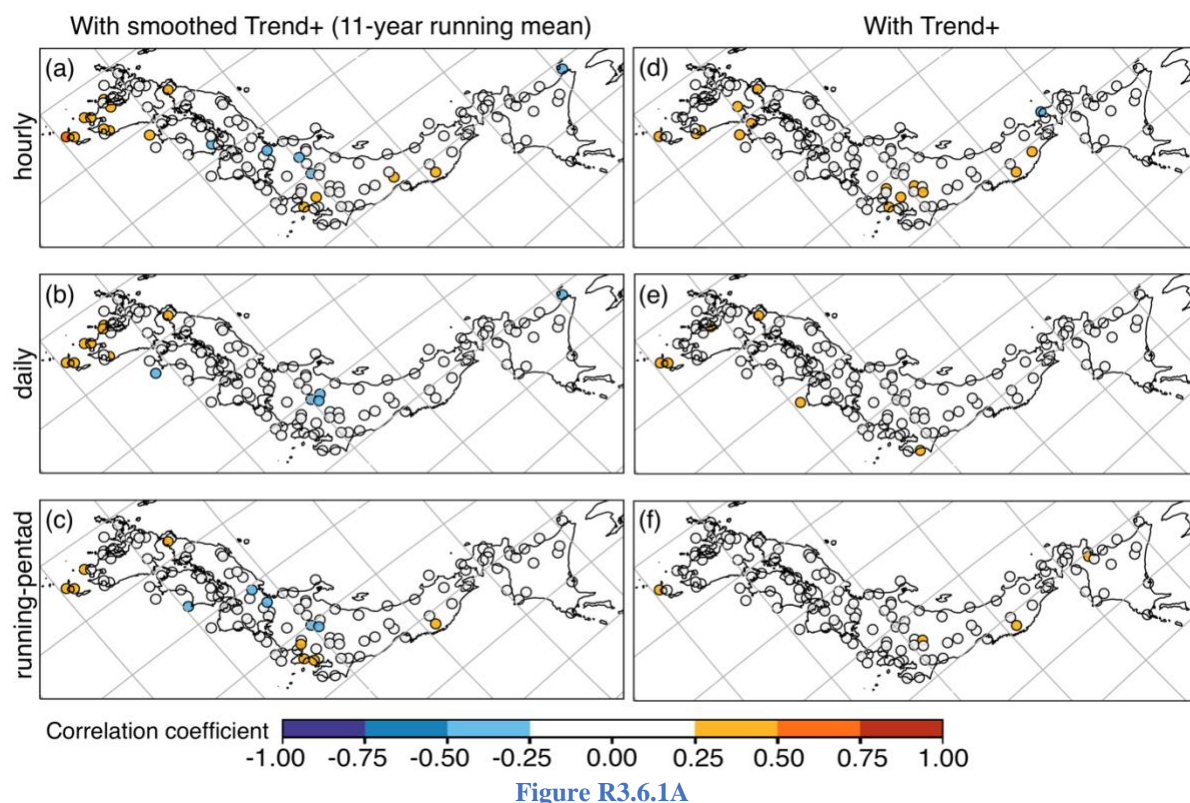


Figure R3.6.1A

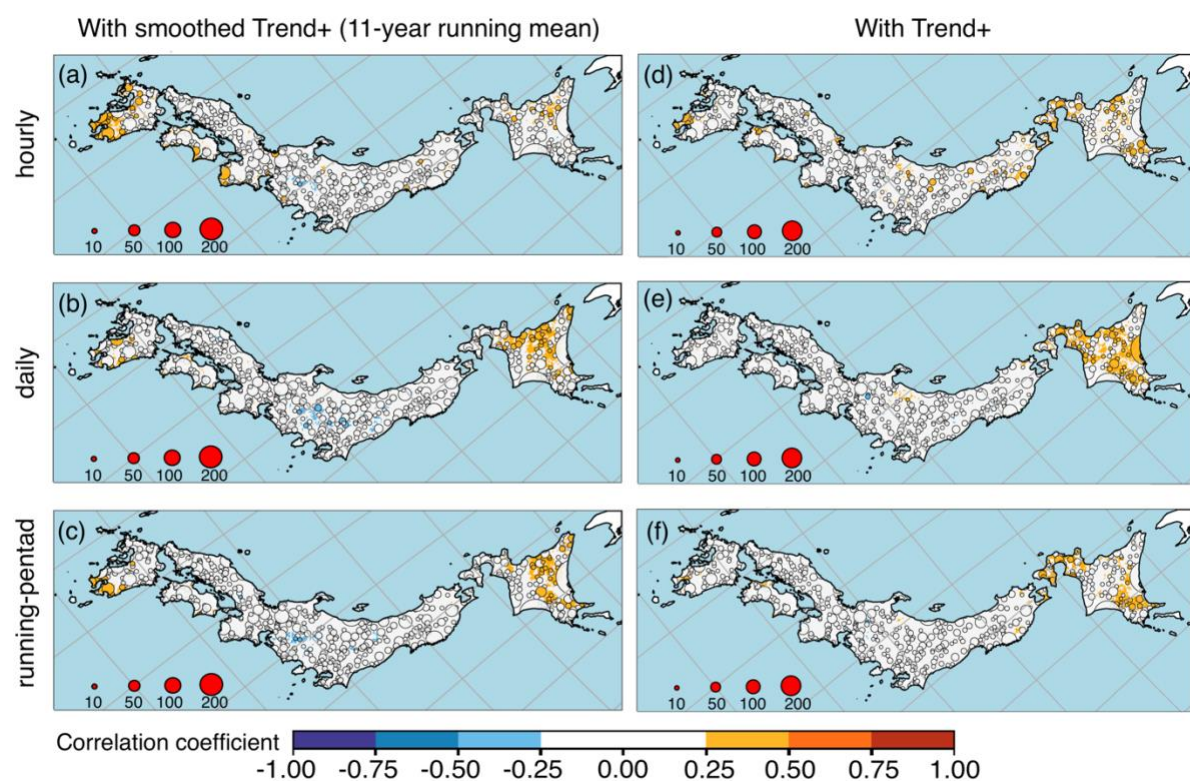


Figure R3.6.1B

7) Line 146 Authors have performed extended-PCA (extended Principal Component Analysis) and not a simple PCA. This must be referred explicitly and for clarity, since an extended vector merging 5 delayed (5 trimesters centered in JJA) of spatially distributed values have been taken, from which the covariance and its eigen-decomposition was computed. This is like a MSSA (Multi Singular Spectrum Analysis) with embedding dimension 5 and trimestral sampling.

