Reply to Reviewer's Comments

Reviewer 2:

GENERAL COMMENTS:

Overall, this paper can be great contribution to our science community. The authors come up with the novel approach for quantifying the contribution of atmospheric circulation to the recent precipitation variability in US. They combine two statistical techniques - Self Organizing Map (SOM) and circulation analogue - to effectively visualizing circulation pattern as in SOM and quantifying the dynamical contribution as exact as in circulation analogue. Using this new SOM-Analogue (SOMA) approach, they investigate the link between circulation patterns and daily precipitation anomalies in different seasons and PDO phases. The results provide the valuable insight how the specific circulation pattern leads to the regional precipitation anomalies in US. However, I think there are some major issues in the development of new SOMA approach and the analysis regarding PDO-related contributions. I would suggest the authors to address below major/minor comments before the publication.

MAJOR (SPECIFIC) COMMENTS:

Decomposition of dynamic and thermodynamic contribution

The variables Z500 and IVT are used for elucidating the dynamic contribution to the precipitation anomalies. However, IVT is multiplication of wind vector (dynamic variable) and moisture (thermodynamic variable), so the SOMA from IVT would also contain the thermodynamic contribution. To decompose the dynamic and thermodynamic contribution, I think dynamic variables (e.g., velocity potential) has to be used instead of IVT. Or, rather than focusing on the decomposition of dynamic/thermodynamic contribution, the authors can express the same results as the contribution from the moist circulation and the residual. In my opinion, the interesting science in this study is to characterize the circulation pattern responsible for the regional precipitation anomalies, and the exact decomposition of thermodynamic and dynamic contribution is kind of secondary interest compared to that.

If the authors want to decompose the thermodynamic/dynamic contribution, then I think new variable should be used for SOMA instead of IVT. Or, authors can express the same results as the contribution from the moist circulation and residual, with revising rest of the manuscript accordingly.

Thank you very much for this helpful comment! The other reviewer also suggested that a decomposition based on wholistic contribution from the SOMA circulation vs. residual is more appropriate. We took your advice and changed P'_{dyn} and P'_{the} to P'_{SOMA} and P'_{RES}. In the texts, we also changed terms like "dynamic contribution" to "moist circulation contribution" or

"SOMA contribution". In addition, we added some texts in the method section to reflect this consideration.

Line 193-199: "It's worth noting that attempting to separate the actual dynamic and thermodynamic components solely through a circulation clustering approach like SOM can be challenging. Each type of circulation, as represented by SOM nodes, inherently encompasses thermodynamic responses. Therefore, the distinction between "dynamic" and "thermodynamic" components can be ambiguous when using these terms. Consequently, we prefer to refer to the precipitation influenced by moist circulation patterns involving Z500 and IVT as P'_{SOMA} instead of P'_{dyn}. This emphasizes that our results regarding circulation contributions are contingent on our chosen set of circulation variables."

Issues regarding the data and SOMA approach

First, is the observational data used for SOMA **detrended or high-pass filtered**? In previous studies using circulation analogue (Deser et al. 2016; Lehner et al. 2018), the dynamic contribution is found after the observational data is detrended or high-pass filtered. This is for eliminating the forced thermodynamic contribution in the timeseries and to focus on the dynamic contribution. In this manuscript, it seems there are no mentioning about such data processing. I think the detrend is needed if it is not done, and it should be mentioned if it is done. Or, you may show that such data processing doesn't matter to your results.

Yes, all data including circulation and precipitation are high-pass filtered with a simple 5-day moving average. For Z500, global mean Z500' are subtracted from original data prior to calculating anomalies to account for thermal expansion of the atmosphere due to warming following prior studies, so there is no apparent trend in Z500 data. We did not use the traditional detrending approach (remove linear trend) because there could be multidecadal variability (both circulation and precipitation) linked to internal variability which could sometimes falsely appear as a weak linear trend. These were mentioned in the data section (2.1) and we revise those sentences (attached below) to make them clearer to the authors the data is high-pass filtered and detrended.

Line 98-102: " To account for the thermal expansion of the warming atmosphere, we subtract the daily global area-weighted mean Z500 from the daily Z500 data at each grid point so that there is no linear trend in the Z500 data due to warming (Christidis and Stott, 2015; Siler et al., 2019; Zhuang et al., 2021b). To mitigate high-frequency synoptic noise, we employ a simple 5-day moving average filter to both Z500 and IVT."

