Replies to the comments from editor and reviewers

Dear editor and reviewers,

Thank you so much for your comments and suggestions. Based on your comments, we concluded that the presentation quality was the major problem of the pervious submission. Improvement on language was necessary. Therefore, we firstly addressed the comments from editors and reviewers and then revised the manuscript throughout the text to avoid similar problems. We then asked one colleague to read the updated text and revised the manuscript again according to the feedbacks to improve the readability. At the last step, we consulted one UK proofreading service company for the further improvement of the manuscript. A final check was done before submitting the revision. I hope the revised manuscript can meet the publication standard of *HESS* at this time.

The comments have been replied one by one in this document and the manuscript with the changed highlighted has been provided as a separated document. The certificate for proofreading service and the modifications are also attached for your check.

Editor Comments

Please address the comments of the referee.

> Reply: The concerns of the reviewers have been addressed. We replied to those comments in this document and the corresponding changes can be tracked in the attachments.

Additionally: in the new version of manuscript you use the term "derivation ratio". Should it be perhaps the "deviation ratio"?

> Reply: Sorry for the mistake. It should be “deviation ratio” as you mentioned. This mistake has been corrected in the revised manuscript.

Comments on the Abstract ================

Estimates of a certain climatic variable are frequently seen -- unclear what do you mean by “seen”... where? By whom?

Parallel datasets

--- what does parallel mean?

> Reply: We tried to emphasize that ensemble estimates have been frequently applied in climatic research. The “parallel datasets” indicates a number of datasets which can provide estimates of the same climatic variable. In order to avoid the confusion, we modified the first sentence in the abstract as:
>
> *“Ensemble estimates based on multiple datasets are frequently applied once many datasets are available for the same climatic variable.”*

Accompanying uncertainties evaluation with the ensemble is recommended while a fundamental flaw is that the uncertainties in temporal variation and spatial heterogeneity are not together considered for the final uncertainty estimate.

--- cumbersome formulation - please reformulate

Reply: We included too much information in this single sentence thus we modified the abstract as:

"Uncertainty that evaluates the difference between the ensemble datasets is always provided along with the ensemble mean estimates to show to what extent the ensemble members are consistent with each other. However, one fundamental flaw of classic uncertainty estimates is that only the uncertainty in one dimension (either the temporal variability or the spatial heterogeneity) can be considered, whereas the variation along the other dimension is dismissed due to limitations in algorithms for classic uncertainty estimates, resulting in an incomplete assessment of the uncertainties."

Ue is higher than classic estimations metrics for the improvement of uncertainty estimation.
--- Unclear, why it is "for improvement"?

Reply: The difference of *Ue* and the classic uncertainty metrics is that *Ue* estimate avoids pre-averaging the variation in either the spatial dimension or temporal dimension. This difference results in a larger estimate of *Ue* compared to the classic uncertainty metrics. In the revision, we highlighted the difference in the abstract, but the improvement was explained in detail in the main text.

"The new methods avoid pre-averaging in either of the spatiotemporal dimensions and as a result, the Ue estimate is around 20% higher than the classic uncertainty metrics."

The new uncertainty estimate is more comprehensive than the classic ones as the components are partially identified by the classic metrics.

--- unclear formulation

Reply: We rewrite this sentence as:

"Decomposing the formula for Ue shows that Ue has integrated four different variations across the ensemble dataset members, while only two of the components are represented in the classic uncertainty estimates."

Multiple precipitation products of different types (gauge-based, merged products and GCMs) are used to better explain and understand the peculiarity of the new methodology
--- unclear how e.g. gauges can explain the new methodology. Consider not using the word "peculiarity"

Reply: We intended to say that the new method is applied to the precipitation products which can be categorized into three groups: gauge-based products, merged products and GCMs. The results of the uncertainty analysis these different precipitation groups help explain the specifics of the new method.

We rewrote the sentence as following.

"The new approach is implemented and analyzed with multiple precipitation products of different types (e.g., gauge-based products, merged products and GCMs) which contain different sources of uncertainties with different magnitudes. Among the multiple gauge-based precipitation products, Ue is the smallest, while among other products Ue is generally larger because other uncertainty sources are included and the constraints of the observations are not as strong as in gauge-based products."

The comments above are about the Abstract. I can see similar problems in some other parts of the manuscript. This raises a concern, that the rest of the added and modified text in the manuscript would be also difficult to understand.

Therefore I encourage you to carefully read and revise the text again, giving attention to every sentence. Please also ask help form professional proof-reading services.

Reply: As the editor suggested, we carefully read and revised the manuscript. The modifications have been highlighted in the attachment. We also consulted proofreading service from a UK company, the certificate and the tracked changes can be found in the attachment as well.

Comments from Reviewers

Reviewer 1.

No comments.

Reviewer 2.

The manuscript entitled "A new uncertainty estimation among multiple datasets and implementation to various precipitation products" was revised based on the reviewer comments. All the comments were responded point by point. The improvement is significant. At the same time, some minor issues still exist, which are listed below.

1. In Line 15 on Page 15 of the revised manuscript with changes marked, "temporal mean (zone C3)" should be changed to "temporal mean (zone C5)".

   Reply: Yes. That should be zone C5 and we have corrected this mistake in the updated manuscript.

2. The meaning of i dimension in caption of Figure 2 is confusing. Does i represent 1, 2 and 3 for three dimensions?

   Reply: Yes, we intended to use i to represent the three dimensions in Figure 2. But since i has been used as the index for the temporal dimension in equations, we alternatively use x,y,z to represent one of the three dimensions (s,t,e) in Figure 2. The changes are highlighted in the caption:

"Partitioning the temporal-spatial-ensemble variance. The original database is re-organized into three dimensions: time, space and ensemble. Zones with different colours represent different processes based on the original database through different dimensions. The labels of the zones are listed on the right; detailed definitions can be found in Appendix A. The grand variance is σ^2 and the grand mean is μ. The subscripts t, s, and e indicate dimensions of time, space and ensemble, respectively. In Zone A, μ_x shows the mean values across the x-dimension (x=t, s or e); in Zone B, σ^2_x indicates the variation across the x-dimension; in Zone C, $\sigma_{x\_y}^2$ indicates the variation across the x-dimension of μ_y (y=t, s or e); in Zone D μ_{xy} indicates the means across the x- and y-dimensions; in Zone E, σ_{xy}^2 indicates the variation across the x- and y-dimensions; in Zone F, $\sigma_x^2(\mu_{yz})$ indicates the variation across the x-dimension of the means across the y- and z-dimensions (z=t, s or e)."

3. The "Enseble" in Figure 2 was wrongly spelled.

   Reply: It has been corrected in the updated figure.

4. Is Figure 4 the average annual precipitation over the period of 1979-2005 and over ensemble datasets of a group? Why does the gridded precipitation show the same color in many subregions?

   Reply: Yes, Figure 4 shows the average annual precipitation over 1979-2005 in grids for each group of the precipitation products (the caption has been revised for less confusion). We use a discrete intervals color bar rather than a continuous color bar in this graphic to reduce the difficulty of capturing the spatial patterns of the precipitation. In this case the grids in a same precipitation interval (400 mm/yr) are marked in the same color. We chose 400mm/yr because it is one of the criterions that distinguishes the climate types (e.g., dry area with precipitation less than 400mm/yr, semi-dry and semi-wet area with precipitation among 400-800mm/yr, wet area with precipitation larger than 800mm/yr). So, we can simply tell the climatic types from the precipitation using this discrete interval color bar.

5. Is the ensemble deviation for the first column in Figure 5 derived by the standard deviation across ensemble precipitation products in a specific group? It should be clearly explained. Similar information should be added in the caption of other figures.

   Reply: Yes, the standard deviation is estimated among the precipitation products within a specific group. For instance, the first row shows the results for gauge-based products and the second row shows the results for merged products and the third row is for GCMs.

   We have revised the captions for Figure 5:

   "The spatial distribution of model uncertainty in annual precipitation among different ensemble products. The uncertainty is expressed as the standard deviation of the annual precipitation across ensemble precipitation products of a specific group (up: gauge-based products, middle: merged products, bottom: GCMs). The left panels are the values of the uncertainty. The right panels are the ratios of ensemble deviation to the ensemble means of the datasets in the corresponding group."

The captions from Figure 6 to Figure 11 are also revised if necessary.

6. Why do the A1, A2 and A3 in Zone A of Appendix A all have l items? Is it a mistake in spelling?

   Reply: Sorry that it is our mistake. The upper bound should be *m, n, l* corresponding to *t, s* and *e* dimensions. The errors have been corrected in the Zone A and in the Zone B as well. A screenshot for the modification is attached for your check.

Appendix A: The algorithms for different expressions in the methodology

Zone A:

A1: ~~$\mu_t[s,e;n\times l];\mu_t[j,k]=\frac{1}{m}\sum_{i=1}^{l} z_{ijk}$~~ $\mu_t[s,e;n\times l];\mu_t[j,k]=\frac{1}{m}\sum_{i=1}^{m} z_{ijk}$

A2: ~~$\mu_s[e,t;l\times m];\mu_s[k,i]=\frac{1}{n}\sum_{j=1}^{l} z_{ijk}$~~ $\mu_s[e,t;l\times m];\mu_s[k,i]=\frac{1}{n}\sum_{j=1}^{n} z_{ijk}$

A3: $\mu_e[t,s;m\times n];\mu_e[i,j]=\frac{1}{l}\sum_{k=1}^{l} z_{ijk}$

Zone B:

B1: ~~$\sigma_t^2[s,e;n\times l];\sigma_t^2[j,k]=\frac{1}{m}\sum_{i=1}^{l}(z_{ijk}-\mu_t[j,k])^2$~~ $\sigma_t^2[s,e;n\times l];\sigma_t^2[j,k]=\frac{1}{m}\sum_{i=1}^{m}(z_{ijk}-\mu_t[j,k])^2$

B2: ~~$\sigma_s^2[e,t;l\times m];\sigma_s^2[k,i]=\frac{1}{n}\sum_{j=1}^{l}(z_{ijk}-\mu_s[k,i])^2$~~ $\sigma_s^2[e,t;l\times m];\sigma_s^2[k,i]=\frac{1}{n}\sum_{j=1}^{n}(z_{ijk}-\mu_s[k,i])^2$

B3: $\sigma_e^2[t,s;m\times n];\sigma_e^2[i,j]=\frac{1}{l}\sum_{k=1}^{l}(z_{ijk}-\mu_e[i,j])^2$

7. Please check the spelling in the manuscript carefully.

   Reply: Following your suggestions, we carefully revised the manuscript and consulted professional proofreading service to avoid the spelling errors in the manuscript.

Manuscript with changes highlighted.

A new uncertainty estimation ~~among~~ with multiple datasets and implementation ~~to~~ for various precipitation products

Xudong Zhou[1,2,3], Jan Polcher[2], Tao Yang[1], and Ching-Sheng Huang[1]

[1]State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Center for Global Change and Water Cycle, Hohai University, Nanjing 210098, China
[2]Laboratoire Météorologie Dynamique du CNRS, IPSL, CNRS, Paris, F 91128, France
[3]Institute of Industrial Science, The University of Tokyo, Tokyo, 〒153-8505, Japan

Abstract.

Ensemble estimates ~~of a certain climatic variable are frequently seen once many parallel~~ based on multiple datasets are frequently applied once many datasets are available for the same climatic variable. Uncertainty that evaluates the difference between the ensemble datasets is always provided along with the ensemble mean estimates to show to what extent the ensemble members are consistent with each other. However, one fundamental flaw of classic uncertainty estimates is that only the uncertainty in one dimension (either the temporal variability or the spatial heterogeneity) can be considered, whereas the variation along the other dimension is dismissed due to limitations in algorithms for classic uncertainty estimates, resulting in an incomplete assessment of the uncertainties. ~~Accompanying uncertainties evaluation with the ensemble is recommended while a fundamental flaw is that the uncertainties in temporal variation and spatial heterogeneityare not together consideredfor the final uncertainty estimate.~~ This study introduces a ~~new~~ three-dimensional variance partitioning approach ~~which avoids pre-averaging in either of the spatio-temporal dimensions. The newly proposed~~ and proposes a new uncertainty estimation (U_e) ~~which integrates~~ that includes the data uncertainties ~~across the spatio-temporal scalesis compared with classical uncertainties metrics. Results show that the~~ in both spatiotemporal scales. The new methods avoid pre-averaging in either of the spatiotemporal dimensions and as a result, the U_e estimate is around 20% higher than ~~classic metrics for the improvement of uncertainty estimation~~the classic uncertainty metrics. The deviation ~~between the metrics is higher~~ of U_e from the classic metrics is apparent for regions with strong spatial heterogeneity and where the ~~temporal~~ variations significantly differ ~~. Decomposing of~~ in temporal and spatial scales. This shows that classic metrics reduce the uncertainty estimate through averaging, which means a loss of information in the variations across spatiotemporal scales. Decomposing the formula for U_e ~~demonstrates that the new uncertainty estimate is more comprehensive than the classic ones as the components are partially identified by the classic metrics. Multiple~~ shows that U_e has integrated four different variations across the ensemble dataset members, while only two of the components are represented in the classic uncertainty estimates. This analysis of the decomposition explains the correlation as well as the differences between the newly proposed U_e and the two classic uncertainty metrics. The new approach is implemented and analyzed with multiple precipitation products of different types (e.g., gauge-based products, merged products and GCMs) ~~are used to better explain and understand the peculiarity of the new methodology. The new uncertainty estimation based on the~~ which contain different sources of uncertainties with different magnitudes. Among the

multiple gauge-based precipitation products, U_e is the smallest, while among other products U_e is generally larger because other uncertainty sources are included and the constraints of the observations are not as strong as in gauge-based products. This new three-dimensional approach is flexible in its structure and particularly suitable for a comprehensive assessment of multiple datasets over large regions within any given period.

Copyright statement.

1 Introduction

With the technical ~~development for monitoring the~~ developments in monitoring natural climate variables and the increasing knowledge of the physical mechanisms in the climate system, many institutes have the ability to provide different kinds of climate datasets. ~~Taken the~~ Taking precipitation, which is the dominant variable in the land water cycle, as an example, there are point measurements, such as GHCN-D (global historical climatology network-daily, Menne et al., 2012), gridded products based on gauge measurements and interpolation (e.g., CRU, Harris et al., 2014), products derived from remote sensing (e.g., the Tropical Rainfall Measuring Mission - TRMM), reanalysis datasets (e.g., NCEP) and ~~those~~ estimates from models (e.g., GCMs). These products ~~are~~ have been developed using different original data, technologies ~~or~~ and model settings for various purposes (Phillips and Gleckler, 2006; Tapiador et al., 2012; Beck et al., 2017; Sun et al., 2018). As a result, ~~differences exist among~~ there are differences between the various products due to ~~the~~ measurement errors, model biases~~or chaotic noises~~, or chaotic noise. The uncertainty is thus regarded as the deviation of these model results from ~~the~~ their real values.

However, the real values are difficult to measure and the uncertainties are difficult to ~~be removed~~ remove from the datasets. ~~Using~~ Thus, using ensembles consisting of multiple datasets to generate a weighted average ~~thus becomes~~ has become very popular in ~~the~~ climate-related ~~researches and the ensemble means are considered as the~~ research. The ensemble means of multiple datasets are considered more reliable estimates than a single dataset. For example, ~~the~~ IPCC uses 42 CMIP5 (Coupled Model Intercomparison Project Phase 5) models to show historical temperature changes and 39 CMIP5 models to average ~~the temperature projection in future~~ future temperature projections in RCP 8.5 scenario (Figure SPM.7 in IPCC, 2013b). Schewe et al. (2014) use nine global hydrological models to evaluate ~~the~~ global water scarcity under climate change. GLDAS (Global Land Data Assimilation System) involves four different land surface models (Rodell et al., 2004) and GRACE (Gravity Recovery and Climate Experiment) provides ~~estimations~~ estimates from three independent institutes (Landerer and Swenson, 2012). Using multiple datasets reduces the dependence on a single dataset and ~~eliminate~~ eliminates the random variations associated to biases or ~~noises in the model estimates~~noise in each single model estimate.