A statement has been added in L155–156 to clarify this. L157 and L160 now uses “extended PCA”. Later mentions of “PCA” in the text have been replaced with “rePCA” (rotated extended Principal Component Analysis).

8) Line 168 The spatial clustering of temporal extremes, particularly the rainfall (e.g. Ma et al. 2020 and references therein) has been studied by different authors, namely by using geostatistical techniques. The reference to those works must be included in the manuscript. The considered metric (distance) to cluster rain-gauges and grid points, used in cluster analysis HDBSCAN depends uniquely on the temporal similarity (Spearman rank correlation) between time-series, being equivalent to the F-madogram. However, for spatially distributed data, a geometrical term, weighing the point-wise distance (e.g. Euclidean) must be added to the statistical distance. The omission of the geometrical term leads to fragmented, topologically complex clusters (even not simply connected, i.e. with ‘holes’). This apparently was remedied by adopting ad-hoc clustering rules. The authors are asked to comment on that.

Yingzhao Ma, Mengqian Lu, Cameron Bracken, Haonan Chen, 2020. Spatially coherent clusters of summer precipitation extremes in the Tibetan Plateau: Where is the moisture from?, Atmospheric Research, Volume 237, 104841, ISSN 0169-8095, <https://doi.org/10.1016/j.atmosres.2020.104841>.

References to previous work have been added in L214–216. The lack of references was not done on purpose and due to ignorance. We came up with the metric with the ad-hoc rules after we intuited that k-means clustering was unsuitable for extremes.

Almost of all the results presented in the manuscript have been recalculated with a new metric that added a geometrical term. It also includes different frequency bands, based on the reviewer’s comment 12. Table R3.8.1 below shows clusters from the original manuscript (first column), after adding a geometric term (second column), and using frequency decomposition (third column). The minimum cluster size of  $c_{\min}=3$  was used. The clusters were similar between the three methods. Increasing  $c_{\min}$  merged the clusters. Figure R3.8.1 shows the dendrogram. We cut at  $c_{\min}=6$  and add cluster  $f$  from  $c_{\min}=5$  to get 8 clusters. More comparisons between the old hourly/daily/running-pentad method and the new method are shown the reply to reviewer’s comment 12.

**Table R3.8.1.** Bold text indicates changes from the first column, red text indicates changes from second column.

From merging results of hourly, daily, and running-pentad (Original manuscript)		Integrated over wavelet decomposition (New)
Triangulation (original manuscript)	Geometric term added (New)	
Asahikawa, Haboro, Kitami-esashi, Omu, Rumoi, Wakkanai	Asahikawa, Haboro, Kitami-esashi, Omu, Rumoi	Asahikawa, Haboro, Kitami-esashi, Omu, Rumoi, Wakkanai
Iwamizawa, Kutchan, Otaru, Sapporo, Suttsu	Iwamizawa, Kutchan, Otaru, Sapporo, Suttsu	Iwamizawa, Kutchan, Otaru, Sapporo, Suttsu
Tomakomai, Urakawa		Tomakomai, Urakawa, <b>Muroran</b> , <b>Obihiro</b>
Esashi, Hakodate, Mutsu	Esashi, Hakodate, <b>Muroran</b>	Esashi, Hakodate, Mutsu, <b>Aomori</b> , <b>Fukaura</b>
Akita, Morioka, Sakata	Akita, Morioka, Sakata	Akita, Morioka, Sakata
Fukushima, Ishinomaki, Sendai	Fukushima, Ishinomaki, Sendai, <b>Yamagata</b>	Fukushima, Ishinomaki, Sendai, <b>Yamagata</b> , <b>Shirakawa</b>
Choshi, Katsuura, Mito, Tateno	Choshi, Katsuura, Mito, Tateno,	Choshi, Katsuura, Mito, Tateno,
Tokyo, Yokohama, Kumagaya	Tokyo, Yokohama, Kumagaya	Tokyo, Yokohama,
	<b>Ajiro, Mishima, Shizuoka</b>	<b>Ajiro, Mishima, Shizuoka</b>
	<b>Hamamatsu, Omaezaki</b>	<b>Hamamatsu, Omaezaki, Irozaki</b>
Chichibu, Kofu, Lake Kawaguchi	<i>Group exists, but unclassified for daily extremes</i>	Kumagaya, <b>Okunikko</b> , Chichibu, Kofu, Lake Kawaguchi,
		<b>Karuizawa, Suwa, Matsumoto</b>
Fushigi, Kanazawa, Takata, Toyama	Fushigi, Kanazawa, Takata, Toyama	Fushigi, Kanazawa, Takata, Toyama,
Fukui, Takayama, Tsuruga		Fukui, Takayama, <b>Wajima</b>
Gifu, Hikone, Kyoto, Nagoya	Gifu, Hikone, Kyoto, Nagoya,	Tsuruga, <b>Iida</b> , Gifu, Hikone, Kyoto, Nagoya,
Himeji, Kobe, Nara, Osaka, Tsu, Ueno	Himeji, Kobe, Nara, Osaka, Ueno	Himeji, Kobe, Nara, Osaka, Tsu, Ueno
Sumoto, Tokushima, Wakayama	Sumoto, Tokushima, Wakayama	Sumoto, Tokushima, Wakayama
Fukuyama, Hiroshima, Kure, Okayama	Fukuyama, Hiroshima, Kure, Okayama	Fukuyama, Hiroshima, Kure, Okayama,
Hagi, Shimonoseki	Hagi, Shimonoseki,	Shimonoseki, <b>Fukuoka</b> , Iizuka
Hirado, Hita, Iizuka, Saga, Sasebo	Hita, Iizuka, Saga, Sasebo	Hirado, Hita, Saga, Sasebo
Matsue, Sakai, Yonago	Matsue, Sakai, Yonago, <b>Tottori</b>	Matsue, Sakai, Yonago, <b>Tottori</b> , <b>Toyooka</b> , <b>Maizuru</b>
Kumamoto, Mount Unzen, Nagasaki	Kumamoto, Mount Unzen, Nagasaki	Kumamoto, Mount Unzen, Nagasaki
Aburatsu, Kagoshima, Makurazaki, Miyazaki, Miyakonojo	Aburatsu, Kagoshima Makurazaki, Miyazaki, Miyakonojo	Aburatsu, Kagoshima, Makurazaki, Miyazaki, Miyakonojo, <b>Tanegashima</b> , <b>Yakushima</b>
		<b>Akune</b> , <b>Hitoyoshi</b> , <b>Ushibuka</b>