Second, in L197-199, to apply circulation analogue to each BMU node, the authors regress total precipitation anomalies onto the circulation anomalies (PC values), and multiply that regression coefficients to the circulation anomalies (PC values) to get dynamical precipitation changes. I think this can overemphasize the contribution from the circulation anomalies, and authors may need to come up with another approach for this step. The detail is written below.

In previous studies for circulation analogue, the linear coefficients or subsamples to estimate the dynamic precipitation anomalies are calculated only using the circulation anomalies, as you wrote in L205-206. However, here, total precipitation anomalies are directly regressed onto the circulation anomalies in same BMU node to get coefficients, and that regression coefficients are used to estimate the dynamic precipitation anomalies. This is like assuming that total precipitation anomalies are similar to anomalies from the moist circulation (or dynamic contribution) even before the decomposition. I think this would artificially overemphasize the dynamic contribution. So, I recommend you to revise step 4 in L197-199. You may use the linear coefficients for reconstructing target-day circulation from analogue-day circulations to getting target-day dynamic precipitation anomalies from the analogue-day precipitation anomalies in same BMU node, following Deser et al. 2016. I think you should avoid to directly link the total precipitation anomalies to the circulation anomalies to get the coefficients for estimating dynamic contribution.

Sorry for the confusion. In the original manuscript, we have a paragraph briefly describing the rationale behind regressing P' onto circulation anomalies (right after the detailed method steps). In order to help readers better understand why we calculate it this way and how it differs from previous methods (especially constructed flow analogue), we move that paragraph above the method steps and revise it to include some equations for better clarification. Related texts are attached here.

Line 201-217: "The basic idea behind SOMA is to incorporate variability in P'_{SOMA} for days with the same BMU node, similar to the flow analogue method, while adhering to the C2007 framework. In contrast to the flow analogue method, where analogue days are determined by minimizing the Euclidean distance of anomalous circulation patterns, in SOMA, the analogue days are directly selected from the daily samples sharing the same BMU. Additionally, the constructed flow analogue approach relies on a linear combination of precipitation anomaly from analogue days to quantify circulation contribution at a target day *j*, i.e., $P'_{j,dyn} = \sum_i a_i P'_i = \sum_i a_i P'_{i,dyn} + \sum_i a_i P'_{i,res}$, where the coefficients a_i are determined by the linear dependency of the circulation pattern on target day *j* and those on analogue days *i*, i.e., $Z'_j = \sum_i a_i Z'_i$. However, it is important to note that the analogue day precipitation anomaly $P'_i = P'_{i,dyn} + P'_{i,res}$ includes both a circulation-contributed component and a residual component. This approach overlooks the potential thermodynamic influences in the precipitation anomaly on analogue days, and the resulting circulation-contributed component $P'_{j,dyn}$ still retains a thermodynamic residual $\sum_i a_i P'_{i,res}$, which may not be negligible with limited samples. To mitigate this issue, we make the assumption that for each group of similar circulation pattern (same SOM node), $P'_{j,dyn} = \sum_i a_i P'_{i,dyn}$. This differs from the equation form in constructed flow analogue and excludes the term $\sum_i a_i P'_{i,res}$. Following constructed flow analogue, where each circulation pattern is linearly dependent on other circulation patterns within the same SOM node, i.e., $Z'_j = \sum_i a_i Z'_i$, we can see that P'_{dyn} and Z' share the same linear form and coefficients. Therefore, P'_{dyn} and Z' can be considered linearly dependent as well, allowing us to treat the calculation of P'_{dyn} as a regression problem, which can be resolved using the steps outlined below."

PDO-related analysis

The precipitation anomalies related to PDO is often linked to the large-scale circulation over the North Pacific (retreat of Aluetian Low/ expansion of subtropical high), and it was less linked to the circulation anomalies within the US. This study can further investigate the teleconnection impact of PDO on the circulation anomalies within US, which will be the valuable contribution for regional teleconnection studies. However, the regional circulation and precipitation anomalies related to PDO will probably sensitive to the detailed pattern of tropical SST pattern (as it is known that the teleconnection within US depends on ENSO diversity), and they will be affected by the other variability (e.g., AMO, ENSO) as well. In this situation, the mere difference of positive and negative phase of PDO with relatively short length of observation would not sufficient to convince the PDO impact in this study. These limitation needs to be addressed in the manuscript.