Along with the ensemble means, uncertainty information is recommended to be presented because the ~~uncertainty level decides~~ level of uncertainty determines the reliability of the ensemble results. In general, uncertainties can be quantified as the range of maximum and minimum values (i.e., $V_{max} - V_{min}$), ~~range of values~~ the value difference at different quantiles (e.g., $V_{5\%} - V_{95\%}$), the consistency of models (ratio of models following a certain pattern to the total number of models), the variation (σ^2) or the standard deviation (σ) among multiple model estimations. These metrics ~~represent different characteristics of the multiple datasets~~describe the differences between multiple model estimates in different aspects. Among the metrics, the standard deviation (σ) is the most used because it has the same ~~magnitude~~ unit as the original dataset~~; it avoids influence of extreme samples and it~~. Moreover, it is less sensitive to extreme samples and to the number of datasets used for the investigation. The ratio of the standard deviation (σ) to the mean value (μ), the so-called coefficient of variance (CV), representing the dispersion or spread of the distribution of various ensemble members (Everitt, 2013), is a ~~unit-less~~ unitless value which also shows the degree of uncertainty efficiently.

Depending on the purpose of the data evaluation, the uncertainty ~~among~~ between the datasets can be displayed ~~over~~ or visualized in space to show the spatial heterogeneity~~of the consistency among multiple datasets~~. For example, the predicted future temperature increase has a higher significance in the northern high-latitudes among different models than in the middle-latitudes (Box TS.6 Figure 1 in IPCC, 2013a). ~~The other~~ Another typical implementation is to evaluate the evolution of the ~~model~~ uncertainty over time. In general, the ~~uncertainty range~~ range of the uncertainty decreases in the historical period over time because more observations ~~are accessible in recent while the range increases for~~ have been accessible recently. But the uncertainty increases in future projections because of the increasing spread of ~~the model simulations~~ model estimates (Figure SPM.7 in IPCC, 2013b)~~. The increasing uncertainty range indicates the~~, indicating a decreasing of consistency ~~and increasing variations~~ but increasing variation among various datasets.

~~The above metrics have been widely used as they show the temporal evolution or spatial distribution~~ The two kinds of ways can easily show the spatial distribution or the temporal evolution of the uncertainty~~easily. But the~~. But a short-coming is apparent~~as only the mean value across~~, as the variation along one dimension (time or space) ~~is used when we~~ has to be collapsed to generate the mean values when we attempt to assess the uncertainty for the other dimension (~~time or space).~~ space or time). For example, the averaging over a specific region ~~(spatial mean)~~ to obtain the spatial mean is estimated at each time step before obtaining the temporal evolution of the model uncertainty ~~can be obtained~~ (red flowcharts in Figure 1). ~~And the~~ In contrast, averaging over a certain ~~period (temporal mean) is estimated at~~ temporal period to obtain the temporal mean is necessary for each grid cell ~~before the spatial distribution of the model uncertainty can be obtained~~ when estimating the spatial variations of model uncertainties (blue flowcharts in Figure 1). ~~While, the averaging~~The averaging, in either dimension, means a loss of ~~the information , for instance the data variation . The~~ information about the variation in the data. Any changes in the variation that leaves the mean values unchanged will not be propagated to the ~~uncertainty estimation if the mean value remains the same. This may result in that the uncertainty among datasets not being~~ global uncertainty estimation. The result of this is that the variations between datasets is not fully considered when estimating the uncertainties. In other words, ~~either~~ neither of the uncertainty estimates ~~cannot represent the full peculiarities among datasets. Therefore, the uncertainty among datasets~~ can represent the whole of the differences between multiple datasets. The uncertainty can be underestimated, and the similarity ~~among them overestimatedwith these two procedures. However current studies have~~ of the datasets thus overestimated. Indeed, the current literature has not paid attention to the ~~ignorance of variation due to the~~ ignoring of variation after averaging as well as its influence on the ~~uncertainty assessment~~ assessment of the uncertainty.

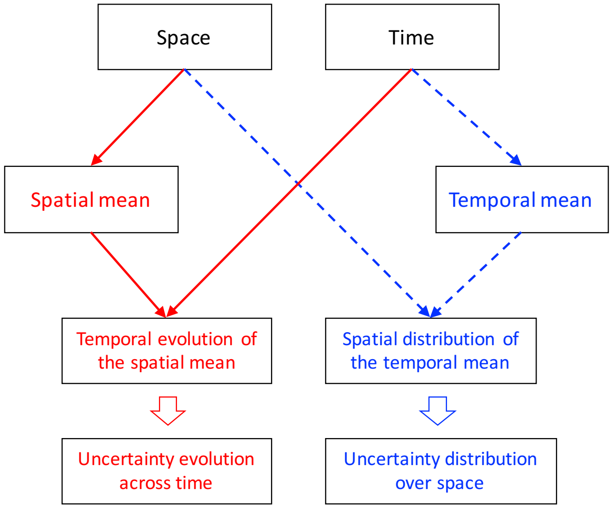

Figure 1. The two classic uncertainty assessments in the current ~~researches as~~ literature: the temporal evolution of the model uncertainty (red) and the spatial distribution of the model uncertainty (blue). ~~Either~~ Each of these estimations of the uncertainty ~~estimates~~ has ~~to do the averaging in~~ average over one of the dimensions~~in~~ , either space or time, ~~and it~~ which will lead to ~~the loss of~~ losing information ~~in~~ about the corresponding dimension.

The total variation among ~~the~~ multiple datasets is contributed ~~by the uncertainties, temporal variation and the spatial heterogeneity.~~ to by the spatial heterogeneity, temporal variability and the model uncertainties. To some degree, the model uncertainty is similar to other dimensions as a variation along a third dimension (ensemble dimension). The key to ~~evaluate the~~ evaluating the model uncertainty is to decompose the variation caused by ~~dataset differences from the others. Though the variation decomposition with method analysis of variance~~ differences between the datasets from the other two contributors. Although decomposing the variation by means of ANalysis Of VAriance (ANOVA) is often seen in hydro-metrological studies, ~~it is always used~~ this is designed to separate the process uncertainties generated in ~~series steps that propagated~~ different model processes that propagate to the final variation. For example, Déqué et al. (2007) ~~separated~~ decomposed the uncertainties of ~~regional climate models~~ Regional Climate Models (RCM) ~~to~~ into four sources of ~~uncertainties (~~uncertainty: sampling uncertainty, model uncertainty, radiative uncertainty and boundary uncertainty~~), and boundary uncertainty plays a greater role~~. Bosshard et al. (2013) decomposed the uncertainty in ~~the~~ river streamflow projections to uncertainties from climate models, statistical ~~postprocessing~~ post-processing schemes and hydrological models. These implementations differ from the ~~scope~~ purpose of the present study ~~and~~ because they fail to separate the uncertainties from the ~~spatio-temporal variations because spatio-temporal averaging has been~~ spatiotemporal variations because spatiotemporal averaging was already applied in the estimation process. Sun et al. (2010, 2012) ~~in~~ for the first time decomposed the total variation ~~to~~ into temporal variation and spatial heterogeneity. ~~However, it~~ They concluded that the variations along the spatial dimension contributed more to the total variation than did the temporal variabilities. However, their method is only valid for ~~the~~ one single dataset and is thus not able to evaluate the uncertainties if multiple datasets describe the same variable. But a generalized method should be based on Sun's work, as one more dimension can be added for a specific analysis of the uncertainties.

In ~~this~~ the present study, we aim to introduce a new approach ~~for uncertainty estimation~~ to estimating uncertainty among multiple datasets. The new uncertainty metric ~~avoids any averaging in~~ should avoid any averaging over time or space~~dimension thus all the information across the~~ , so that all information along each of these two dimensions can be maintained for the ~~uncertainty assessment~~ assessment of the uncertainty. Multiple precipitation products ~~are used to explain the peculiarity~~ will be used to display the results and explain the specifics of the new ~~methodology. In section~~ method. In Section 2, the detailed ~~methodology~~ method of the three-dimensional variance partitioning approach is introduced. The characteristics of multiple precipitation datasets and estimations of ~~the two~~ two other classic uncertainty metrics are shown in ~~section~~ Section 3. The results of the new approach for ~~the~~ precipitation products are discussed in ~~section~~ terms of the types of precipitation datasets in Section 4. The differences between the new uncertainty estimation and two selected classic metrics ~~introduced previously~~ used in uncertainty analysis are analyzed and discussed in ~~section~~ Section 5. ~~The discussion and conclusions are followed in the end of this article.~~ A discussion and some conclusions follow in Section 6.

2 ~~Methodology~~ Method and datasets

2.1 Mathematical Derivation

~~The database consists of multiple datasets that record~~ Multiple datasets recording the same climatic ~~variables in spatio-temporal scale. The database has to be organized in three dimensions of~~ variable should be reorganized into a three dimensional database, using the dimensions (1) **time** with a regular time interval (e.g. monthly or annual), (2) **space** with regular spatial units~~where~~ , with all the grids ~~are~~ re-organized ~~in a new~~ into one dimension from the original ~~latitude-longitude grids,~~ longitude–latitude grids, and (3) **ensemble** ~~with different ensemble datasets regarded~~ as the third dimension describing the different ensemble datasets. Thus, the dataset array can be ~~reformed as~~ re-organized to be

$$\mathbf{Z} = [z_{ijk}] \tag{1}$$

with the i-th time step ($i = 1, 2, \ldots, m$), j-th grid ($j = 1, 2, \ldots, n$), and k-th ensemble member or ensemble model ($k = 1, 2, \ldots, l$).

We define the three dimensions ~~as~~ to be time, space and ensemble dimension, and the means for these three dimensions ~~are called~~ to be the temporal mean, spatial mean and ensemble mean~~, respectively~~. The corresponding variances are ~~named time variance, space variance~~referred to as the temporal variance, spatial variance, and ensemble variance~~, respectively. The~~ . We also define the grand mean (μ), grand variance (σ^2) ~~across time, space and ensemble dimensions as well as~~ and the total sum of squares (SST) ~~are defined as.~~ (or total variation) across the entire database:

$$\mu = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{l} z_{ijk}/(mnl) \tag{2}$$

$$\sigma^2 = \frac{SST}{mnl} \tag{3}$$

$$SST = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{l}(z_{ijk}-\mu)^2. \tag{4}$$

The total variation ~~is contributed by the variation in all~~ receive contributions from the variations along all three dimensions (Eq. 4). ~~Thus, it should~~ It can be reformulated as an ~~express of variations in~~ expression in terms of the variations along each of the three different dimensions. ~~The~~ For instance, the derivation of the total ~~squares~~ variation can start from the third ensemble dimension. For a specific k^{th} ensemble member, the grand mean is formulated as $\mu_{ts}[k] = \sum_{i=1}^{m}\sum_{j=1}^{n} z_{ijk}/(mn)$, leading to the total ~~squares rewritten as~~ sum of squares being rewritten as

$$SST = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{l}(z_{ijk}-\mu_{ts}[k]+\mu_{ts}[k]-\mu)^2. \tag{5}$$

The SST can be further expanded and rearranged as

$$\begin{aligned} SST = \quad & \textstyle\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{l}(z_{ijk}-\mu_{ts}[k])^2 \\ & +2\times \textstyle\sum_{k=1}^{l}(\mu_{ts}[k]-\mu)\underbrace{\left[\sum_{i=1}^{m}\sum_{j=1}^{n}(z_{ijk}-\mu_{ts}[k])\right]}_{=0} \\ & +\underbrace{\left[\sum_{i=1}^{m}\sum_{j=1}^{n}\right]}_{=mn}\textstyle\sum_{k=1}^{l}(\mu_{ts}[k]-\mu)^2 \end{aligned} \tag{6}$$

$$SST = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{l}(z_{ijk}-\mu_{ts}[k])^2 + mn\sum_{k=1}^{l}(\mu_{ts}[k]-\mu)^2 \tag{7}$$

$$SST = mn\sum_{k=1}^{l}\sigma_{ts}^2[k] + mnl\sigma^2(\mu_{ts}), \tag{8}$$

~~Where~~ where $\sigma^2(\mu_{ts})$ is the variation of the grand mean for each ~~member of the ensemble ,~~ ensemble member and $\sigma_{ts}^2[k]$ ~~,~~ is the grand variance in ~~space and time for~~ the spatial and temporal dimensions for the ensemble member k~~,~~. Moreover, $\sigma_{ts}^2[k]$ can be split using the mean of the spatial variation at each time step $\overline{\sigma_s^2[k,:]}$ and the variation of the spatial mean $\sigma^2(\mu_s[k,:])$, denoted as

$$\sigma_{ts}^2[k] = \overline{\sigma_s^2[k,:]} + \sigma^2(\mu_s[k,:])$$

~~The detailed derivation of~~ in Eq. (9) ~~is shown in Eqs .~~ with its derivation given in Eqs (10)~~-~~–(17).

$$\sigma_{ts}^2[k] = \overline{\sigma_s^2[k,:]} + \sigma^2(\mu_s[k,:]). \tag{9}$$

For a specific dataset k, the grand mean $\mu_{ts}[k]$ ~~through space-time scale is~~ at the spatiotemporal scale is

$$\mu_{ts}[k] = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} z_{ijk}. \tag{10}$$

The total ~~square for difference~~ sum of squares of the differences from the grand mean of this ensemble member is

$$SST[k] = \sum_{i=1}^{m}\sum_{j=1}^{n}(z_{ijk} - \mu_{ts}[k])^2 \tag{11}$$

and the grand variance σ_{ts}^2 is

$$\sigma_{ts}^2[k] = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(z_{ijk} - \mu_{ts}[k])^2. \tag{12}$$

The derivation can start from either the ~~space~~ spatial dimension or the temporal dimension. If the derivation starts from the ~~space~~ spatial dimension, Eq. (11) can be rewritten by incorporating the spatial mean of each time step $\mu_s[k,i] = \sum_{j=1}^{l} z_{ijk}/n$

$$SST[k] = \sum_{i=1}^{m}\sum_{j=1}^{n}(z_{ijk} - \mu_s[k,i] + \mu_s[k,i] - \mu_{ts}[k])^2. \tag{13}$$

~~It~~ This can be expanded and then rearranged as

$$\begin{aligned} SST[k] = &\sum_{i=1}^{m}\sum_{j=1}^{n}(Z_{ijk} - \mu_s[k,i])^2 \\ &+ 2\times\sum_{i=1}^{m}(\mu_s[k,i] - \mu_{ts}[k]) \times \underbrace{\left[\sum_{j=1}^{n}(Z_{ijk} - \mu_s[k,i])\right]}_{=0} \\ &+ \underbrace{\left[\sum_{j=1}^{n}\right]}_{=n}\sum_{i=1}^{m}(\mu_s[k,i] - \mu_{ts}[k])^2 \end{aligned} \tag{14}$$

$$SST[k] = \sum_{i=1}^{m}\sum_{j=1}^{n}(Z_{ijk} - \mu_s[k,i])^2 + n\sum_{i=1}^{m}(\mu_s[k,i] - \mu_{ts}[k])^2 \tag{15}$$

$$\begin{aligned} SST[k] &= n\sum_{i=1}^{m}\sigma_s^2[k,i] + nm\sigma^2(\mu_s[k,:]) \\ &= nm\overline{\sigma_s^2[k,:]} + mn\sigma^2(\mu_s[k,:]). \end{aligned} \tag{16}$$

The grand variance of this specific dataset is Eq. 17 (identical to Eq. 9).

$$\sigma^2_{ts}[k] = \frac{SST[k]}{mn} = \overline{\sigma^2_s[k,:]} + \sigma^2(\mu_s[k,:]). \tag{17}$$

Here, $\overline{\sigma^2_s[k,:]}$ is the mean of the spatial variation at each time step and $\sigma^2(\mu_s[k,:])$ is the variation of the spatial mean.