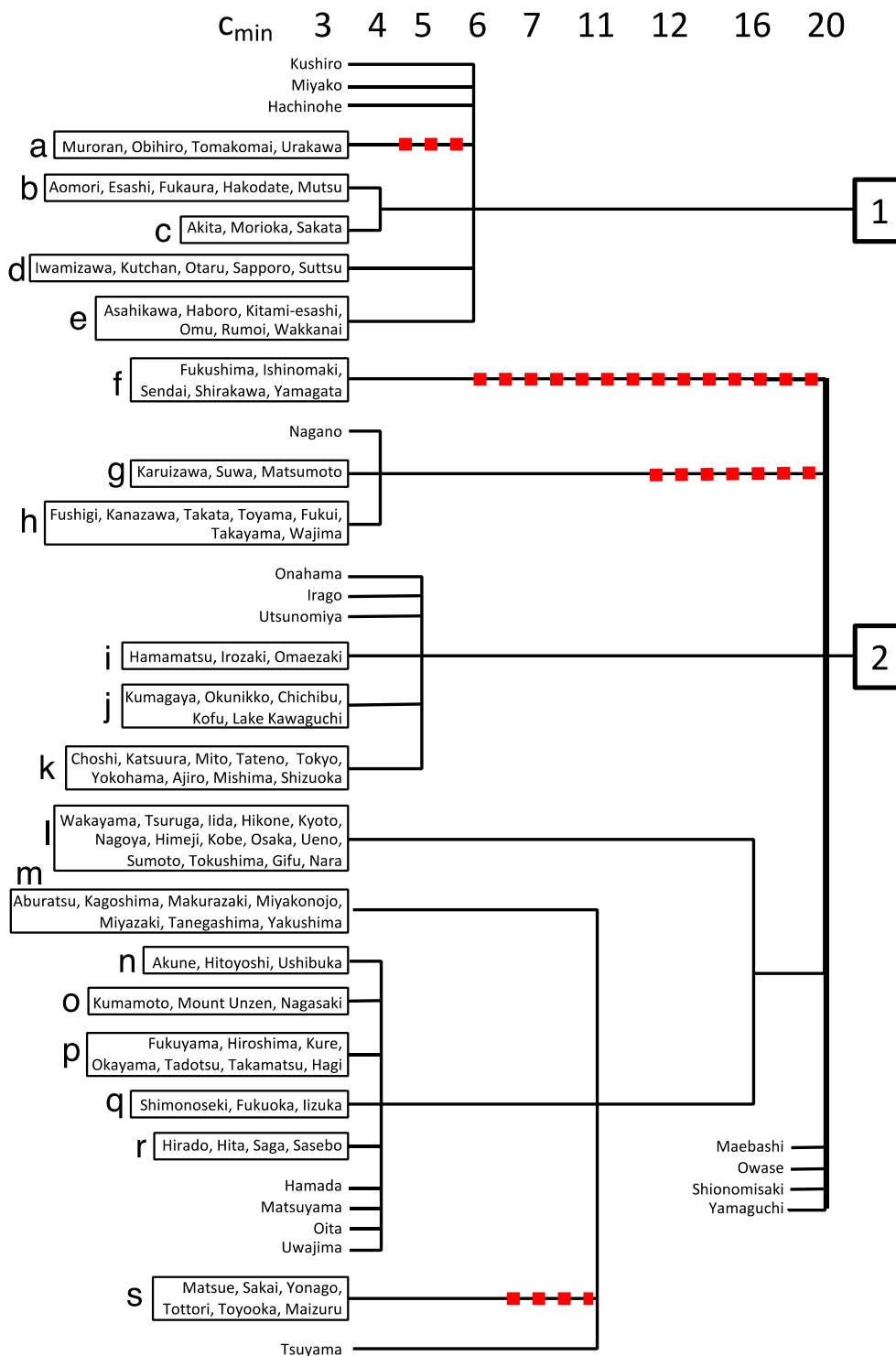


Figure R.3.8.1

9) Line 175. Authors use the procedure ‘percentiles were calculated for each individual ensemble member, then the ensemble mean was taken’. This is an alternative to take quantiles of a super-sample (collecting all sub-samples). Please comment on that.

During the verification of climatological extremes (for the whole 59 year period) we have compared taking ensemble mean and taking super-sample. The results were similar and in that case we felt the super-sample method was better, with percentiles of individual ensemble members forming the range of uncertainty about the super-sample value.