Thank you for pointing out these limitations. And we agree that it would be impossible to disentangle the PDO contributions from various climate modes simply by using our current form of methodology. We added discussion about these limitations in the discussion section (attached below). Hope this can address your concerns.

Line 471-482: "Furthermore, the identified PDO-related SOM nodes and P' is likely also modulated by tropical SST variability. This is due to the well-established understanding that ENSO and PDO can generate similar atmospheric and oceanic anomaly patterns (e.g., Hu and Huang, 2009). Additionally, other internal climate variability modes, such as the Atlantic Multidecadal Oscillation (AMO; e.g., Hu et al., 2011), North Atlantic Oscillation (NAO; e.g., Whan and Zwiers, 2017), and Interdecadal Pacific Oscillation (IPO; e.g., Dai, 2013), can also influence these patterns. The SOM circulation patterns defined by Z500′ and IVT′ simply represent regional-scale manifestation of larger-scale variability simultaneously influenced by multiple internal climate modes. Attempting to isolate the individual contribution of these modes using statistical methods with limited data, such as SOM or SOMA, can be a complex challenge. Therefore, our results related to PDO contribution serves as a preliminary starting point which demonstrate the combined contribution of the SOM node patterns statistically linked to the PDO. To gain a more detail quantification of the PDO's influence excluding the effects of other climate variability modes, further research integrating both observational data and climate model output is needed."

In spite of above limitations, I think the analysis in this manuscript can be developed further to shed light on understanding the PDO teleconnection impact. The figure 10 is a good starting point where PDO-related local circulation/precipitation anomalies can be hypothesized and analyzed in detail. I suggest the authors to provide the figure for precipitation anomalies corresponding to figure 10, showing the precipitation anomalies linked to each circulation changes between two PDO phase (for three target regions). And then, authors may select the circulation node which shows strong precipitation anomalies, linking the PDO to that circulation and precipitation anomalies. If the PDO-induced teleconnection can explain those circulation pattern with the previously known mechanisms, authors can somewhat convince that their PDO-related precipitation anomalies are indeed PDO-induced. In summary, I think the current analysis/discussion for the PDO impact is not enough, and authors need to investigate the significant circulation pattern in Fig. 10 and try to link those to the PDO-related teleconnections to ensure the PDO impact.

Thanks for the suggestion. We have modified our method of identifying PDO-related nodes to include the contrast in circulation related precipitation (P'_{SOMA}) in addition to node frequency (f), so as to ensure the nodes we selected has PDO impact on precipitation. Related texts are also attached below:

Line 346-361: "To further determine how much of the dynamic contribution to the multi-decadal changes of P' can be linked to PDO changes, we first investigate whether the moist circulation and precipitation anomalies associated with each node may be modulated by PDO phases. We assess whether the difference in frequency-weighted P'_{SOMA} between positive and negative PDO phases is statistically significant through a Monte-Carlo test. Detailed steps are as following: a) for node k in month m, calculate node frequency for positive PDO phase (monthly PDO index > 0.5) and negative phase (monthly PDO index < -0.5) for all years, i.e., $f_{PDO+}(k,m)$ and $f_{PDO-}(k,m)$, as well as average SOMA contributed P', i.e., $P'_{SOMA,PDO+}(k,m)$ and $P'_{SOMA,PDO-}(k,m)$. b) the frequency weighted P'_{SOMA} difference between positive and negative PDO phases for node k in month m can be calculated as $\Delta_{PDO}P'_{SOMA}(k,m) = f_{PDO+}(k,m)\times P'_{SOMA,PDO+}(k,m) - f_{PDO-}(k,m)\times P'_{SOMA,PDO-}(k,m)$, and overall PDO impact for node k can be represented by summing this difference across all months, i.e., $\Delta_{\mathrm{PDO}}P'_{\mathrm{SOMA}}(k) = \sum_{m=1}^{12}[\Delta_{\mathrm{PDO}}P'_{\mathrm{SOMA}}(k,m)]$. c) randomly shuffle the sequence of years for the PDO index data and recalculate the abovementioned statistics. d) repeat step c) for many times (10000 here), if $\Delta_{PDO}P'_{SOMA}(k,m)$ is larger than 90% of all simulated values, then for node k in month m, P'_{SOMA} is significantly larger during positive PDO phase than negative phase; in contrast, if $\Delta_{PDO}P'_{SOMA}(k,m)$ is smaller than 90% of all simulated values, then P'_{SOMA} is considered significantly smaller during positive PDO phase. e) similar conclusions can be drawn for node k overall by assessing $\Delta_{\mathrm{PDO}}P'_{\mathrm{SOMA}}(\mathrm{k})$ instead of $\Delta_{PDO}P'_{SOMA}(k,m)$. The resulting nodes statistically related to PDO (for each month and overall) are shown in Fig. 9."