Or if we started the derivation from the time dimension, the grand variance can be split using the average of the temporal variation from all regions $\overline{\sigma^2_t[:,k]}$ and the ~~space~~ spatial variation of the temporal mean $\sigma^2(\mu_t[:,k])$:

$$\sigma^2_{ts}[k] = \overline{\sigma^2_t[:,k]} + \sigma^2(\mu_t[:,k]). \tag{18}$$

With Eq. (9) or Eq. (17) and Eq. (18), we ~~can have~~ obtain

$$\sigma^2_{ts}[k] = \frac{1}{2}\left\{[\sigma^2(\mu_t[:,k]) + \overline{\sigma^2_s[k,:]}] + [\sigma^2(\mu_s[k,:]) + \overline{\sigma^2_t[:,k]}]\right\}. \tag{19}$$

Substituting Eq. (19) into Eq. (8) results in

$$\begin{aligned} SST = & \frac{mn}{2}\sum_{k=1}^{l}[\sigma^2(\mu_t[:,k]) + \overline{\sigma^2_s[k,:]}] \\ & + \frac{mn}{2}\sum_{k=1}^{l}[\sigma^2(\mu_s[k,:]) + \overline{\sigma^2_t[:,k]}] + mnl\sigma^2(\mu_{ts}). \end{aligned} \tag{20}$$

The first term on the right-hand side of Eq. (20) can be transformed to ~~:~~

$$\frac{mn}{2}\sum_{k=1}^{l}[\sigma^2(\mu_t[:,k]) + \overline{\sigma^2_s[k,:]}] = mnl\left[\frac{\overline{\sigma^2_{s\_t}} + \overline{\sigma^2_s}}{2}\right], \tag{21}$$

where $\overline{\sigma^2_{s\_t}}$ is the mean ~~of space~~ value across ensemble members of the spatial variation of the temporal mean~~across each ensemble member,~~ , and $\overline{\sigma^2_s}$ represents the grand mean of σ^2_s, which is the grand variance across ~~time~~ the temporal and ensemble dimensions. Eq. (20) then becomes ~~:~~

$$SST = mnl\left[\frac{\overline{\sigma^2_{s\_t}} + \overline{\sigma^2_s}}{2}\right] + mnl\left[\frac{\overline{\sigma^2_{t\_s}} + \overline{\sigma^2_t}}{2}\right] + mnl\sigma^2_e(\mu_{ts}), \tag{22}$$

where $\overline{\sigma^2_{t\_s}}$ ~~is the mean of time~~ denotes the mean value across ensemble members of the temporal variation of the spatial mean~~across ensembles~~, $\overline{\sigma^2_t}$ ~~represents~~ denotes the grand mean of σ^2_t, the grand variance across space and ensemble dimensions~~.~~ , and $\sigma^2_e(\mu_{ts})$ ~~represents the variation~~ denotes the variation across ensemble members of the spatial-temporal means ~~(μ_{ts}).~~ .

Similarly, the global derivation of SST can start from any of the other two dimensions ~~. And the SST derived from time and space dimensions are formulated , respectively,~~ (i.e., space or time). This derivation can then be formulated as

$$SST = mnl\left[\frac{\overline{\sigma^2_{s\_e}} + \overline{\sigma^2_s}}{2}\right] + mnl\left[\frac{\overline{\sigma^2_{e\_s}} + \overline{\sigma^2_e}}{2}\right] + mnl\sigma^2_t(\mu_{se}) \tag{23}$$

$$SST = mnl\left[\frac{\overline{\sigma^2_{e\_t}} + \overline{\sigma^2_e}}{2}\right] + mnl\left[\frac{\overline{\sigma^2_{t\_e}} + \overline{\sigma^2_t}}{2}\right] + mnl\sigma^2_s(\mu_{et}), \tag{24}$$

~~Where~~ where each variable is defined in ~~the~~ Appendix A. Averaging these three expressions of SST defined in Eqs ~~.~~ (22)~~-~~–(24) leads to

$$
\begin{aligned}
SST = \quad & \tfrac{mnl}{3}[\tfrac{\overline{\sigma^2_{t\_s}}+\overline{\sigma^2_{t\_e}}}{2}+\overline{\sigma^2_t}+\sigma^2_t(\mu_{se})] \\
& +\tfrac{mnl}{3}[\tfrac{\overline{\sigma^2_{s\_t}}+\overline{\sigma^2_{s\_e}}}{2}+\overline{\sigma^2_s}+\sigma^2_s(\mu_{et})] \\
& +\tfrac{mnl}{3}[\tfrac{\overline{\sigma^2_{e\_t}}+\overline{\sigma^2_{e\_s}}}{2}+\overline{\sigma^2_e}+\sigma^2_e(\mu_{ts})].
\end{aligned}
\tag{25}
$$

With the total ~~degree of freedom (~~number of degrees of freedom being $m \times n \times l$~~)~~, the grand variance is expressed as

$$
\begin{aligned}
\sigma^2 = \quad & \underbrace{\frac{1}{3}[\frac{\overline{\sigma^2_{t\_s}}+\overline{\sigma^2_{t\_e}}}{2}+\overline{\sigma^2_t}+\sigma^2_t(\mu_{se})]}_{V_t} \\
& +\underbrace{\frac{1}{3}[\frac{\overline{\sigma^2_{s\_t}}+\overline{\sigma^2_{s\_e}}}{2}+\overline{\sigma^2_s}+\sigma^2_s(\mu_{et})]}_{V_s} \\
& +\underbrace{\frac{1}{3}[\frac{\overline{\sigma^2_{e\_t}}+\overline{\sigma^2_{e\_s}}}{2}+\overline{\sigma^2_e}+\sigma^2_e(\mu_{ts})]}_{V_e},
\end{aligned}
\tag{26}
$$

where V_t, V_s and V_e ~~represent the time, space~~ denote the temporal, spatial and ensemble variances, respectively. ~~To~~ An illustration of the present approach is shown in Figure 2 to facilitate the understanding of the partitioning results~~, an illustration of the present approach is shown in Figure 2~~. The original database, consisting of multiple datasets, is re-organized into three dimensions (grey in the centre). Zones with different colors represent different processes of the original database from different dimensions (see the details in the caption of Figure 2 and Appendix A).

Note that the ensemble variance V_e ~~is estimated based on the combination of variation~~ in Eq. (26) is a combination of several variations across the ensemble ~~dimension~~members. The four components are the variations of temporal and spatial values ($\overline{\sigma^2_e}$, zone B3), temporal mean ($\overline{\sigma^2_{e\_t}}$, zone ~~C3~~C5), spatial mean ($\overline{\sigma^2_{e\_s}}$, zone C6) and the grand variance of the spatiotemporal mean for a single ensemble member ($\sigma^2_e(\mu_{ts})$, zone F3). Similarly, the other variances only rely on the variances in the corresponding dimension, which shows the independence ~~in~~ of the three dimensions. This also is an illustration of the fact that the uncertainty across ensemble members is similar to the temporal variation and spatial heterogeneity.

2.2 ~~Metrics definition~~ Definitions of the metrics for model uncertainty

~~Since the temporal variation or~~ Although the total variation is a result of contributions from the spatial heterogeneity~~is natural in the climate variables and the purpose of this study is to evaluate the model uncertainty among~~ , temporal variability, and the uncertainties across different datasets, we ~~focus mainly~~ mainly focus on the variance in the ensemble dimension because the spatial or temporal variation is natural for climatic variables. The uncertainty among ~~the ensemble member~~ ensemble members is normalized as the ratio of the square root of the ensemble variance (V_e) ~~divided by the~~ to the grand mean value of the datasets (μ).

$$
U_e = \sqrt{V_e}/\mu \tag{27}
$$

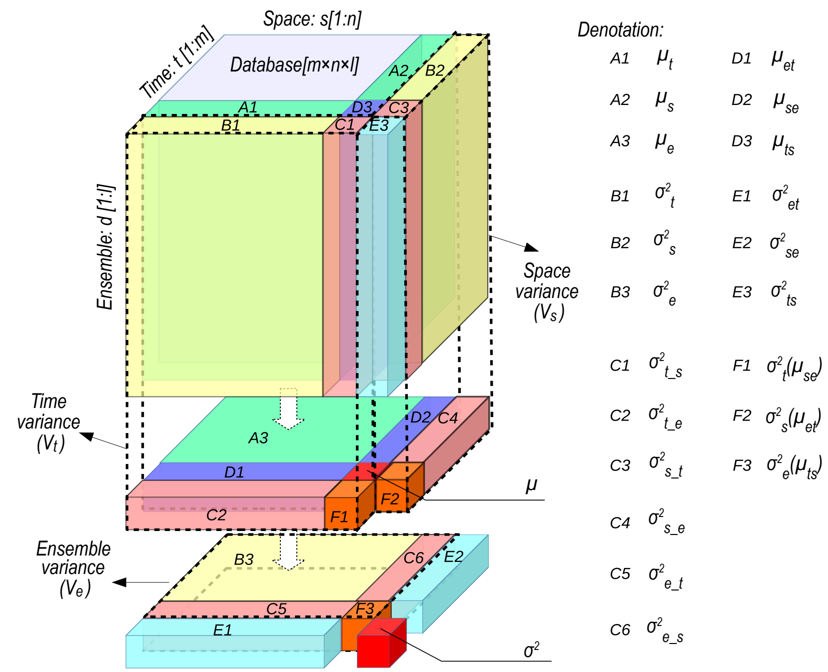

Figure 2. ~~The illustration of~~ Partitioning the ~~partitioning time-space-ensemble~~ temporal-spatial-ensemble variance~~method~~. The original ~~dataset~~ database is ~~reorganized~~ re-organized into three dimensions~~of~~: time, space and ensemble. Zones with different colours represent different processes based on the original database through different dimensions. The ~~denotations~~ labels of the zones are listed ~~to~~ on the right; detailed definitions can be found in Appendix A. The grand variance is ~~defined as~~ σ^2 and the grand mean ~~as~~ is μ. The subscripts t, s, and e ~~represent~~ indicate dimensions of time, space and ensemble, respectively. In Zone A~~(μ_i) indicates~~, μ_x shows the ~~means of~~ mean values across the ~~i dimension; zone B~~ x-dimension (~~σ_i^2~~x=t, s or e); in Zone B, σ_x^2 indicates the variation ~~for i dimension~~across the x-dimension; ~~zone~~ in Zone C~~($\sigma^2_{i\_j}$)~~, $\sigma^2_{x\_y}$ indicates the variation across ~~i dimension of~~ the ~~means~~ x-dimension of ~~μ_j; zone D~~ μ_y (~~μ_{ij}~~y=t, s or e); in Zone D μ_{xy} indicates the means across ~~i~~ the x- and ~~j dimensions~~y-dimensions; ~~zone~~ in Zone E~~(σ^2_{ij})~~, σ^2_{xy} indicates the variation across ~~i~~ the x- and ~~j dimensions~~y-dimensions; ~~zone~~ in Zone F~~($\sigma_i^2(\mu_{jk})$)~~, $\sigma_x^2(\mu_{yz})$ indicates the variation across ~~i dimension~~ the x-dimension of the means across ~~j~~ the y- and ~~k dimensions~~z-dimensions (z=t, s or e). ~~The detailed definitions of these denotations can be found in Appendix A.~~

Two ~~classical~~ classic metrics are also introduced for comparison. For each basic spatial unit (~~grid cell in this study~~ in the present study this means a grid cell), we can estimate the ~~long-term~~ temporal mean of the target variable ~~for each dataset~~ in each ensemble dataset as $\mu_t[j,k]$, $j = 1,...,n$ represents the ~~space~~ spatial unit, and $k = 1,...,l$ represents the ~~number of datasets. Then for each spatial unit,~~ index of the dataset. Then we can estimate the ~~ensemble~~ variations across different ensemble datasets of the mean values as $\sigma^2(\mu_t[j,:])$ (expressed as $\sigma^2_{e\_t}[j]$ in this study). The spatial distribution of the $\sigma^2_{e\_t}$ shows the magnitude of the model uncertainty over space and its root $\sigma_{e\_t}[j]$ is the model deviation at each ~~space~~ spatial unit. The ~~overall estimation of the model uncertainty~~ estimate of this model deviation over the entire region can be expressed as ~~:~~

$$N.s.std = \sqrt{\overline{\sigma^2_{e\_t}}}/\mu = \frac{1}{\mu}\sqrt{\frac{1}{n}\sum_{j=1}^{n}\sigma^2_{e\_t}[j]}. \tag{28}$$

~~$\sigma^2_{e\_t}[j]$ has different values for~~ For each spatial unit~~and the~~, $\sigma^2_{e\_t}[j]$ $(j = 1,...,n)$ can take a different value. The values for all the grid cells are averaged to obtain $\overline{\sigma^2_{e\_t}}$, which shows the general magnitude of the ensemble variation over space. The quantity $N.s.std$ is normalized as the ratio of the square root of the ~~mean of~~ averaged variations $\sqrt{\overline{\sigma^2_{e\_t}}}$ to the ~~average value~~ grand mean of all the datasets μ.

Similarly, the model uncertainty can ~~also~~ be normalized as the ratio of the square root of the averaged ensemble variation ~~at all~~ but at different time steps $\overline{\sigma^2_{e\_s}}$ to the entire means~~(Eq. 29).~~:

$$N.t.std = \sqrt{\overline{\sigma^2_{e\_s}}}/\mu = \frac{1}{\mu}\sqrt{\frac{1}{m}\sum_{i=1}^{m}\sigma^2_{e\_s}[i]}, \tag{29}$$

where ~~the $\sigma^2_{e\_s}[i], i = 1,...,m$ is the ensemble variation of the spatial mean of each dataset~~ $\sigma^2_{e\_s}[i]$ $(i = 1,...,m)$ is the variation across different datasets of the spatial means of each product at each time unit $\mu_s[i,k], (i = 1,...,m, k = 1,...,l)$. ~~It has different values at different time steps.~~

The two uncertainty estimates (Eqs ~~.~~28 and 29) correspond to the two classic metrics presented in the Introduction. ~~And we will compare the~~ We will compare U_e with ~~the~~ these two classic metrics ($N.t.std$ and $N.s.std$) to show their relations and ~~perecularities~~differences.

2.3 Study area and data description

~~China is large in its area with different climate types encountered in the mainland (Kottek et al., 2006). To~~ Mainland China has been selected as the study area because of its large area and different types of climate (Kottek et al., 2006). Ten different subregions have been defined to facilitate the comparisons and ~~analyses that have spatial variations, ten different subregions are defined in Figure 3 as the~~ analysis of the strong spatial variations. The subregions are (1) Songhua River Basin, (2) Liao River Basin, (3) Hai River Basin, (4) Yellow River Basin, (5) Huai River Basin, (6) Yangtze River Basin, (7) Southeast China, (8) South China, (9) Southwest China, (10) Northwest China, see Figure 3. The entire Chinese mainland is numbered as the 11~~st~~th region. Most of the ~~regions~~ subregions are natural river basins~~, and~~; this definition is more ~~proper when considering water resources~~ appropriate for water resource analysis than definitions using ~~longitude-latitude grids or that are~~ longitude–latitude grids or those based on administrative regions.

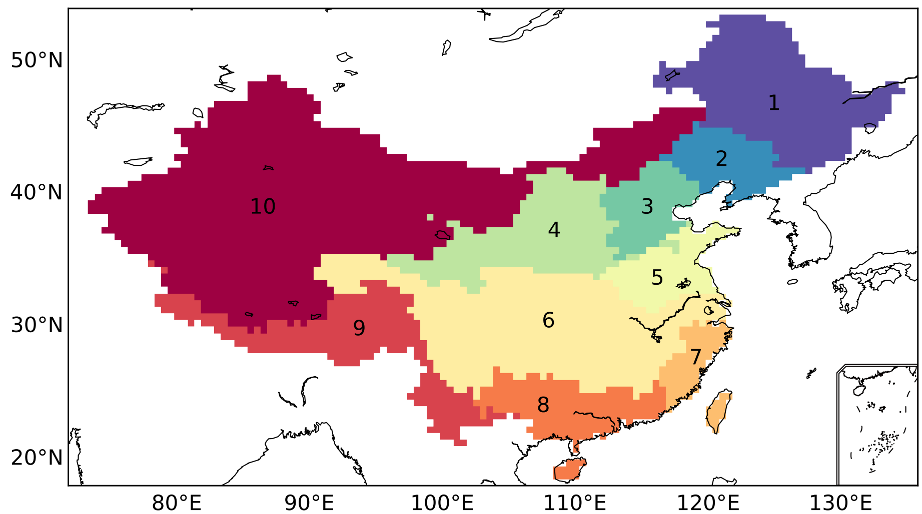

Figure 3. Ten subregions are ~~identified~~ defined in this study. These subregions are mainly ~~divided as the~~ river basins (~~regions 1-8~~Regions 1–8)~~and~~, but 9 ~~as the southwestern~~ is Southwest China and 10 ~~as the northwestern~~ is Northwest China. ~~The~~ Region 11 ~~represents~~ is the ~~whole~~ entirety of the Chinese mainland.