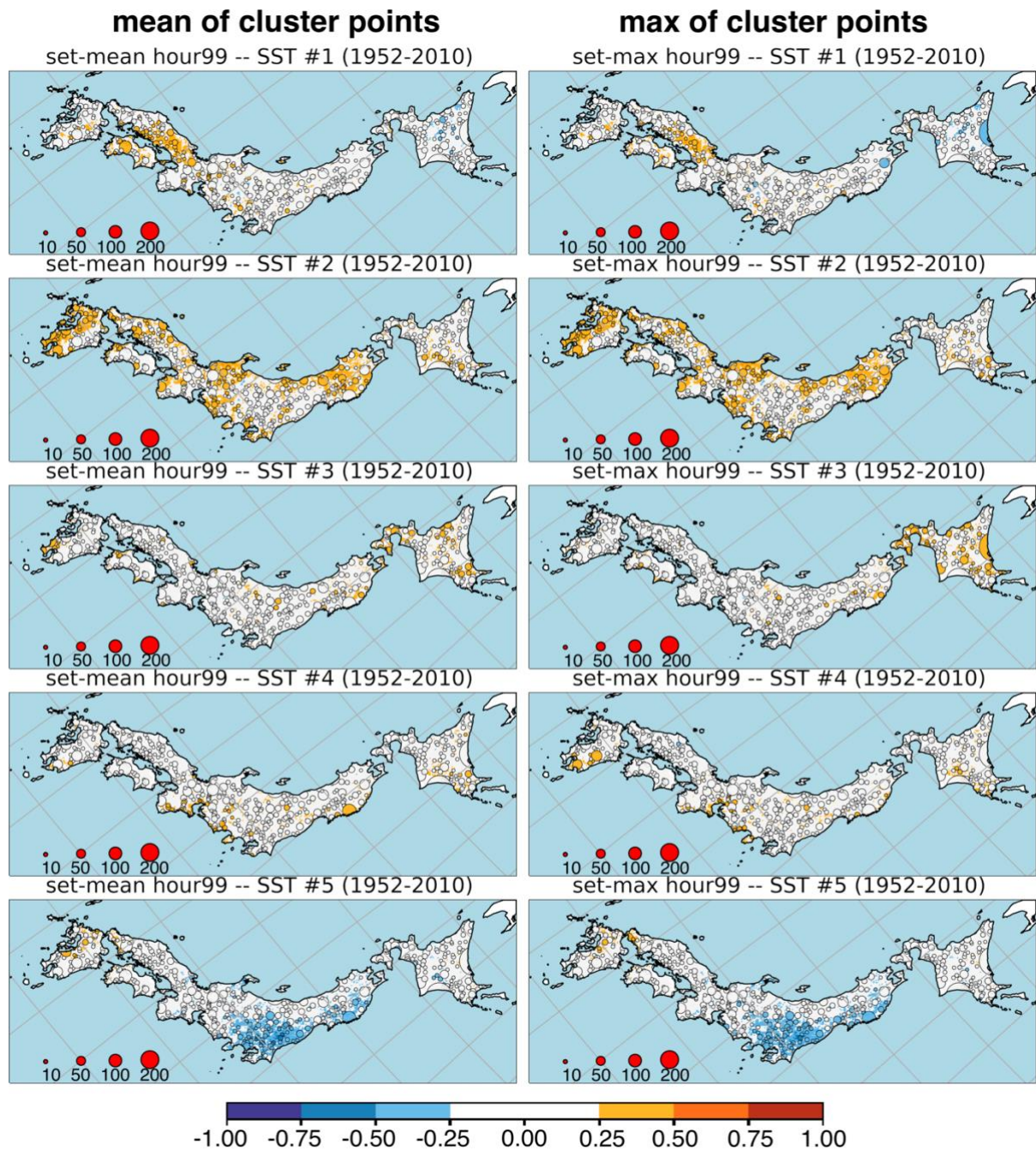
In this study a percentile is calculated for each year. Each year has its own SST and each ensemble member a perturbation of the SST, but we wish to do a correlation between the “real” SST interannual-variability and the model results. When a percentile super-sample is taken, the ensemble members do not participate equally. We have the concern that one or a few ensemble members may dominate the tail, but they do not reflect a response to the “real” SST. Instead they could reflect the response to a subset of the perturbed SSTs, and this subset could change every year. Then extreme rainfall time series could reflect the interannual variation of subsets, rather than the interannual variation of the real SST. To guard against such a possibility, we take an ensemble mean of percentiles, to force equal participation by each ensemble member. This is different from the climate case above, where the interannual variation is included in the samples.

10) Line 184 Say here the value of CC significant at  $\alpha=0.05$ .

The threshold values are now explicitly stated in L210–212 and L276–278.

11) Line 202-205. Authors say: ‘At each timestep, the maximum rainfall in each set was selected. For example, in a cluster of three points, this would be the point with the highest rainfall and the other two points would be discarded. This was based on the concept that a set included locations which experienced rainfall from the same event, and the maximum rainfall of that event was sampled by an observer that could “see” the entire location rather than only one specific point.’ This procedure seems ad-hoc and not well supported statistically. The maximum of the extremes within a cluster is taken, instead for instance the cluster average of maxima. What is the representativeness of the maximum (among the extremes) within the cluster?

Actually, I have calculated the maximum, mean, median and the 99th/90th percentiles of the clusters. But my statistical knowledge is quite poor and I could not decide which was best to use. The time series of the statistic is correlated against other time series (SST mode scores, monsoon indices). Also, multiple clusters from d4PDF correspond to one rain-gauge cluster, so a further statistic had to be calculated from many clusters. I did not understand how to interpret such calculations when performed on a statistic, so I decided to just pick the point with the maximum value in each cluster. In my mind the downstream calculations then reduced to the typical interpretation. As the reviewer points out, this can produce a time series unrepresentative of the cluster, since the points inside each cluster were not perfectly correlated in time. We have switch to the cluster mean for both rain-gauges and d4PDF. Figures R3.11.1A– R3.11.2B show comparison of plots in the original manuscript using cluster mean (left column) and using cluster max (old column). The results are qualitatively similar, but the correlation values of large clusters are affected, such as the very big cluster in north Hokkaido (most extreme right side of the panels).



