Dear Reviewer:

Thank you so much for your kind suggestions, we reply to your comments shortly here and we will make corresponding changes in the revised manuscript.

The manuscript entitled "A new uncertainty estimation technique for multiple datasets and its application to various precipitation products" introduced the variance partitioning method into uncertainty quantification of ensemble precipitation datasets, which considers both temporal and spatial uncertainties, and thus established a more comprehensive uncertainty metric as compared with the classic metrics. The deviation of the mathematical framework is rigorous and complete, while lots of work, including various precipitation products in multiple regions, was conducted in the validation of the new metric. On the other hand, some theoretical questions are needed to be explained clear and the English writing of this manuscript needs improvement. The detailed problems are listed as follows:

(1) According to the definition of the new uncertainty metric, it's one of components partitioned from the SST over time, space and ensembles. Thus, the new uncertainty Ve interacted with other components (Vt and Vs), as the authors discussed in Section 6.1. Given an ensemble of precipitation, if we replace one year's data to make the inter-annual variation larger, then the Ve obtained will correspondingly decrease.

Reply: Yes, the Ve is one of the components partitioned from the total grand variance. If any of the values change in the dataset, all the Ve, Vt, Vs will change correspondingly. In our case discussed in Section 6.1, if we evaluate the real precipitation datasets with the monthly values, the Ve decreases compared to the Ve evaluated by the datasets at the annual scale.

But if we evaluate the method with assumed data as we enlarge the inter-annual variation, Ve is not necessarily decreasing. Ve (or Ue) is an estimation of the difference between multiple datasets. Expanding the temporal variation of a piece of data (or all data) does not mean the difference among multiple datasets decreases or not. It is not very easy to illustrate it in three-dimensional datasets, but we can explain it easily with one-dimensional time series. Assume we have two time series (T_1, T_2) with the same fluctuation (t_i) but different variation amplitudes (k_1, k_2).

$$T_1 = k_1 t_i \ ; \ T_2 = k_2 t_i; \ and \ k_1 < k_2$$

The Ve evaluates the similarity between the two time series. If we increase the variation of T_1 (by increasing k_1), the similarity of the two series will increase and Ve (or Ue) will decrease. Instead, if we increase the variation of T_2 (by increasing k_2),, the similarity decreases but Ve (or Ue) increases. So it is not the inter-annual variation change determines the Ve but its difference between the changed data to other series determines the ensemble variance. The conclusion in 1-D series can be upgraded to the 3-D database applied in the manuscript.

However, this decrease of Ve resulted from regular temporal variation instead of variability of ensemble precipitation datasets. In summary, how to separate the influence of normal spatio-temporal variation from the ensemble variability representing the new uncertainty estimation?

Reply: If we only enlarge the temporal variation of a piece of data, the variability of ensemble precipitation datasets will keep the same. But because the data has been partly changed, the difference between the changed dataset and remained others will change. Ve (or Ue) will change as a result.

There are two kinds of uncertainties. The first is the uncertainties between any database and its real values. The second is the uncertainties among multiple datasets which show the different performance of various datasets. The former one is difficult to evaluate because we never know the real values. Thus, we rely on the second uncertainty estimation as if different datasets show small uncertainties around the ensemble values, the ensemble estimation has high credibility. The present study only focuses on the latter uncertainty which is evaluated by Ve (or Ue). The changes of normal spatio-temporal variation can be revealed in the changes of Ve (or Ue). But the variability of ensemble datasets (the first kind of uncertainty) is not evaluated and it will remain the same if the spatio-temporal variation changes do not affect the ensemble variability.

In addition, is the classical N.t.std or N.s.std affected by the same temporal or spatial variation?

Reply: Taken the sample the reviewer has proposed, if part of the inter-annual variation is enlarged, the N.s.std will keep the same because it requires averaging the time series across different time steps.

$$N.s.std = \frac{\sqrt{\overline{\sigma_{e_t}^2}}}{\mu} = \frac{1}{\mu}\sqrt{\frac{1}{n}\sum_{j=1}^{n}\sigma_{e_t}^2[j]}$$

$$\sigma_{e_t}^2[j] = \sigma^2(\mu_t[j,:])$$

$$\mu_t[j,k] = \frac{1}{m}\sum_{i=1}^{l} z_{ijk}$$

Averaging temporal variation

But N.t.std changes correspondingly. This is the one of the shortcomings of *N.s.std* (or *N.t.std*) as the changes of temporal variation is not necessarily revealed in the result.

(2) Following Comment (1), the authors should clearly define the reasonable variation resulted from temporal dynamic or spatial heterogeneity and variability associated with uncertainty investigated in the present study of biased ensemble precipitation datasets. Also, as the interaction between the spatio-temporal variation and ensemble data variability exists, is it part of the uncertainty Ve?