~~Thirteen~~ Precipitation is one of the climatic variables sensitive to large-scale atmospheric cycles and the local topography. Thirteen different precipitation datasets from ~~different sources are~~ various sources have been collected for comparison (Table 1). These datasets ~~are~~ have been categorized into three groups according to the ~~methodologies used to generate~~ methods they used for generating the products, ~~i.e.~~namely, gauge-based products, merged products and General Circulation Models (GCMs). The gauge-based products (~~i.e.,~~ namely, CMA, GPCC, CRU, CPC and UDEL) use ~~observed data~~ data observed from global precipitation gauges~~, while the density of~~. The density of the ground observation gauges, the representativeness of the gauges, and the interpolation algorithms for converting the gauge observations to ~~gridded dataset vary~~ a gridded dataset differ from product to product. ~~CMA (stands for~~The CMA (China Meteorological Administration) dataset ~~uses the densest~~ has the densest distribution of gauges and probably has the best quality to capture the spatiotemporal variations of the precipitation over the study area. ~~But CMA~~ The CMA dataset is excluded when estimating the ~~ensemble means of~~ uncertainty among the gauge-based products~~and~~: it is chosen as the reference ~~datasets~~ dataset for comparison.

Among the merged precipitation products, the CMAP, GPCP and MSWEP use different sources of precipitation data (~~e.g.~~namely, gauge observations, satellite remote sensing, and atmospheric model re-analysis). These different precipitation sources are averaged using different weights. Thus, the differences ~~among~~ between the three merged products are associated with the precipitation sources and the weight of the gauge observations. ERA-Interim is a re-analysis product~~, while~~: it uses near-real-time assimilation with data from global observations (Dee et al., 2011). Thus, the forecasting model is constrained by the observations and forced to follow the real system to some degree. Because of ~~the usage~~ its use of observations, ERA-interim ~~is also belonging to the~~ also belongs to the category of merged products.

GCM precipitation is ~~model estimation, therefore, the~~ a pure model estimation because observations are not used to constrain the simulations. The implemented physical and numerical ~~choices~~ processes will affect the accuracy of the model results. ~~In addition, observations are not used to constrain the simulations.~~ The lack of constraints on the GCMs will cause them ~~not following~~ to not follow the actual synoptic variability and explore other trajectories in the solution space. Kay et al. (2015) repeatedly ~~run~~ ran the same GCM with a very small ~~difference~~ shift in the initial conditions~~, and there is a spread of~~. But the small difference leads to a spread in the model outputs after a number of ~~time steps of running~~ running time steps (see Figure 2 in Kay et al., 2015). Therefore, the uncertainty ~~estimated is due~~ in GCMs can be attributed to the differences in the model ~~settings~~ structures, parameter settings, and the initial conditions as well. There are more than 20 ~~datasets of GCMs, while only four are randomlytaken to match the number of~~ kinds of different GCMs; only 4 of them have been chosen, randomly, to maintain the same number of datasets using the gauge-based products ~~and~~ as those using merged products.

All the products of the three precipitation types~~including CMA~~, including CMA, are in gridded format. ~~Though~~ Although they differ in ~~the~~ their original spatial resolution, all products ~~are interpolated to~~ have been interpolated to a 0.5^{o} spatial resolution to unify the spatial units. Annual average values are summed ~~up~~ based on their original time steps (daily or monthly) and the overlap time span of all the datasets is ~~selected~~ from 1979 to 2005 for ~~the maximum coverage of~~ all products.

3 Characteristics of precipitation and model quantified uncertainties with classic metrics

3.1 Spatial patterns of ~~ensemble~~ annual precipitation

The ~~ensemble means of the~~ long-term annual ~~precipitation (1979-2005~~mean precipitation (1979–2005) obtained by averaging the precipitation from multiple datasets in the corresponding precipitation group ~~are~~ is mapped in Figure 4. The ~~long-term~~ annual mean precipitation obtained from the CMA ~~data~~ dataset is 589.8 $\mathrm{mm\ yr^{-1}}$ (1.6 $\mathrm{mm\ day^{-1}}$) over ~~mainland China~~the entire Chinese mainland. The gauge-based precipitation has the least bias (-4.1mm $\mathrm{yr^{-1}}$, -0.7% in ~~proportion~~percentage) compared to the CMA precipitation. ~~Precipitation~~ The precipitation in the merged products and GCMs is larger than that of the CMA by 63.1 and 232.0 $\mathrm{mm\ yr^{-1}}$ (with the bias ~~as~~ equal to +10.7% and +39.3%), respectively.

The spatial pattern of the annual precipitation shows a decreasing gradient from ~~the southeastern~~ Southeast China (>1600 $\mathrm{mm\ yr^{-1}}$) to ~~the northwestern~~ Northwest China (<400 $\mathrm{mm\ yr^{-1}}$) ~~. All the ensemble means of the~~ in CMA and all other three precipitation groups~~capture the spatial gradient, while they have different ability to express in some details~~. They have different abilities to display the spatial gradient of the precipitation in some detail. For instance, some areas have abrupt precipitation changes rather than ~~following~~ follow the general gradient in CMA. This is probably caused by the sudden changes in topography (e.g., the northern Tienshan Mountain, the Qilian Mountains), ~~while it~~ which is not captured in the gauge-based products ~~. As we know, the precipitation gauges are mainly distributed on the lower altitude and therefore, they have difficulty in capturing the precipitation events over mountains~~because some of the key gauges are not included in the production of the gauge-based products. The abrupt changes can be somehow represented by merged products and GCMs because the local variation due to topographic changes can be observed by other methods or by model algorithms. The precipitation in the merged products and the GCMs is higher than ~~CMA in Himalayas~~ that of CMA in the Himalayas, and particularly the GCMs

Table 1. The precipitation datasets used in this study. Three different precipitation groups ~~are~~ have been identified according to the way the precipitation dataset is generated.

No.	Type	Name	Long name	Institute	Reference
1		CMA	China Meteorological Administration dataset	China Meteorological Administration, Beijing, China	
2		GPCC	Global Precipitation Climatology Centre	the World Climate Research Programme (WCRP) and to the Global Climate Observing System (GCOS)	Schneider et al. (2017)
3	Gauge-based	CRU TS	Climatic Research Unit Time-Series	Climatic Research Unit (CRU) / Ian Harris, Phil Jones, UK	Harris et al. (2014)
4		CPC	CPC Global Unified Gauge-Based Analysis of Daily Precipitation	NCEP/Climate Prediction Center, Maryland, USA	Xie et al. (2007)
5		UDEL	University of Delaware Air Temperature & Precipitation Global (land) precipitation and temperature	University of Delaware, Delaware, USA	Willmott and Matsuura (2012)
6		CMAP	CPC Merged Analysis of Precipitation	NOAA CPC, Maryland, USA	Xie et al. (2003)
7	Merged Products	GPCP	Global Precipitation Climatology Project	GSFC (NASA), Maryland, USA	Adler et al. (2018)
8		MSWEP	Multi-Source Weighted-Ensemble Precipitation	Princeton University, Princeton, NJ, USA	Beck et al. (2017)
9		ERA-I	ERA-Interim	European Centre for Medium-Range Weather Forecasts, Reading, UK	Dee et al. (2011)
10		HadCM3	Hedley Centre Coupled Model Version 3	Met Office Hadley Centre, Exeter, UK	
11	GCMs	IPSL-CM5A-LR		~~Insitut~~ Institut Pierre Simon Laplace, Paris, France	
12		CMCC-CM		~~Cetro~~ Centro Euro-Mediterraneo per ~~I Cambiamenti~~ i Cambiamenti Climatici, Lecce, Italy	
13		MIROC5		AORI, Chiba, Japan, NIES, Ibaraki, Japan, JAMSTEC, Kanagawa, Japan	

show higher precipitation in the ~~northern~~ North Tibet Plateau as well as the southern part of the Hengduan Mountains. These differences show the general characteristics of the three types of precipitation products.

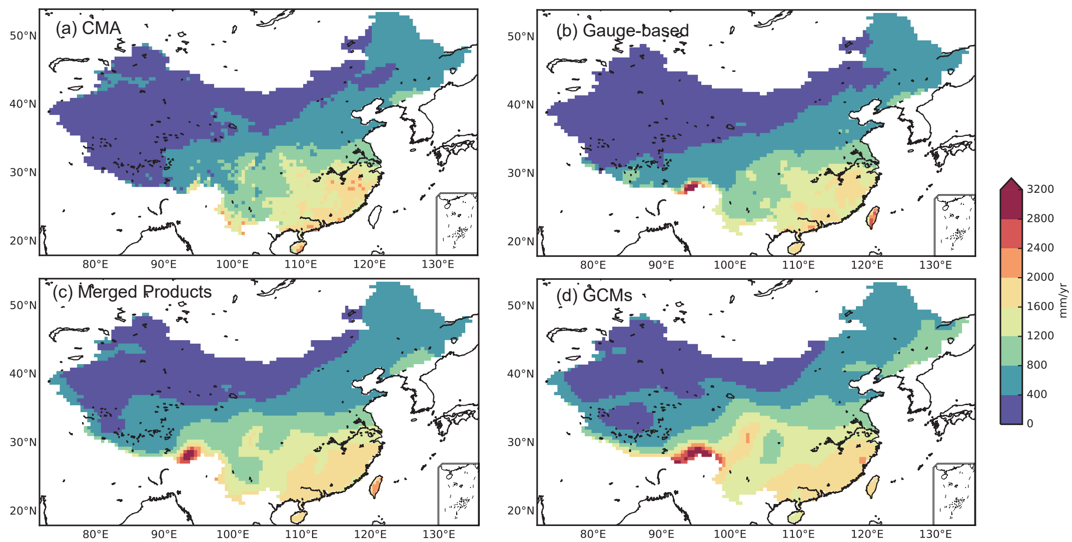

Figure 4. ~~Long-term~~ The annual precipitation over a long-term period (~~1979-2005~~1979–2005) ~~annual~~ for each group of precipitation ~~in different precipitation groups~~datasets. (a) Annual precipitation of CMA dataset, (b) ensemble means of the annual precipitation over the precipitation products in gauge-based ~~products~~ precipitation excluding CMA, (c) ensemble mean of the annual precipitation of all merged products, (d) ensemble means of the annual precipitation of all GCMs. The observations in Taiwan are not ~~included~~ released in the CMA dataset.

3.2 Spatial distribution of model uncertainties

In addition to ~~differences of the~~ the precipitation differences in its long-term annual ~~precipitation, differences are found among~~ means, differences can be found between datasets within the same precipitation group. The spatial distribution of the model uncertainty for each precipitation group, which is expressed as the ensemble deviation ~~across multiple products~~ of the annual precipitation ~~, is calculated for each group and~~ from different precipitation products, is mapped in Figure 5.

Among the datasets based on gauge observations~~(Figure 5-a)~~, the ensemble deviation value is small in most of the land area of China (<50 $\mathrm{mm\ yr^{-1}}$~~). It~~ , Figure 5-a). Although the deviation is higher in the south of China (50-100 $\mathrm{mm\ yr^{-1}}$)~~but~~ , the area is not continuous in space. The highest deviation occurs along the Himalayas, indicating a high variation among the observed datasets. Regarding the merged precipitation products, the deviation shows high values (>200 $\mathrm{mm\ yr^{-1}}$, Figure 5-c) in ~~the southwestern~~ Southwest China (e.g., the Tibet Plateau, Yunnan Province, Guangxi Province). Moderate deviation is found in ~~the northeastern China, northern China and southeastern China.~~ Northeast China, North China and Southeast China. The deviation of precipitation has a correlation with the topology, which indicates that the performance of the technologies

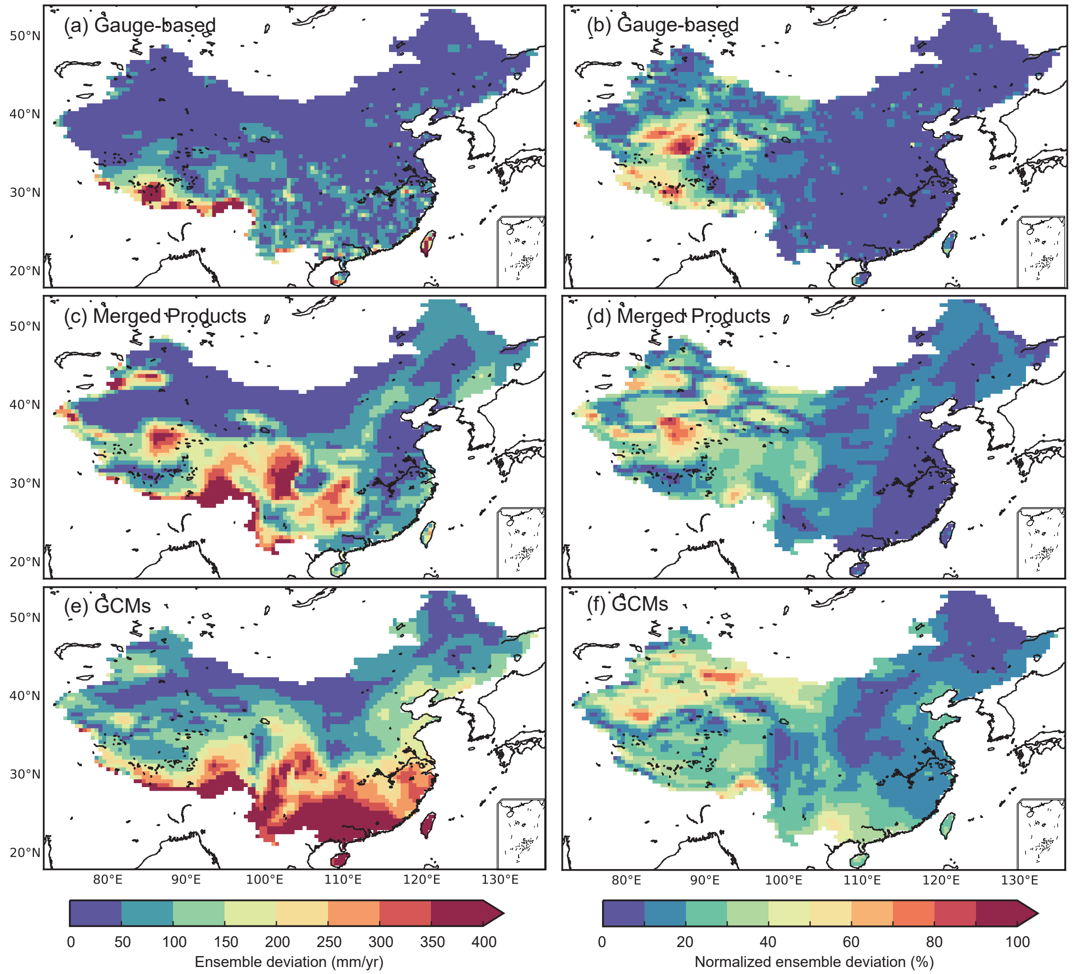

Figure 5. The spatial distribution of ~~the~~ model uncertainty ~~, which~~ in annual precipitation among different ensemble products. The uncertainty is expressed as the ~~ensemble~~ standard deviation ~~across multiple products~~ of the ~~long-term~~ annual precipitation ~~in each~~ across ensemble precipitation products of a specific group (up: gauge-based products, middle: merged products, bottom: GCMs). The left panels ~~) and~~ are the ~~normalized value as the ratio~~ values of the uncertainty. The right panels are the ratios of ensemble deviation to the ensemble means of the datasets in the corresponding group~~(right panels)~~.

used for the merged products are subject to the topologies as well. Compared to the gauge-based and merged products, the deviation among the selected GCMs has the highest value (>400 $\mathrm{mm\,yr^{-1}}$, Figure 5-e) in ~~the southern~~ South China, indicating a significant model uncertainty of the annual precipitation between different GCMs.