**Figure R3.11.1A.** d4PDF correlation between scores of Pacific SST modes 99th percentile hourly extremes.

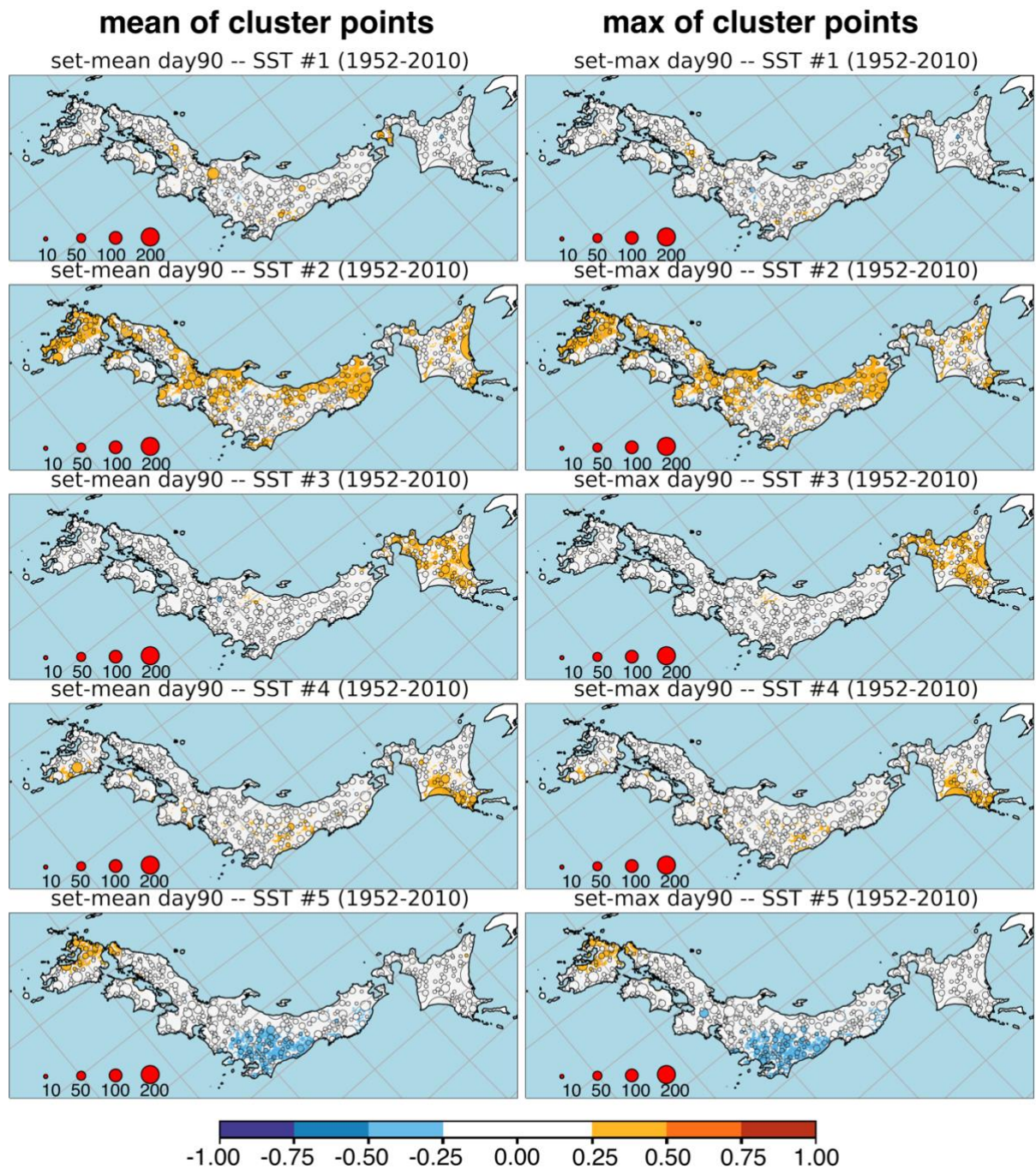
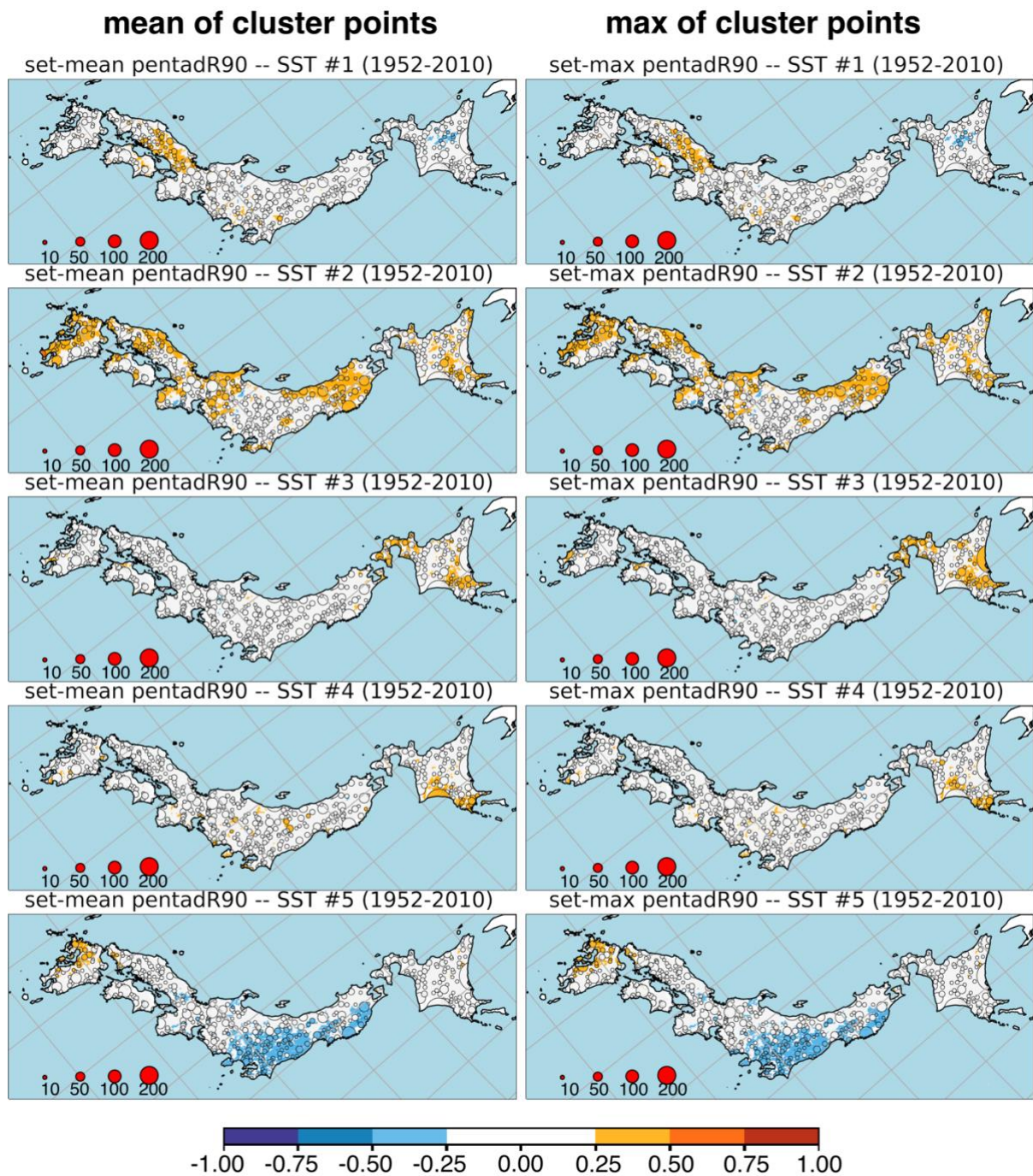


Figure R3.11.1B. d4PDF correlation between scores of Pacific SST modes 90th percentile daily extremes.