Reply: As we explained in the comment (1), there are two kinds of uncertainties and Ve (or Ue) only evaluates one of them. We will explain it better in the revised manuscript.

As the reviewer mentioned in the comment (1), Ve is interacted with Vt and Vs. But we don't think the spatio-temporal variation is interacted with the ensemble data variability. The counter-sample is what the reviewer has suggested in the comment (1) as we enlarged the spatio-temporal variation but the ensemble data variability remains the same. Ve (or Ue) will capture the changes of the spatio-temporal variation of any piece of data but it has no association with the ensemble variability changes.

(3) The authors said the new uncertainty metric Ve contained both temporal and spatial uncertainties at the same time, while the classical metrics (N.t.std or N.s.std) contained only one source of uncertainties. Why the comparison of Ve with classical metrics was conducted by using each of classical metrics rather than the sum of N.t.std and N.s.std?

Reply: Each classical metric has its physical meanings as the *N.s.std* represents the uncertainties across space and *N.t.std* represents the uncertainties across time. So, the comparison of Ue to only one of them helps investigate the metric performance on the same physical meaning.

Though it is possible to compare Ue with an combination of *N.s.std* and *N.t.std*, the two classical metrics cannot be summed up with the same weights due to the different size of their data samples (m and n). Since the Ue is in general linearly correlated with each of the two classical metrics, Ue will be linearly correlated with a new metric which combines the two classical metrics. These can be added into the manuscript after the comparisons of Ue with single metric.

(4) Many literatures in the field of hydro-meteorology have studied on the variance decomposition method in multi-source uncertainty investigation. What's the difference between the new uncertainty estimation partitioned from grand variance and previous studies should be highlighted in Abstract and Introduction.

Reply: We will investigate more literatures and then add the comparisons to the revised manuscript.

(5) There exist many grammar mistakes in the manuscript. For example, "an new uncertainty ..." in Abstract should be "a new uncertainty ...", the expression of "which has been included the model variation" in Abstract is wrong, "of" was omitted after "because" in Line 30 on Page 2. In addition, please check Line 16 on Page 5 and Line 8 on Page 25.

(6) Despite the grammar mistakes, multiple improper or incomplete English expressions also tended to hinder the readability of the paper. For example, the mean precipitation value in Line 26 on Page 10 may be not only derived from "The long-term annual mean precipitation" but also from the lumping of spatial grids? To make reviewers and readers fully understand this study, the English should be improved considerably throughout the manuscript.

Reply: Thanks for pointing out the mistakes and the improper expressions. We have revised the ones the reviewer has mentioned, and we will check others throughout the manuscript before submitting a new version.

(7) The gauge precipitation provided by CMA was taken as the benchmark. Although the CMA data was excluded from gauge-based group, other gauge-based products also contained part of the gauge data from CMA. This is expected to clearly state. Were the gauge-based data downloaded in grid or gauge format? Are all the precipitation data in daily time scale?

Reply: The CMA was downloaded in grid format and the original data is in daily scale. But we summed up the daily values to monthly and annual scale for use in the present study.

Yes, some of the CMA gauges are included in other global observation systems. We will clarify it in the revised manuscript.

(8) Why is there no content in the section of 2.4?

Reply: Sorry for the mistake. We moved the statements for "underlying of the uncertainties" to previous section. The section 2.4 will be removed in the revised manuscript.

(9) In Figure 6, since the curves plotted represented the uncertainty, what does the band of ±standard deviation around the curves mean, the uncertainty of uncertainty? Please explain.

Reply: The colored solid line represents the ensemble mean of precipitation in each precipitation data group over the same subregion. The shaded area (±standard deviation around the curves mean) represents the uncertainty between different datasets in that data group.

$$\mu_i = \frac{1}{l}(P_{1,i} + P_{2,i} + P_{3,i} + P_{4,i})$$

$$\sigma_i = \sigma(P_{1,i}, P_{2,i}, P_{3,i}, P_{4,i})$$

$P_{1,i}$ represent the mean precipitation in precipitation dataset (the first among the four) over a specific region at time step i. $P_{2,i}, P_{3,i}, P_{4,i}$ represents the values for other three precipitation datasets in the same precipitation groups (gauge-based, merged products, GCMs). The curve is the mean of the four (μ_i) and the shaded area is their standard deviation (σ_i). We will revise the caption of Figure 6 to avoid the confusions.

(10) In Figure 12, the quantile of the box is increasing from bottom to top for normal boxplots, while there is inverse order of quantiles here. Please check it.

Reply: Thanks, we did not notice the order here and we will revise it in the new version.