The ratio of the ensemble deviation to the mean value, which shows the model uncertainty with no ~~unit~~units, is very low in East China (<10%, Figure 5-b)~~in the eastern China. While, it~~. It is higher in ~~the western~~ West China especially in the Himalayas and the ~~northern~~ North Tibet Plateau. Similar to that of the gauge-based products, the uncertainty in the merged products has ~~the~~ higher values in the ~~west than that in the east~~ West than in the East of China (Figure 5-d). The area with ~~the derivation~~ a deviation ratio less than 10% is mainly distributed in ~~the southeastern~~ Southeast China and is apparently smaller than that of the gauge-based products, showing a decreasing similarity among different merged products. The area with a moderate ~~derivation~~ deviation ratio (10%~~-40~~–40%) increases compared to that of the gauge-based products, and the area is mostly in ~~the middle~~ central and western China. The uncertainty estimated in the GCMs shows similar patterns in ~~western~~ West China to that of the merged products but with higher magnitudes in ~~the eastern~~ East China (Figure 5-f). Only the area in the ~~northeastern~~ Northeast and part of ~~the middle~~ central China features small uncertainty, less than 10%, and the ~~derivation~~ deviation ratio rises significantly in ~~the southern~~ South China (e.g., the Pearl River basin), which corresponds to the high standard ~~deviation of~~ deviations in the GCMs shown in Figure 5-e.

The magnitude of the ensemble deviation demonstrates the model uncertainty among ~~different precipitation~~ the different products in the same precipitation group and ~~it~~ shows the ability ~~of the precipitation estimation with different methodologies~~to estimate the precipitation with different methods. For all products, the ensemble deviation is relatively larger where the precipitation is higher, especially along the mountains and the subtropical regions. The ~~derivation~~ deviation ratio is higher in ~~the northwestern China~~Northwest China, where the precipitation is among the lowest in China. Particularly for the gauge-based products, ~~the higher ratio occurs~~ higher ratios occur where the gauge density is low and the orographic effect is apparent (e.g., the Tibet Plateau and ~~the~~ other mountainous area). For the merged products and the GCMs, the deviation ratio increases especially in ~~the southeastern~~ Southeast China, showing decreasing similarities among different ~~GCMs~~precipitation products. Because the deviation ratio has taken into account both the variation and the means (which may have a systematic bias), the ~~derivation~~ deviation ratio is better than the absolute ensemble deviation ~~to represent the uncertainty. Thus~~ at representing the uncertainty, and it is the most commonly used in ~~the~~ geographic studies.

3.3 Temporal evolution of model uncertainties

Figure 5 shows the spatial distribution of the ensemble deviation among different precipitation products. However, the temporal evolution of the deviation~~among the various products is~~, which shows the performance of product over time and its changes, are not captured because the temporal variation has been averaged ~~before estimating the~~ in order to estimate the spatial ensemble deviation in Figure 5. In this ~~section~~subsection, we examine the temporal evolution of ~~model uncertainty of the~~ the uncertainties in regional annual precipitation ~~across different~~ among different ensemble products. The analysis is based on the ten subregions defined in Figure 3 and the ~~whole~~ entire Chinese mainland.

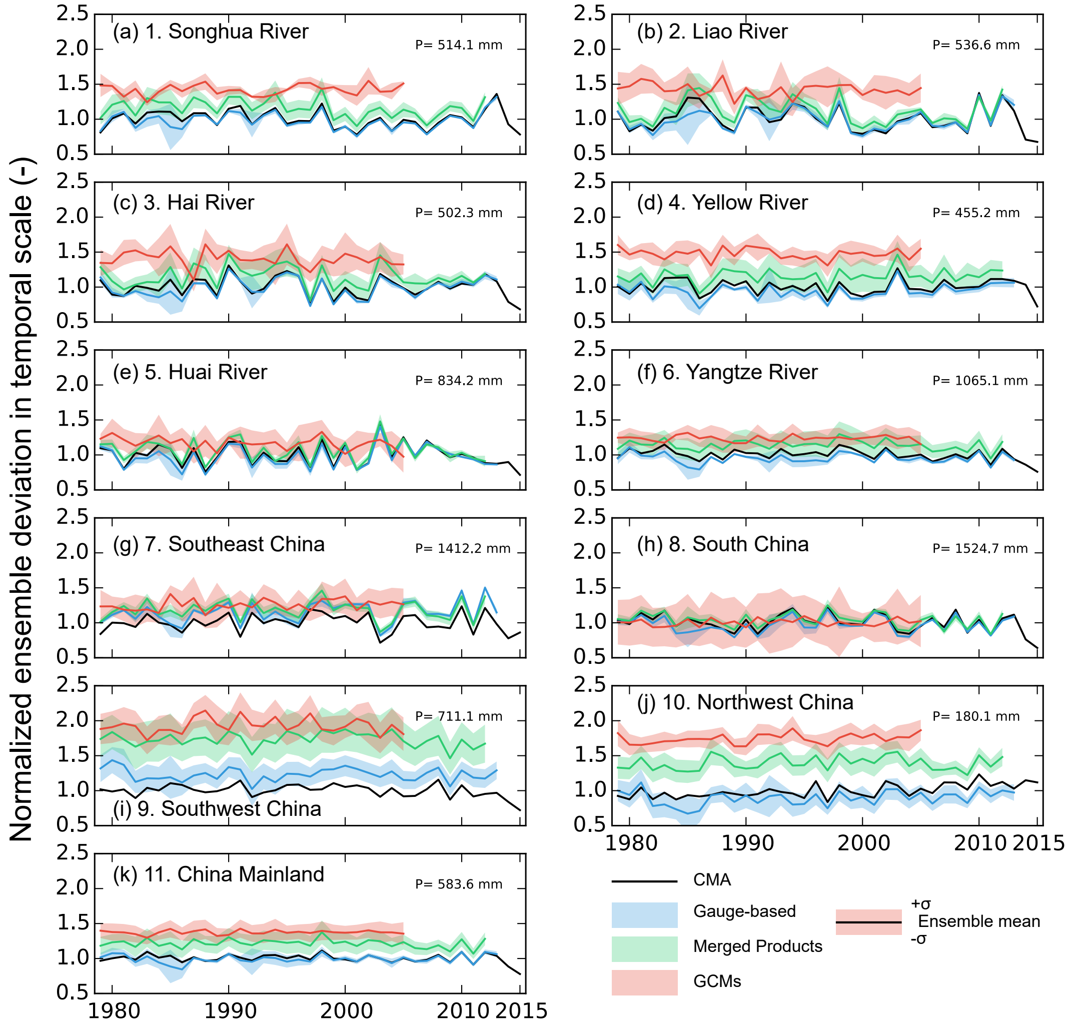

Figure 6. The temporal evolution of the model uncertainty~~, which~~. The uncertainty is expressed as the normalized ensemble deviation of annual precipitation across ensemble datasets in each precipitation group for specific subregions. The value on the top right of each panel is the annual regional precipitation estimated in CMA dataset (~~1979-2015~~1979–2015). The annual precipitation is normalized as the ratio to the CMA long-term annual precipitation. The solid curve represents the ensemble mean of precipitation in each precipitation data group over the subregion. The width of the shaded area represents the standard deviation of the annual precipitation in each year among the datasets within that group (divided by the annual precipitation of the corresponding group). The shaded area ~~distributes~~ is equally distributed in the two sides of the ensemble mean values for the corresponding precipitation group.

The annual precipitation of each precipitation group has been normalized as the ratio to the long-term annual ~~means of CMA~~ mean of the CMA in each subregion (black line in Figure 6). The magnitude of the annual precipitation in the gauge-based products (the blue ~~curve~~solid line) is similar to that of CMA except in ~~the southwestern~~ Southwest China (Figure 6-i) for the overestimation along the Himalayas (Figure 4-a,b). The precipitation in the merged products (the green ~~curve~~solid line) is higher in ~~the southwestern and northwestern~~ Southwest and Northwest China, in accordance with Figure 4-c. The annual precipitation of the GCMs (the red ~~curve~~solid line) is apparently higher than that of the gauge-based products ~~or~~ and merged products for almost all regions, which agrees with the spatial patterns in Figure 4-d.

The ensemble deviation ~~(shaded area) shown~~ across time scale is shown in the shaded area in Figure 6~~represents the variations of the~~. It is estimated as the deviation of regional annual precipitation among different products in the same ~~precipitation group . The normalized deviation facilitates the~~ group at a specific time step for each subregion. The deviation in normalized to facilitate comparisons between different ~~regions~~subregions. High deviations are found in ~~all three precipitation groups in the southwestern~~ Southwest China (Figure 6-i) in all three precipitation groups because of the large differences along the Himalayas. The deviations among the gauge-based products and the merged products in other regions are small and getting smaller with time. ~~It~~ This is mainly because more observations are ~~integrated and technologies improve~~ included and technologies have improved with time to control the ~~data quality~~ quality of the data. A large deviation is found in the merged products in ~~10-northwest~~ 10-Northwest China (Figure 6-j) and ~~the~~ 4-Yellow River Basin (Figure 6-d), where a dry climate dominates and the annual precipitation is among the lowest~~and dry climate dominates~~. The model deviation of GCMs varies ~~among~~ between regions as it is ~~smallest in the~~ at its smallest in 1-Songhua River Basin (Figure 6-a) and ~~the~~ 6-Yangtze River Basin (Figure 6-f), while it is among the highest in ~~the 8-south China and the west~~ 8-South China and West China (9,10), agreeing with the deviation maps in Figure 5.

Despite ~~of the difference in mean values~~ their mean values and magnitudes of deviation, the temporal evolution of the gauge-based products and merged products agree well with ~~that~~ those of the CMA dataset, while the temporal evolution of ~~GCMs ensemble~~ the members of the category of GCMs is weaker and not well correlated with that of the CMA. The main reason is that GCMs are not constrained in their synoptic variability and the sequence of ~~the~~ wet and dry years can be very different from that of the observations. ~~So, a smoother result can be~~ A smoother result is thus obtained when we build the ensemble ~~means~~ mean from the GCMs. ~~While this is different for the~~ Unlike the weak variation in GCMs, the gauge-based and merged products ~~, as they~~ have a strong co-variance and the ensemble mean preserves this co-variance.

For the entire ~~mainland of China~~ Chinese mainland (Figure 6-k), the ensemble deviation remains stable ~~for~~ in different precipitation groups. In contrast, the annual precipitation spans the ~~largest~~ strongest spatial heterogeneity in the mainland compared to those divided by subregions (Figure 4). However, the spatial variation has been collapsed ~~when estimating~~ because the regional precipitation ~~for~~ has to be obtained before the temporal analysis. It is therefore interesting to ~~see~~ evaluate how the uncertainty ~~estimate~~ changes when the variations ~~in~~ along both the time dimension and ~~in the space~~ the spatial dimension are considered ~~together~~ in the precipitation datasets.

3.4 Variations ~~in~~ along the ~~time~~ temporal and ~~space~~ spatial dimensions

~~The precipitation varies in time and space; however, it is averaged either in the time dimension to obtain the spatial patterns of model uncertainty (Figure~~ 5~~) or in the space dimension to obtain the temporal evolution of the model uncertainty (Figure~~ 6~~). But the deviations in the time and space dimensions are indeed very rarely compared~~ Previous subsections provide the deviation analysis in either temporal scale or the spatial scale. However, the two are seldom compared with each other. Herein, the standard deviation of the temporal and spatial variations in the precipitation datasets are compared in Figure 7 in ten subregions and the ~~Chinese~~ China mainland for different precipitation groups.

The gauge-based products provide similar annual regional precipitation to CMA over the China mainland and ~~ten specific regions~~ the ten specific subregions except for the region ~~7-southeast~~ 7-Southeast China (Figure 7-g) and region ~~9-southwest China (~~9-Southwest China (Figure 7-i). While the merged products ~~show~~ provide larger precipitation estimations for most of the regions. It might indicate the degraded ability of remote sensing, ~~the important data source~~ one of important data sources in the merged products, to estimate the precipitation amount in storms as the storms mainly contribute to the total precipitation for the two subregions. The regional precipitation ~~is larger in merged products than that of observations and the magnitude of the deviation in~~ in GCMs is even larger except in the region ~~8-south~~ 8-South China (Figure 7-h). These results indicate the degraded ability of merged products and GCMs in reproducing the total value of the annual precipitation.

Regarding the ~~variations in time and space dimensions, the~~ temporal and spatial deviations, regions 9, 10 and 11 have the largest ~~ratio of the~~ spatial standard deviation (as a ratio to the mean), indicating the ~~most significant~~ strongest spatial heterogeneity over the regions. ~~The 7-southeast~~ Regions 7-Southeast China and the 3-Hai River have the smallest variations, either because of ~~the~~ their small area or because ~~of~~ the homogeneity in ~~the subregion as the spatial correlation is highin the area. The relative~~ these subregions is high. However, the spatial deviation in most of the subregions is larger than the temporal deviation. The ratio of the temporal ~~standard~~ deviation to the spatial ~~standard~~ deviation is among the smallest in the ~~regions~~ subregions 9, 10 and 11 (k=0.1, 0.12 and 0.05, respectively. k is the ratio of the temporal deviation to the spatial deviation), showing an apparent difference between the ~~variation in the time and space~~ variations along the two dimensions. While, the difference between ~~variation in~~ the variations along the two dimensions is small in the 3-Hai River basin (k=1.15) and ~~7-southeast~~ 7-Southeast China (k=0.90), mainly ~~because~~ due to the relatively strong variability of the annual precipitation in different years.