**Figure R3.11.1C.** d4PDF correlation between scores of Pacific SST modes and 90th percentile running-pentad extremes.

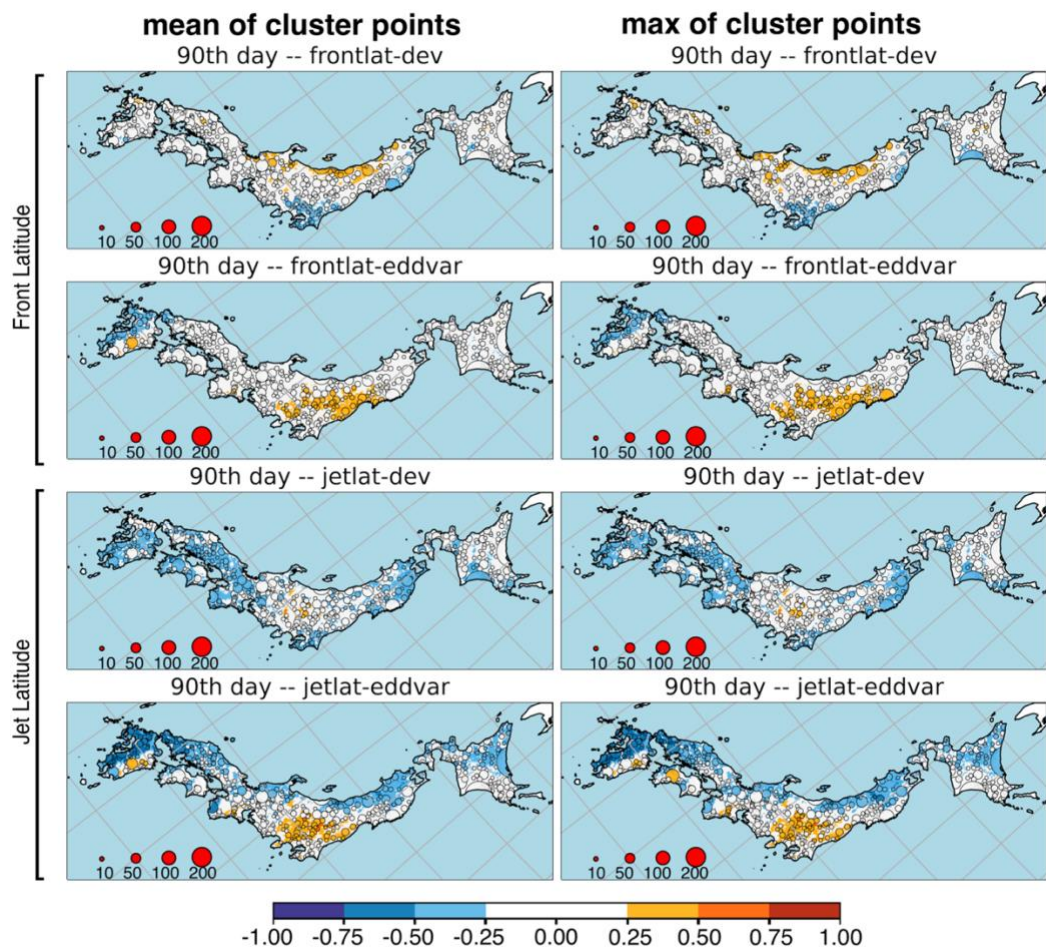


Figure R3.11.2A. d4PDF correlation between regional monsoon indices related to latitude and 90th percentile daily extremes.

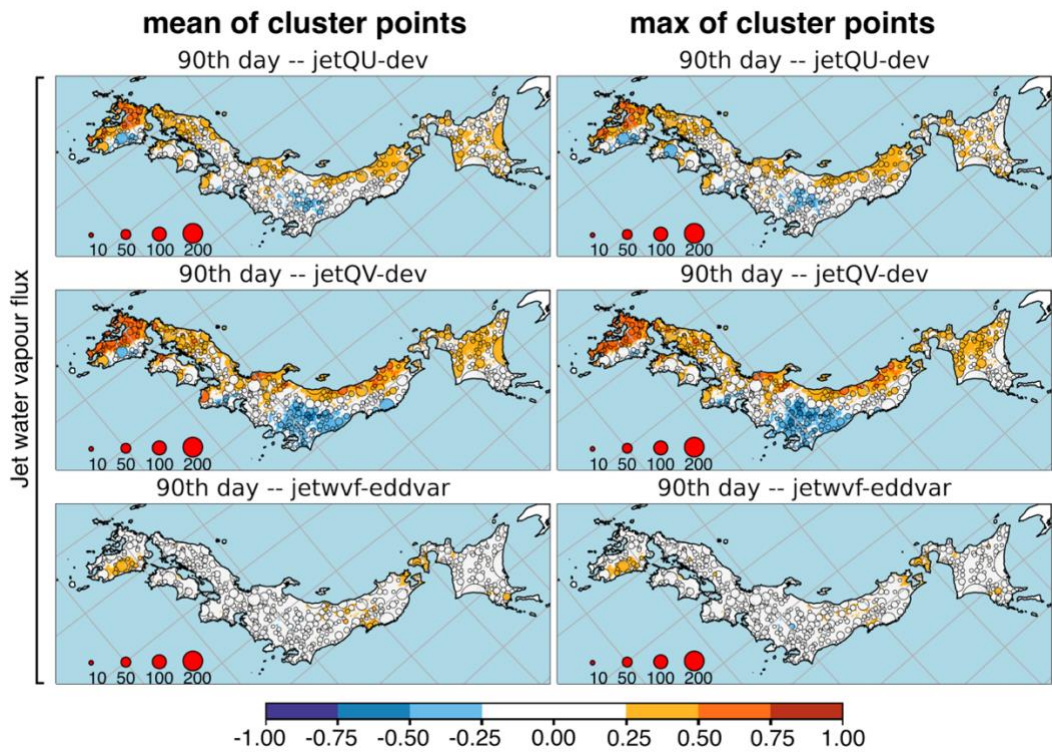


Figure R3.11.2B. d4PDF correlation between regional monsoon indices related to water vapour flux and 90th percentile daily extremes.