Line 371-382: "A significant $\Delta_{PDO}P'_{SOMA}$ value could be due to differences in node frequency, node mean P'_{SOMA}, or both, between the two PDO phases. Our result in Fig. 9 indicates that, for most nodes, PDO phases can have very different influences on precipitation through moist circulation conditions identified in this study, for different seasons and regions. For example, node A1, characterized by geopotential high anomalies along the west and east coasts and low anomalies over the central US, mirrors a Pacific North America (PNA)-like pattern, previously linked to Pacific variability and its impact on precipitation in the western US and GP (e.g., Ciancarelli et al. 2014; Zhuang et al. 2021a). Interestingly, node A1 does not display a significantly higher frequency during the positive PDO phase compared to the negative phase (even lower in January to April; Fig. 10). However, it stands out as one of the nodes exhibiting the most substantial contrast in P'_{SOMA} between the two PDO phases over the SGP. This contrast arises because the composite circulation pattern for node A1 during the positive PDO phase exhibits stronger negative Z500' over the central US, which can promote upward motion and is generally linked to reduced convection inhibition (Fig. S5), thereby leading to larger P'_{SOMA}."

MINOR (TECHNICAL) COMMENTS:

L60: Is there any need to use the term "surface parameter"? I think you can just directly mention the precipitation instead of "surface parameter such as precipitation". It would be great if the terminology in this paragraph is adjusted accordingly.

We use "surface parameter" because this method and similar ones are applicable to variables other than precipitation as shown in previous studies, such as temperature and VPD. But we agree simpler terminology should be used for better flow. So in this revision, we keep "surface parameter" in Line 59 when SOM is first introduced but simply use P' afterwards for elsewhere in the manuscript. Hope this can address your concern.

L60-63: I think it is hard for readers to understand Cassano et al. 2007 (C07) SOM method by reading few sentences here. I suggest you to add and explain the equation 1 and 3 of C07, since they are quite simple. Those equations may be added on the line with simple explanations.

Thanks for the suggestion. We have revised Line 59-64:

"The idea is that the total precipitation anomaly for a certain node *K* during a period (P'_K) can be decomposed as the sum of a circulation or dynamic component ($P'_{K,dyn}=\Delta f_K P'_K$) controlled by the mean frequency change of SOM nodes, a thermodynamic component ($P'_{K,the}=f_K \Delta P'_K$) related to the mean state change of precipitation for the same SOM node, and an interaction term ($P'_{K,int}=\Delta f_K \Delta P'_K$) related to both. Detailed descriptions of this method can be referred to Cassano et al. (2007) and Horton et al. (2015)."

L138: I think it is better to represent the fraction of variances only for the Z500 and IVT combined variability since the EOF analysis is done simultaneously to two variables.

Yes, we agree. PC number cutoff is based on the combined variance, we just added more information about how much variance in each variable is explained. We revise the sentence to better reflect this:

Line 144-145: “Specifically in this study, we retain the top 26 PCs which explain 87.3% of the combined variance in Z500′ and IVT′ (or 90.6% and 77.4% of the variance in Z500′ and IVT′, respectively).”

L164: “issmall” to “is small”?

Corrected.

L194-196: Why do you find the analogue day within 91 calendar windows? This might be due to the seasonal dependence of circulation-precipitation relationship, and it is better to mention this somewhere in the text.

Thanks, added in Line 226-227:

“3) For any day t0, assuming its BMU is node i, we find all days tj (j = 1, 2, …, M) that have this node i as their BMU and are within the 91-day calendar window centered on the day t0 but not in the same year as t0, to account for the seasonal dependence of circulation-precipitation relationship, i.e., BMU(xtj)=BMU(xt0), |DOY(tj) – DOY(t0)| ⩽ 45 d, Year(tj)≠Year(t0).”

L317: “Fig. 9 (1st column) shows the period mean P′ for different seasons and regions” à There are no seasonal information in Fig. 9. You may need to change the text.

Thanks, “different seasons” is removed.

L323: “Table2” to “Table1”?

Corrected.