In addition to the differences ~~across~~ between regions, the variations in different precipitation groups also vary in magnitude. Excluding the CMA dataset~~which only consists of~~, which consists of only one single product, the total variation (the sum of the spatial and temporal ~~variation~~variations) in the gauge-based products ~~are~~ is higher than that of the other two groups. ~~The~~ This difference demonstrates that ~~on one hand~~ the gauge-based products may have the largest ~~variation over space or on the other hand the correlation among~~ spatial variation, and the correlations between the different gauge-based products are high~~so that the~~, so that this variation is preserved when ~~doing~~ passing to the ensemble. ~~On the contrary~~In contrast, the GCMs have the smallest variations, either because the precipitation estimated in the GCMs are more ~~homogenous over space~~spatially homogenous than those of other precipitation products, or because the ~~spatial patterns~~ precipitation estimations in different GCMs are not consistent ~~and the spatial correlation is lower since there is no constrain in~~ in time or space since there are no

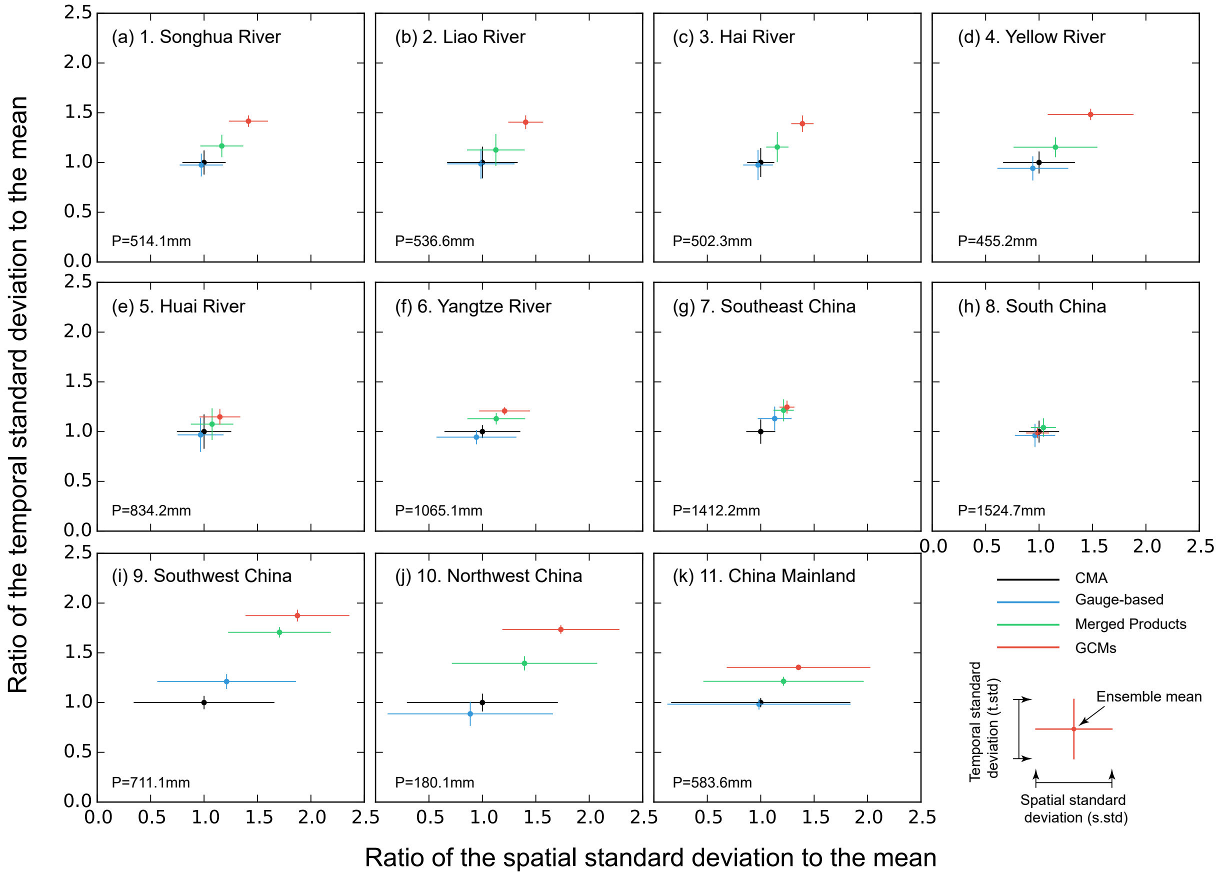

Figure 7. The spatial standard deviation (horizontal) and temporal standard deviation (vertical) of the annual precipitation across ensemble datasets in each of the different precipitation groups for ~~ten regions and~~ each subregion. The P value in the ~~mainland China~~bottom left is the annual precipitation of CMA. The cross ~~center~~ centre represents the long-term means of the regional annual precipitation in ratio to the CMA mean value. The horizontal error bar represents the spatial standard deviation (spatial variation of the long-term annual precipitation at all the grids). The vertical error bar represents the temporal standard deviation (temporal variations of region-averaged annual precipitation in different years). ~~The P values in the left bottom is the annual precipitation of CMA.~~

constraints on the GCM simulation. The inconsistent precipitation patterns will be further eliminated when carrying out an ensemble averaging over multiple datasets.

4 Variances in precipitation products

4.1 Variances in three dimensions

~~We have introduced the general~~ In the preceding section, we introduced the spatial and temporal characteristics of the ~~precipitation in different groups and their variations in different dimensions in the above section~~annual precipitation. The variations in the precipitation in two dimensions among different precipitation products in the same precipitation group were estimated by two classic methods. In this section, we will present the ~~results that~~ uncertainty results estimated by the newly proposed ~~variance approach~~ approach to the variance. As introduced in the ~~methodology section,~~ methods section, the input annual precipitation to the approach is re-organized into three dimensions~~as~~: (1) **time**, 27 years from 1979 to 2005, (2) **space**, ~~the number of~~ 0.5^o grids in a specific region and (3) **ensemble**, the number of ~~the models in a same~~ models in each precipitation group (four models ~~in all~~ for each of the three groups). Note that the estimated variance is for a specific subregion because it is an analysis based on regions and a long-term scale.

The grand variance (V, total value of the variance for all three dimensions) and its three components (i.e., variance in time ~~,~~ ~~space~~ V_t, space V_s and ensemble dimension V_e) for all the subregions ~~are~~ is mapped in Figure 8. The grand variance ~~(total value of the variance for all three dimensions) is similar for data groups of~~ is similar in space in the precipitation groups of the gauge-based products and the merged products (Figure 8-a,b,c), while the grand variance in ~~GCMs is large and is approximating twice the values of~~ the GCMs is larger and is approximately twice the V in the other two groups in regions ~~9-south China and 10-southwest~~ 9-South China and 10-Southwest China. The differences are mainly constituted by the ~~space~~ spatial variance and ensemble variance (Figure 8-i,l).

The ~~time variance (~~temporal variance V_t ~~)~~ is the smallest among all three ~~variance proportions, and there are~~ variances, and it has very little differences ~~of V_t in the northern~~ in North China (Figure 8-d,e,f). ~~V_t~~ But it is higher in the gauge-based products ~~is higher than that~~ than in the merged products and GCMs in regions ~~8-southeast China and 9-south~~ 8-Southeast China and 9-South China, indicating a relatively strong temporal variation in the annual precipitation series~~which consists~~, in accordance with the larger uncertainty ranges shown in Figure 6-h,i. Similar patterns of the ~~space variance (~~spatial variance V_s ~~)~~ are found in the gauge-based products and merged products (Figure 8-g,h). ~~The 7-Yangtze~~ Regions 7-Southeast River basin and ~~9-southwest~~ 9-Southwest China have the largest V_s because the precipitation significantly varies in space in these two ~~regions. V_s~~ subregions: it is higher in ~~the precipitation of GCMs especially in the 9-southwest~~ GCM precipitation especially in 9-Southwest China, indicating the strong spatial heterogeneity in the GCM models over the Himalayas (Figure 8-i). The ensemble variance ~~(V_e)~~ is relatively small in most regions ~~in~~ for gauge-based products (Figure 8-j), ~~with the highest V_e occurring in 9-southwest China. It indicates~~ indicating that the model variation ~~between~~ among datasets in the observation group is small. ~~Similar small values of~~ A similarly small V_e ~~are~~ is found in the northern regions ~~in~~ among the merged products as well as in the GCMs for the regions in ~~the northern~~ North China, while the intra-ensemble variations are large in the ~~south~~

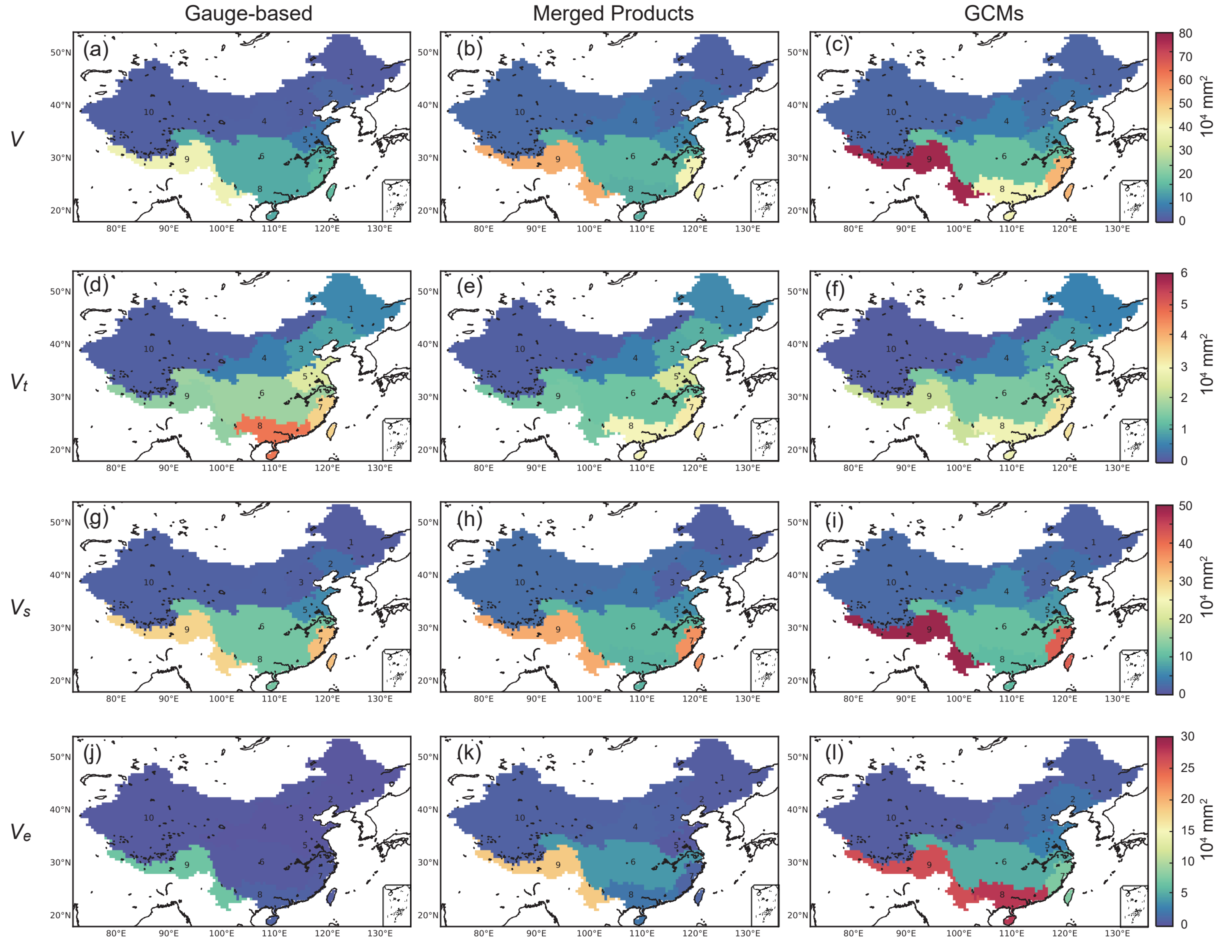

Figure 8. ~~The maps~~ Maps of the estimated grand variance (~~V~~V) and variances in different dimensions (V_t, V_s, V_e) ~~for~~ across the ensemble datasets in each of the three different precipitation groups.

~~especially the 9-southwest China and 8-south China in the GCMs~~ GCMs, especially in the South, especially 9-Southwest China and 8-South China (Figure 8-k,l).

~~In conclusion,~~ One can conclude that the grand variance and individual variance for each of the three different dimensions are generally larger in the ~~dataset~~ precipitation group consisting of GCMs. The variations for the gauge-based products and merged products are similar in values and spatial distribution. However, in addition to the variances, the ~~uncertainty~~ deviation defined as the ratio of the square root of the variance to the mean (~~i.e.~~e.g., U, U_t, U_s, U_e) contains extra information ~~of~~ about the regional means, and will be discussed in the following section.

4.2 Deviations in three dimensions

In contrast to the spatial ~~patterns of the variance magnitude distributed in~~ gradient of the magnitude of the variance distributed over the ten subregions (Figure 8), the larger values of the total deviation ($U = \sqrt{V}/\mu$) ~~occur in the northwest, and lower values occur in the southern China in general~~ occurs in the Northwest, but a lower value generally occurs in South China (Figure 9). ~~A possible reason is the~~ The decreasing tendency of ~~precipitation magnitude~~ magnitude of the precipitation from the southeast to the northwest ~~(Figure 4). Although the variances are among the lowest in the northwest China, the total deviation~~ results in a shift of the spatial gradient compared to Figure 4. The total deviation U is the highest in ~~this region~~ Northwest China (U=0.89, Figure 9-a,b,c) for all three precipitation groups~~because of the low precipitation rate in the northwest. U~~, but is relatively small in the northeastern 1-Songhua River (U=0.27) ~~in the northeast~~ and 8-South China (U=0.29) for the gauge-based products~~and~~. Subregion 6-Yangtze River has a relatively lower U in the merged products and GCMs in the ~~east~~ eastern part of China.

~~The variations in time and space dimension~~ Deviations along the temporal and spatial dimensions are inherent, ~~and~~ as they show the temporal evolution and spatial heterogeneity of the ~~characteristics in different~~ precipitation products. ~~It is found that the~~ The results show that U_t is small and contributes very little to the total U, indicating the weak fluctuation of annual precipitation compared to ~~spatial variations~~ the spatial heterogeneity (Figure 9-d,e,f). The smallest ~~U_t value~~ value of U_t for the GCMs is in accordance with the weakest temporal variations in Figure 6. The ~~relative variance in space~~ deviation in the spatial dimension (U_s) contributes the most to the total ~~variance~~deviation, especially in ~~the northwestern~~ Northwest China (U_s=0.77 for the gauge-based products, Figure 9-g). The high ~~values indicate~~ U_s indicates the strong spatial heterogeneity of precipitation in the region~~compared to the mean values. It indicates~~, demonstrating that the ability to describe the precipitation ~~significant varies~~ varies significantly in different places in the subregions. However, because the spatial variations ~~characterized by GCMs in the northwestern China is~~ obtained by the GCMs in Northwest China are less significant than with the other two groups, the value of U_s for region ~~10-southwest~~ 10-Southwest China (=0.51) is smaller than that of the gauge-based and merged products.

The variations ~~in time and space~~ along the temporal and spatial dimensions show the natural precipitation patterns but the deviation of the values ~~at same spatiotemporal points show the ability of the products~~ among multiple products (U_e) shows the ability to consistently represent the spatiotemporal patterns. ~~The relative variance in the ensemble dimension (~~Therefore, U_e ~~) shows the variations among different products in~~ indicates the uncertainty of the precipitation products among ensemble members of the same group. For the gauge-based products, ~~the~~ U_e is smaller than 0.1 for regions in ~~the eastern~~ East China, indicating that the model ~~differences~~ variations are relatively small compared to the annual means. The ~~U_e value~~ value of U_e is higher for ~~the 9-southwest~~ 9-Southwest China (=0.30) and ~~10-northwest~~ 10-Northwest China (=0.37), showing large variations even in the gauge-based products. For the merged products, U_e is similar to that of the gauge-based products in ~~the western~~ West China (=0.36), while it is larger in the ~~east especially for the~~ East, especially for 6-Yangtze River and 4-Yellow River (more than two times larger than U_e of the gauge-based products).