12) Line 220. Authors say: ‘Since there were over a hundred rain-gauges, we would like to group them into regions that persisted across the three time resolutions’. The obtained clusters depend on the time resolution (hour, day or pentad), what seems expected, since the associated meteorological systems have different spatio-temporal scales. Authors shall provide a clear justification for their choice. For instance, if instead we had taken the time-scales: hour, 6 hours, 3 days, would the clustering output be the same? Ideally, a uniform (or time-scale integrated) criterion should be adopted. Comment on that.

A multi-frequency metric has been defined. The hourly rainfall time series was decomposed into different frequency components using wavelets. (Due to the highly intermittent nature of hourly rainfall, the Fourier decomposed components looked unreasonable. In contrast, wavelet decomposed components showed matching peaks with the original time series.) Figure R3.12.1 shows a comparison of  $c_{min}=6$  groups formed using the old method (panel a) and using the new metric (panel b). Cluster 8 in panel b was manually added. Table R3.12.1 shows how the clusters are similar. The new method is better because it covers the whole frequency range without choosing certain frequency band from a priori meteorological knowledge. Also, this allows for the clusters in d4PDF to be the same between the hourly, daily and running-pentad analysis.

Unfortunately, the problem of many fragmented clusters still occurs for d4PDF. Preliminary investigations using single frequency bands or single ensemble member suggest that this may be due to a combination of high frequency component (also associated with systems of small spatial scale) and multiple ensemble members (each responding to a different perturbed SST). I am unable to solve the problem at this point in time.

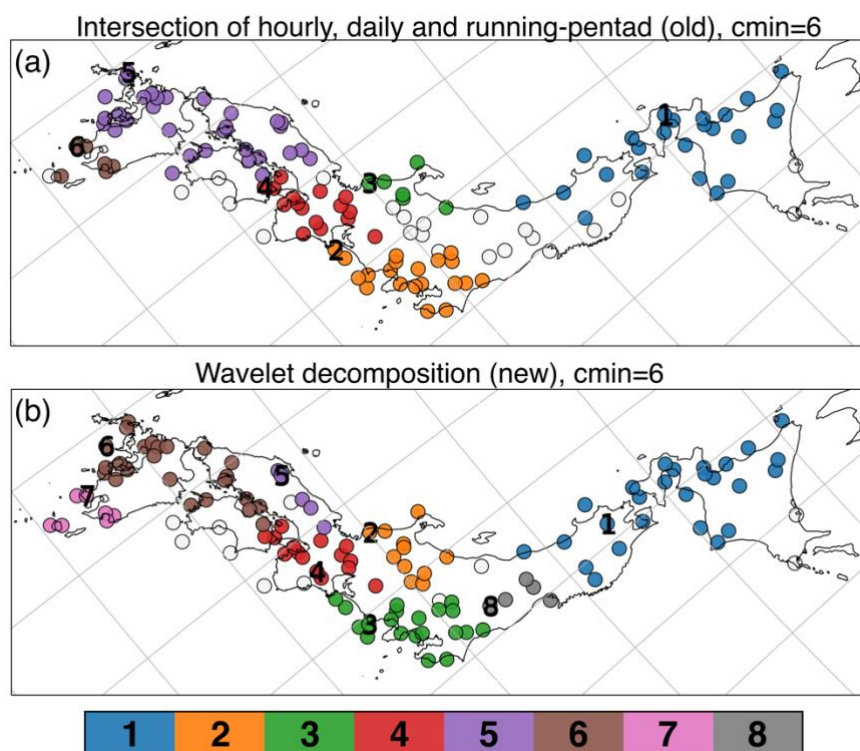


Figure R3.12.1

13) Line 220-221. Authors have performed cluster analysis for each of the three analyzed extremes and then they have grouped the respective clusters into common regions. Explain the rationale of that.

We wished to use the same set of clusters when comparing between correlations for hourly, daily, and running-pentad extremes. The methodology has changed into a single cluster analysis with a multi-frequency metric (response to reviewer’s comment 12).

14) Line 241 Authors use both the symbol R and CC for the Spearman correlation. Choose only one symbol.

Only R is now used.

15) Lines 275 The seasonal anomalous Jlat, Flat, QU and QV were obtained as residuals from a sinusoidal and logistic fit of daily values. The result can be biased due to choice of the fit. Why not computing explicitly the daily-basis seasonal cycle from averages along the 61 days with the same Julian day (or within a smoothing window of 2-3 days)?

The smoothing window was our first method of choice. Figure R3.15.1 shows our attempts at smoothing the monsoon-related variable for JRA55 (left column) and d4PDF (right column). We did not feel satisfied with the results of even a 15-day window, so we did fits instead.

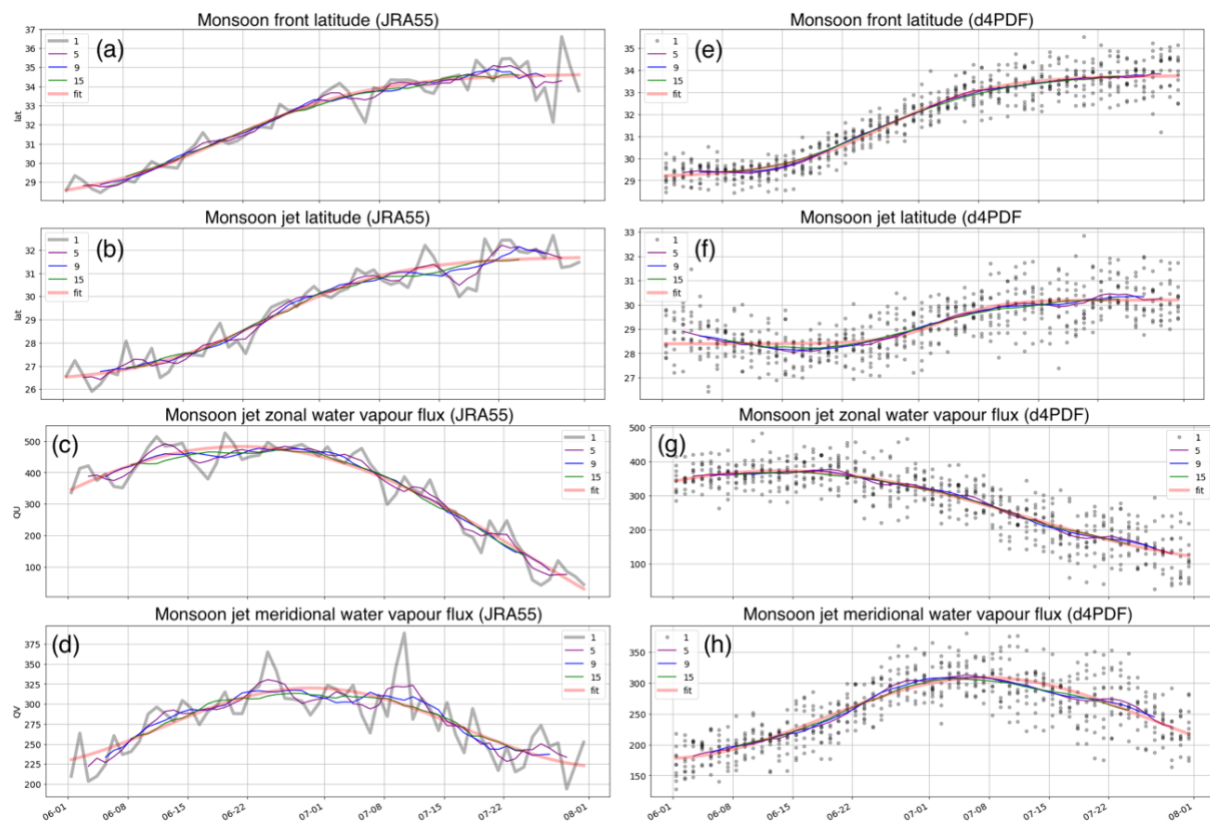


Figure R3.15.1

16) Line 285 Emphasize again that extended-PCA is applied, not PCA, to help the reading. The extended-PCA (Ext-PCA) analysis in the North-Pacific, shown by authors, have omitted other relevant modes of SST Pacific variability like the Northern Pacific Gyre Oscillation (NPGO) (DiLorenzo et al. 2008) that can be also candidates as drivers of extremes. Moreover, the Ext-PCA looks for variability all along the year (all trimesters) and not uniquely focused in the summer season. For instance, a PCA of SST in (JJA) would capture uniquely the difference between summers, eventually producing more suited modes driving the summer rain in Japan. Comment on that.

Di Lorenzo E., Schneider N., Cobb K. M., Chhak, K., Franks P. J. S., Miller A. J., McWilliams J. C., Bograd S. J., Arango H., Curchister E., Powell T. M. and P. Rivere, 2008: North Pacific Gyre Oscillation links ocean climate and ecosystem change. *Geophys. Res. Lett.*, 35, L08607, doi:10.1029/2007GL032838.

“PCA” in the text have been changed to “rePCA” (rotated extended Principal Component Analysis). A comparison of the SST modes with NPGO has been added in Section 3.1 L351–253 (reviewer’s comment 17). Rainfall extremes have been correlated with NPGO, TPI, SOI in Section 3.3 L565–567 and Figure 11. The reason for extended PCA was specifically to select for all-year variability, and clarify the timescale of the influential phenomena. Climate indices are calculated by month, but the SST climate mode are slow, such as ENSO developing and decaying over two years. A monthly decomposition may include both the direct anomaly of the slow mode, and the indirect anomaly due to the summer (monsoon) response to the mode. If it is primarily the summer response, different climate modes may project onto similar regional SST (or atmospheric) anomalies that drive a similar monsoon response. This seems to be the case based on the results of this study, so indeed analysis of regional summer markers should produce more suitable modes.

17) Line 323 The so-called TREND+ mode seems to contain interannual variability. It is not clear how much is attributed to the trend and global warming. It seems to be a mixed mode including the linear trend (maybe not the largest part), the PDV (or more commonly the PDO: Pacific Decadal Oscillation) and the NPGO.

A comparison of the SST modes with NPGO has been added in Section 3.1 L364–366. The NPGO was significantly correlated with the Trend+ mode with  $R=0.38$ . This is the best  $R$  value amongst the five modes. In order to isolate the response to warming alone, Section 3.3 L550–569 and Figs 11–12 now has additional analysis using Trend+ smoothed by an 11-year moving average.

18) Line 476. The relationship between monsoon parameters and rainfall extremes is done uniquely for daily extremes. Explain the omission of the other scales.

Figures have been added for hourly extremes as Figs 13–14, for running-pentad extremes as Figs 17–18, and summarised regionally for the three resolutions as Fig. 19. These figures were left out in the previous manuscript because I thought there would be too many figures, but they are unavoidable to properly discuss the results.

19) Lines 481–483 Authors say: ‘The spatial patterns of correlation were similar between d4PDF and rain-gauges for the  $\mu$ FLat,  $\mu$ JLat, and  $\mu$ QV, showing correlation along the Sea of Japan coast and anti-correlation along the Pacific coast (Figs 8e, 8g, 9e)’. This statement is based on very few rain-gauges, while in some others the agreement is neutral (weak correlations). For other parameters the similarity is quite inexistent, so any apparent similarity seems to be unfair.

I may have made some unfair claims on similarity due to psychological bias. To prevent this, the manuscript now focuses on rain-gauges using objective criteria such as statistical significance within the rain-gauge cluster and field significance for rain-gauges or clusters. To compare d4PDF against rain-gauges, the box-and-whisker plots for the 8 analysis regions are used, such as Figs 6, 10, 12, and 19. The Discussion now focuses on explaining how the regional monsoon acts as an intermediary between the Pacific SST modes and rainfall extremes, for each case of rain-gauges and d4PDF.

20) Line 503. Authors refer to ‘scores of the Pacific SST’. Do you mean PCs (ENSO+, ENSO-, TREND+ etc.)? Rewrite and clarify. A Table with correlations between PCs and monsoon indices should be presented with analysis of significance and discussed afterwards. No physical explanation is provided for the correlations found at the end of the Discussion Sec. 4. Improvement needed.

The sentence has been rewritten as “correlations between the monsoon indices and standardised principal components of the Pacific SST modes were calculated”. Table 2 has been added showing the correlations. Figure 20 has been added to show the bi-weekly progression of the monsoon in JRA55 and d4PDF. We offer the explanation that differences between JRA55 and d4PDF was due to the slower northward movement of the monsoon in d4PDF.