For the GCM precipitation, ~~the uncertainty~~ U_e increases compared to the other two groups in the eastern ~~regions~~subregions, corresponding to the higher ~~ensemble variations in GCM~~ spatial model uncertainty in GCMs over the eastern regions shown in Figure 5. ~~While, it decreases in 10-northwest~~ It decreases in 10-Northwest China (U_e=0.25) and a possible reason for this is

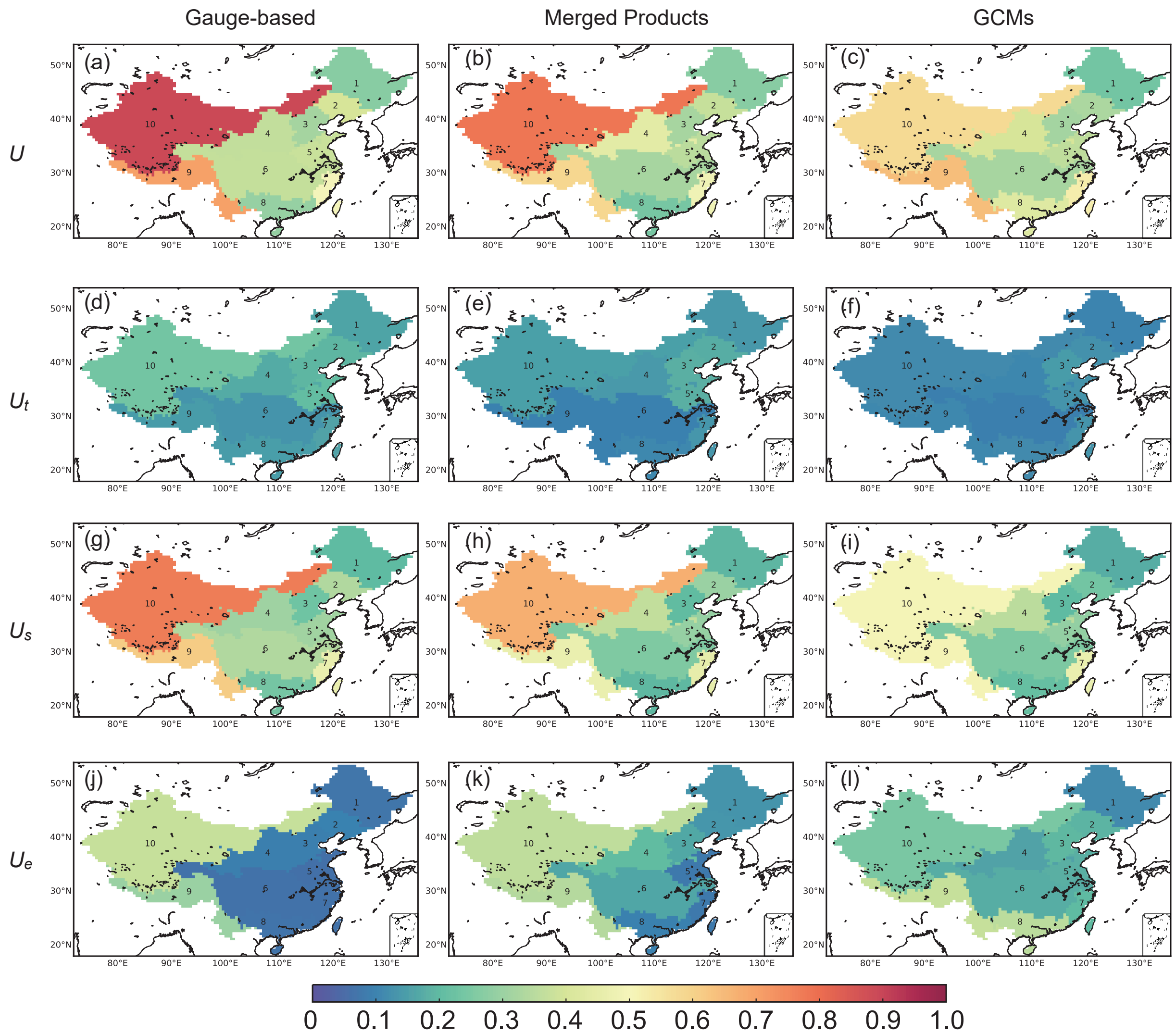

Figure 9. ~~The maps~~ Maps of deviations (U, U_t, U_s, U_e) estimated as the ratio of the square root of the corresponding variances (i.e., V, V_t, V_s, V_e) to the regional mean (μ) ~~for~~ among the ensemble datasets in each of the three different precipitation groups. ~~Among which~~Of these, ~~the~~ U_e is considered ~~as~~ to be the model uncertainty.

that the spatial homogeneity of the variations in ~~the region 10-northwest~~ 10-Northwest China (Figure 5-f) is stronger than that of the other groups (Figure 5-b,d,f). In the GCMs, the highest U_e occurs in ~~the southwestern China~~Southwest China, where both the means and the variations are higher (Figure 4 and 5). ~~In conclusion, the~~ One can conclude that U_e is linked with the magnitude of the model uncertainties in Figure 5 and Figure 6~~. It indicates that the U_e~~, indicating that it is to some degree correlated ~~to~~ with the classic metrics~~as the~~, as higher U_e covers the grid cells or regions with higher model uncertainty.

5 ~~Uncertainty and~~ Comparison of the uncertainty U_e with the classic metrics~~comparison~~

5.1 Deviation from the classic uncertainty metrics

~~The new estimates of the uncertainty U_e provide a comprehensive evaluation of the uncertainty over space and time, and the values are affected by both the temporal variation and the spatial homogeneity (spatial correlations) among the different examined products.~~ In this section, we will compare the uncertainty ~~(U_e)~~ among ensemble members estimated by the three-dimensional partitioning approach with the two classic metrics (defined as $N.s.std$ in Eq. 28 and $N.t.std$ in Eq.29), to explain how these three metrics are related and differ with each other.

~~htbpThe relation of the U_e to two classic metrics as (a) the normalized spatial standard deviation - $N.s.std$ and (b) the normalized temporal standard deviation - $N.t.std$.~~

As shown in Figure 10, U_e is correlated ~~to both the~~ with both $N.s.std$ and $N.t.std$~~, especially~~. The correlation is stronger when U_e is smaller than 0.2, where the regions from 1 to 8 are generally included for all three precipitation groups. ~~The~~ But U_e is in general larger than ~~the~~ $N.s.std$ and $N.t.std$ for the products. ~~And the~~ This deviation is because the ~~variations of the other dimension have~~ variation along one dimension has been collapsed when calculating the ~~spatial deviation (or temporal deviation). For the regions~~ deviation along the other dimension. For subregions 9, 10 and 11, ~~the values of the~~ $N.s.std$ and $N.t.std$ deviate the most from the 1:1 line of ~~the~~ U_e. Taking subregion ~~9-southwest~~ 9-Southwest China in the gauge-based products as an example, the temporal variance is 62.4 $\mathrm{mm\ yr^{-1}}$ while the spatial variance is 571.8 $\mathrm{mm\ yr^{-1}}$ (Figure 7-i). The difference between $N.s.std$ and U_e is 0.058 (=0.297-0.239, ~~changing~~ deviation ratio is 24.3%) when the temporal variation is collapsed~~while the~~. The difference between $N.t.std$ and U_e is 0.126 (=0.297-0.171, ~~changing~~ deviation ratio is 73.4%) when the spatial variation ~~, which is~~ is collapsed. The deviation is significantly larger than ~~the temporal variation, is collapsed~~ that between U_e and $N.s.std$, showing that the collapse will induce a deviation related to the magnitude of the collapsed dimension.

~~These regions~~

These subregions (9, 10, 11) feature strong spatial heterogeneities (Figure 7-i,j,k) in the annual mean precipitation (Figure 4). The ~~spatial correlation of the annual precipitation and the temporal correlation of the regional precipitation is also weaker in these three regions than other regions (*not shown in the results*). The~~ averaging process before estimating ~~classical~~ the classic metrics will cause a significant ~~smooth~~ smoothing of the datasets when the spatial ~~correlation among datasets are very low. The spatial variation across space is also~~ heterogeneity among the datasets is very strong, because the spatial variation is significantly higher than ~~temporal variations (Figure 7). Because the~~ the temporal variation, as shown in Figure 7. The estimation of $N.t.std$~~needs the averaging in spatial dimensionwhich may include~~, which needs an averaging over the spatial

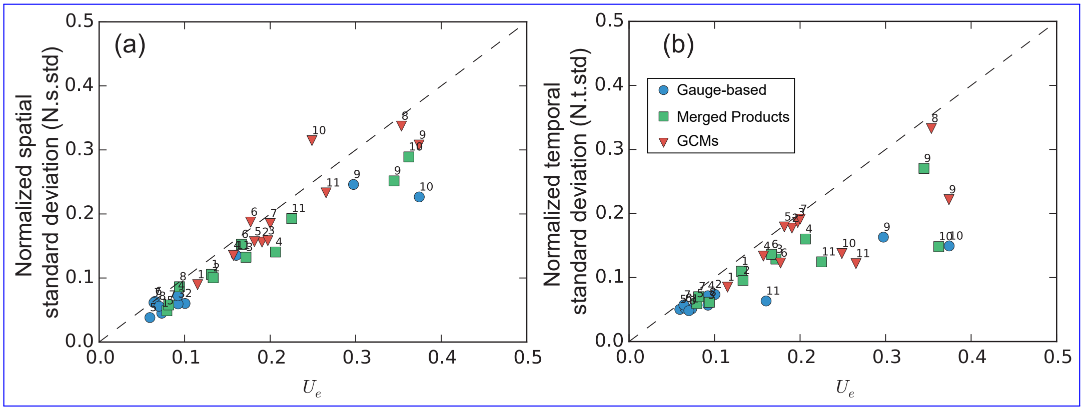

Figure 10. The relation of U_e to the two classic metrics (a) the normalized spatial standard deviation $N.s.std$ and (b) the normalized temporal standard deviation $N.t.std$. The two metrics are estimated with Eqs 28 and 29 among the ensemble datasets in each of the three different precipitation groups.

dimension, will lose more information than that in the time dimension~~, the~~. The deviation between $N.t.std$ and U_e (Figure 10-b) is larger than that between $N.s.std$ and U_e (Figure 10-a). The priority of the precipitation types also changes~~from the~~, from model dominated (the model uncertainty in GCMs are larger than the other) to ~~the~~ region dominated (~~uncertainty in~~ the uncertainties in the specific regions 9, 10, and 11, are larger than in the other regions no matter ~~in~~ which precipitation data is used). This indicates that ~~difference of model uncertainty over space has been~~ the difference in model variation over space can be reflected in the new uncertainty U_e.

Each ~~classical~~ classic metric has its physical ~~meanings as the~~ meaning: $N.s.std$ represents the uncertainties ~~across~~ over space and $N.t.std$ represents the uncertainties across time. The comparison of U_e with each of them demonstrates the metric performance on the same physical meaning. It is possible to compare U_e with a combination of the two ~~classical~~ classic metrics, but the combination ~~can~~ could be far more complex than a simple sum of the two classic metrics. However, ~~the~~ a qualitative comparison is accessible because U_e has a linear correlation with either of them. ~~The correlation will also remain~~ This correlation will persist, and occur between U_e and a combination of the two classic metrics by summing ~~up them~~ them up with certain weights.

5.2 Decomposition of the ensemble uncertainty

We now decompose the ensemble variance to ~~explore the possible~~ determine the reason for the deviation of U_e from ~~the~~ $N.s.std$ and $N.t.std$. As shown in Eq. (26), the ensemble variance ~~(V_e) is formulated as~~ is expressed by

$$V_e = \frac{1}{3}\left[\frac{\overline{\sigma^2_{e\_t}} + \overline{\sigma^2_{e\_s}}}{2} + \overline{\sigma^2_e} + \sigma^2_e(\mu_{ts})\right]. \tag{30}$$

~~It~~ This combines four components which stand for the variation of different estimates across the ensemble dimension (i.e., the variance of original temporal and spatial values - $\overline{\sigma^2_e}$, of the temporal mean - $\overline{\sigma^2_{e\_t}}$, of the spatial mean - $\overline{\sigma^2_{e\_s}}$ and of the grand mean - $\sigma_e^2(\mu_{ts})$). Among ~~which, the~~ these, $\overline{\sigma^2_{e\_t}}$ is the mean of the ~~square of spatial standard~~ squares of the spatial deviation in Figure 5-a,c,e for all grids in a specific region and $\overline{\sigma^2_{e\_s}}$ is the mean of the ~~square~~ squares of the temporal ~~standard~~ deviation in Figure 6 for each time step in a specific region. These two components are closely related to the two classic metrics $N.s.std$ (Eq. 28) and $N.t.std$ (Eq. 29), respectively.

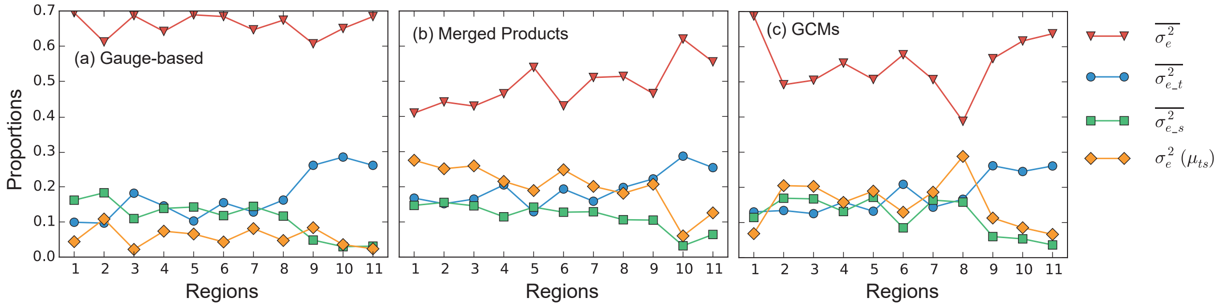

Figure 11. The ~~proportion~~ proportions of the four components in Eq. (30) to ~~the~~ V_e among the ensemble datasets in each of the three different precipitation groups~~,~~: (a) gauge-based products, (b) merged products and (c) GCMs. The contribution is normalized so that ~~the~~ their sum ~~of them~~ is 1.0 for each region. Among the four components, ~~the~~ $\overline{\sigma^2_{e\_t}}$ and $\overline{\sigma^2_{e\_s}}$ are associated with the two classic ~~metric~~ metrics $N.s.std$ and $N.t.std$, respectively.

By decomposing ~~the~~ Eq. (30), the contributions of the four components to the ensemble variance ~~(V_e)~~ are shown in Figure 11. For all three precipitation groups, $\overline{\sigma_e^2}$ is the dominant component simply because all the information on variations among the original datasets is retained in the uncertainty estimation. ~~While, the~~ The other three components ~~are estimations after~~ result from estimations after an averaging is performed~~in~~, either over time, space, or the full spatiotemporal dimensions, which ~~indicates~~ means a loss of information. The contribution of ~~the~~ $\overline{\sigma^2_{e\_t}}$ and $\overline{\sigma^2_{e\_s}}$ is approximating 0.15 for regions from 1 to 8. ~~While the~~ But $\overline{\sigma^2_{e\_t}}$ increases for ~~the region~~ regions 9, 10 and 11, indicating that ~~the spatial heterogeneity is significant for~~ there is significant spatial heterogeneity in these regions. ~~On the contrary~~In contrast, $\overline{\sigma^2_{e\_s}}$ decreases because the spatial averaging has collapsed the spatial variations. The very small contribution of $\overline{\sigma^2_{e\_s}}$ related to $N.t.std$ is the cause for larger deviations between $N.t.std$ and U_e in these subregions (Figure 10-b).

Although ~~all the components~~ any component can be used as ~~metrics~~ a metric for evaluating the variations among multiple datasets, there are limitations for each of the variations. For the variation of the temporal mean $\overline{\sigma^2_{e\_t}}$ and spatial mean $\overline{\sigma^2_{e\_s}}$, the collapse of a dimension has ignored part of the information~~(also introduced in the Introduction)~~. Moreover, the variation of the grand mean $\sigma_e^2(\mu_{ts})$ has ignored both the temporal variability and spatial heterogeneity, which further decreases its applicability ~~in uncertainty assessment~~ to the assessment of uncertainty. The variation $\overline{\sigma_e^2}$ is estimated based on the original data without averaging, and thus it represents the most information. However, it ~~cannot account for~~ does not take into account the systematic uncertainty (bias in the mean values) which is expressed ~~as~~ by $\sigma_e^2(\mu_{ts})$.

Therefore, ~~all the four components represent the model variations from different aspects and neither~~ none of the single ~~component~~ components is able to represent ~~all~~ the others. ~~Integration of different components (~~The integrated metric V_e ~~)~~ is therefore a solution ~~to indicate~~ that represents all metrics to different degrees. What is interesting is that the variability of the proportions of $\overline{\sigma^2_{e\_t}}$ and $\overline{\sigma^2_{e\_s}}$ (or $\overline{\sigma^2_e}$ and $\sigma^2_e(\mu_{ts})$) are opposite and the sum of their proportions is stable, around 0.3 (or 0.7). This indicates a complementary relation between the two pairs of elements ($\overline{\sigma^2_{e\_t}}$ & $\overline{\sigma^2_{e\_s}}$; $\overline{\sigma^2_e}$ & $\sigma^2_e(\mu_{ts})$). On the other ~~word~~hand, some of the information is ignored in one of the components but ~~remained~~ remains in the other one within the same pair. ~~And therefore, it indicates that the variation in~~ Therefore, the variation along the time dimension and that ~~in the space~~ along the spatial dimension should be considered together~~as~~, as is done in the estimation of the ensemble variance ~~(~~V_e~~)~~. The normalized ~~metric (~~uncertainty U_e ~~)~~ derived from the integrated variation ~~(V_e), which has better ability to demonstrate the uncertainties compared to~~ , which is better able to determine the uncertainties than are the classic metrics, should be ~~a properer choice for the~~ the more proper choice for an uncertainty analysis.

5.3 ~~Metrics differences~~ Differences between the metrics in value and proportion

Figure 10 shows that ~~the~~ U_e is generally higher than the uncertainty identified by the two classic metrics ~~.~~ (i.e., $N.s.std$ and $N.t.std$). Figure 12 then ~~summaries~~ summarizes the magnitude of the ~~changes from~~ deviations of the classic metrics ~~to~~ from the new uncertainty ~~identified by~~ U_e. We can ~~find~~ see that the two classic metrics generally underestimate the uncertainty by around 0.03 (Figure 12-a). The variation of the underestimation of $N.t.std$ is larger than that of ~~the~~ $N.s.std$, showing a larger deviation between ~~the~~ U_e ~~with~~ and $N.t.std$. ~~Applying~~ Employing the new uncertainty metric will increase the ~~estimation of~~ estimated uncertainty by around ~~20-40%~~20%–40% for half of the cases~~compared to the~~, when compared to $N.s.std$ (Figure 12-b). For nearly 25% of the cases, the new U_e increases the ~~estimation of~~ estimated uncertainty by more than 50%. In ~~the~~ extreme cases, U_e is ~~larger than twice~~ more than double $N.t.std$ (Figure 12-b). The results show that the ~~known uncertainty estimated by~~ widely applied uncertainty estimates from the two classic metrics ~~, which have been widely applied to climatic analysis, have~~ have underestimated the uncertainty among different models / datasets. ~~The~~ Such an underestimation may especially occur for ~~assessment of temporal evolution of~~ the temporal assessment of the uncertainties ($N.t.std$), which is very commonly seen in scientific reports and articles to illustrate the temporal evolution of the variables of interest.

6 Discussion and ~~Conclusion~~Conclusions

6.1 Features and applicability of the approach

The total variation of the database which consists of multiple datasets is contributed by the ~~spatio-temporal~~ spatiotemporal variations as well as the uncertainties among ~~the~~ ensemble datasets. While the uncertainty assessment with current approaches (e.g., eqs. 28 and 29) needs either the temporal ~~variation~~ variability or the spatial heterogeneity to be averaged which means a loss of information~~and bias in uncertainty estimation. The proposed~~. The variance partitioning approach proposed in this study works in three dimensions. It uses all the information ~~across the time and the space~~ over both the temporal and the spatial dimensions among the multiple ensemble members~~, thus it~~. It avoids the collapse of variation along any dimension, and thus the proposed uncertainty estimate U_e provides a more accurate ~~uncertainty estimation. The proposed~~ estimate of the

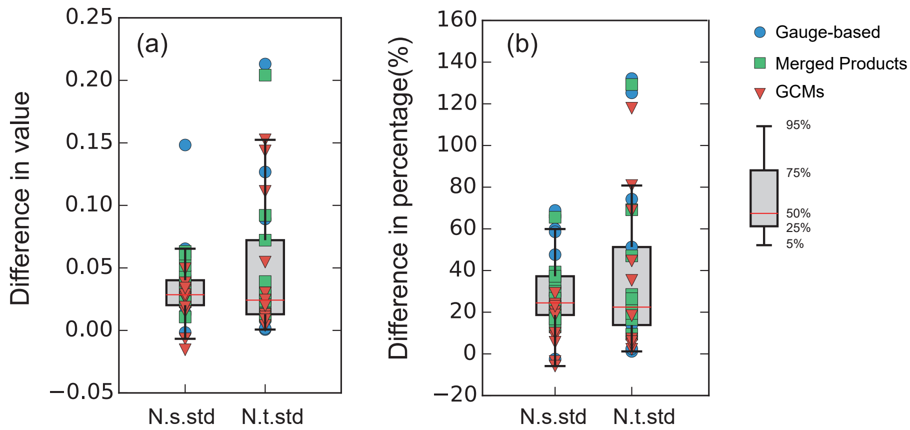

Figure 12. The changes in (a) value and (b) percentage when using U_e as the new uncertainty metric compared to classic metrics $N.s.std$ (Eq. 28) and $N.t.std$ (Eq. 29).

uncertainty. The estimate U_e is especially suitable for ~~the~~ an overall assessment among multiple datasets over a certain period and over a specific space. ~~Though, the compensation is that the~~ Even though the trade-off is that U_e cannot provide the temporal evolution or spatial heterogeneity for users' consideration~~. In~~ , in many cases we would like to know the general performance of the ensemble models ~~with a~~ based on a global single estimate.

The results of ~~the~~ this partitioning approach can be affected by the choice of the time step intervals. For example, the ~~time variation or time variance proportion~~ temporal variance or proportion of temporal variance will significantly increase if the time interval is chosen ~~as~~ to be one month. The inter-annual ~~variation~~ variability of precipitation will result in higher V_t ~~and lower V_s or V_e. It depends~~ . The changes depend on how significant the inter-annual variability is compared to the intra-annual variations. Moreover, only changes in the temporal variation (~~increase or reduce the variation magnitude while remain the average values) will be captured in the~~ the average values remain but the magnitudes of the variation increase or decrease) can be captured by U_e ~~while~~ . But $N.s.std$ will ~~keep~~ remain the same because the temporal ~~variation~~ variability has been neglected in the averaging process. ~~The case will be the same for~~ It is the same with $N.t.std$ if different spatial resolutions of the measurements are used.

The proposed approach has a flexible structure that ~~potentially deals~~ can deal with different problems, from a global scale to regional studies. The temporal dimension can also span from ~~global to regional dimensions. The time dimension can consider intervals from~~ daily, monthly, annual ~~or to decadal analysis in~~ to decadal analyses with different scopes. The ensemble dimension is applicable from ~~2~~ two members (i.e., model evaluation between simulations and observations) to any number of multi-models (consensus evaluation, Tebaldi et al., 2011; McSweeney and Jones, 2013). The present approach is also applicable to any variables that are organized in ~~the three dimensions~~three dimensions, such as climatic variables (e.g., temperature, evaporation), hydrological variables (e.g., soil moisture, runoff) or environmental variables (e.g., drought index). Based on these advantages, ~~the~~ this three-dimensional partitioning approach can ~~widely be applied in the~~ be widely applied in hydro-climatic analysis.

6.2 ~~Conclusion~~Conclusions

A new three-dimensional partitioning approach ~~is~~ has been proposed in this ~~study~~ paper to assess the model uncertainties among multiple ensemble datasets. The new uncertainty metric (U_e) is estimated with an overall consideration of temporal and spatial variations as well as the differences among the ensemble products. ~~Results show~~ The results have shown that U_e is generally larger than the ~~classical~~ classic uncertainty metrics $N.s.std$ and $N.t.std$, which require a collapse ~~in either of the time or space~~ of the variation along either the temporal or spatial dimension. The deviation occurs where the spatial variations are significant but being averaged in the $N.t.std$ estimation. The ~~decomposing of the~~ decomposition of the total variance V_e shows the complementary relation ~~of~~ between the two classic metrics, and therefore the new uncertainty U_e (derived from V_e) is a more comprehensive ~~estimation of uncertainty~~estimate of the uncertainty among multiple ensemble products.

Thirteen precipitation datasets generated by different ~~methodologies are~~ methods have been categorized into three groups (~~i.e.~~namely, gauge-based products, merged products and GCMs) and the model uncertainty in the ensemble products ~~in the same group is~~ has been analyzed with the new ~~and~~ approach and with the two classic uncertainty metrics ~~. The GCMs are identified with the largest model uncertainty with the classical metrics~~for each precipitation group. Using the classic metrics, in most regions, ~~while the new estimation~~ the GCMs have been indicated as having the largest model uncertainty. But the new estimator U_e indicates that the largest model uncertainty occurs in specific regions no matter ~~in~~ which precipitation group ~~. The spatial heterogeneity of~~ is considered. The impact of spatial heterogeneity on the model uncertainty ~~over space~~ has been represented well in the new uncertainty metric ~~. Thus~~(U_e). In addition to the theoretical analysis of the components of U_e, the overall model uncertainty ~~(U_e) is~~ can be used as a new uncertainty estimate which involves more information and should receive more attention in the ~~uncertainty assessment field~~field of uncertainty assessment.

Appendix A: The algorithms for different expressions in the methodology

Zone A:

A1: ~~$\mu_t[s,e;n\times l];\mu_t[j,k]=\frac{1}{m}\sum_{i=1}^{l} z_{ijk}$~~$\mu_t[s,e;n\times l];\mu_t[j,k]=\frac{1}{m}\sum_{i=1}^{m} z_{ijk}$

A2: ~~$\mu_s[e,t;l\times m];\mu_s[k,i]=\frac{1}{n}\sum_{j=1}^{l} z_{ijk}$~~$\mu_s[e,t;l\times m];\mu_s[k,i]=\frac{1}{n}\sum_{j=1}^{n} z_{ijk}$

A3: $\mu_e[t,s;m\times n];\mu_e[i,j]=\frac{1}{l}\sum_{k=1}^{l} z_{ijk}$

Zone B:

B1: ~~$\sigma_t^2[s,e;n\times l];\sigma_t^2[j,k]=\frac{1}{m}\sum_{i=1}^{l}(z_{ijk}-\mu_t[j,k])^2$~~$\sigma_t^2[s,e;n\times l];\sigma_t^2[j,k]=\frac{1}{m}\sum_{i=1}^{m}(z_{ijk}-\mu_t[j,k])^2$

B2: ~~$\sigma_s^2[e,t;l\times m];\sigma_s^2[k,i]=\frac{1}{n}\sum_{j=1}^{l}(z_{ijk}-\mu_s[k,i])^2$~~$\sigma_s^2[e,t;l\times m];\sigma_s^2[k,i]=\frac{1}{n}\sum_{j=1}^{n}(z_{ijk}-\mu_s[k,i])^2$

B3: $\sigma_e^2[t,s;m\times n];\sigma_e^2[i,j]=\frac{1}{l}\sum_{k=1}^{l}(z_{ijk}-\mu_e[i,j])^2$

Zone C:

C1: $\sigma_{t\_s}^2[e;l];\sigma_{t\_s}^2[k]=\sigma^2(\mu_s[k,:])$

C2: $\sigma_{t\_e}^2[s;n];\sigma_{t\_e}^2[j]=\sigma^2(\mu_e[:,j])$

C3: $\sigma_{s\_t}^2[e;l];\sigma_{s\_t}^2[k]=\sigma^2(\mu_t[:,k])$

C4: $\sigma_{s\_e}^2[t;m];\sigma_{s\_e}^2[i]=\sigma^2(\mu_e[i,:])$

C5: $\sigma^2_{e\_t}[s;n]; \sigma^2_{e\_t}[j] = \sigma^2(\mu_t[j,:])$

C6: $\sigma^2_{e\_s}[t;m]; \sigma^2_{e\_s}[i] = \sigma^2(\mu_s[:,i])$

Zone D:

D1: $\mu_{et}[s;n]; \mu_{et}[j] = \frac{1}{lm}\sum_{k=1}^{l}\sum_{i=1}^{m} z_{ijk}$

D2: $\mu_{se}[t;m]; \mu_{se}[i] = \frac{1}{nl}\sum_{j=1}^{n}\sum_{k=1}^{l} z_{ijk}$

D3: $\mu_{ts}[e;k]; \mu_{ts}[k] = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} z_{ijk}$

Zone E:

E1: $\sigma^2_{et}[s;n]; \sigma^2_{et}[j] = \frac{1}{lm}\sum_{k=1}^{l}\sum_{i=1}^{m}(z_{ijk} - \mu_{et}[j])^2$

E2: $\sigma^2_{se}[t;m]; \sigma^2_{se}[i] = \frac{1}{nl}\sum_{j=1}^{n}\sum_{k=1}^{l}(z_{ijk} - \mu_{se}[i])^2$

E3: $\sigma^2_{ts}[e;l]; \sigma^2_{st}[k] = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(z_{ijk} - \mu_{ts}[k])^2$

Zone F:

F1: $\sigma^2_{t}(\mu_{se}) = \frac{1}{m}\sum_{i=1}^{m}(\frac{1}{nl}\sum_{j=1}^{n}\sum_{k=1}^{l} z_{ijk} - \frac{1}{m}\sum_{i=1}^{m}(\frac{1}{nl}\sum_{j=1}^{n}\sum_{k=1}^{l} z_{ijk}))^2$

F2: $\sigma^2_{s}(\mu_{et}) = \frac{1}{n}\sum_{j=1}^{n}(\frac{1}{lm}\sum_{k=1}^{l}\sum_{i=1}^{m} z_{ijk} - \frac{1}{n}\sum_{j=1}^{n}(\frac{1}{lm}\sum_{k=1}^{l}\sum_{i=1}^{m} z_{ijk}))^2$

F3: $\sigma^2_{e}(\mu_{ts}) = \frac{1}{l}\sum_{k=1}^{l}(\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} z_{ijk} - \frac{1}{l}\sum_{k=1}^{l}(\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} z_{ijk}))^2$

The t, s, e in the algorithms represents the three dimensions **time**, **space** and **ensemble**, with the size of m, n, l and index with i, j, k, respectively. Each expression is shown with its size and the meaning of each dimension. For example, for the A1: $\mu_t[s,e;n \times l]$, the μ_t has a size of $n \times l$. The first axis represents the space dimension, and the second is the ensemble dimension. While C1 ($\sigma^2_{t\_s}[e;l]$) has only one ensemble dimension with its size as l. F1 ($\sigma^2_t(\mu_{se})$) is only a single value.

Author contributions. XZ initialized the ideas presented in this paper with supervising from JP and TY. XZ prepared the simulations, the figures and the manuscript. CSH participated in the data preparation. All authors contributed to the discussion and revising the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This study was supported by the National Natural Science Foundation of China (grant nos. 41561134016 and 51879068); the CHINA-TREND-STREAM French national project (ANR grant no. ANR-15-CE01-00L1-0L); the National Key Research and Development Program (2018YFC0407900) and the China Scholarship Council (CSC, 201506710042). The work was supported by computing resource of the IPSL ClimServ cluster at École Polytechnique, France.

References

Adler, R. F., Sapiano, M. R., Huffman, G. J., Wang, J. J., Gu, G., Bolvin, D., Chiu, L., Schneider, U., Becker, A., Nelkin, E., Xie, P., Ferraro, R., and Shin, D. B.: The Global Precipitation Climatology Project (GPCP) monthly analysis (New Version 2.3) and a review of 2017 global precipitation, Atmosphere, 9, https://doi.org/10.3390/atmos9040138, https://www.esrl.noaa.gov/psd/data/gridded/data.gpcp.html, 2018.

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., Van Dijk, A. I., Weedon, G. P., Brocca, L., Pappenberger, F., Huffman, G. J., and Wood, E. F.: Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling, Hydrology and Earth System Sciences, 21, 6201–6217, https://doi.org/10.5194/hess-21-6201-2017, http://gloh2o.org, 2017.

Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., Schär, C., and Schar, C.: Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections: UNCERTAINTY SOURCES IN CLIMATE-IMPACT PROJECTIONS, Water Resources Research, 49, 1523–1536, https://doi.org/10.1029/2011WR011533, 2013.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., Mcnally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, https://www.ecmwf.int/en/forecasts/datasets/archive-datasets/reanalysis-datasets/era-interim, 2011.

Déqué, M., Rowell, D. P., Lüthi, D., Giorgi, F., Christensen, J. H., Rockel, B., Jacob, D., Kjellström, E., De Castro, M., and Van Den Hurk, B.: An intercomparison of regional climate simulations for Europe: Assessing uncertainties in model projections, Climatic Change, 81, 53–70, https://doi.org/10.1007/s10584-006-9228-x, 2007.

Everitt, B.: The Cambridge Dictionary of Statistics, vol. 53, Cambridge University Press, https://doi.org/10.1017/CBO9781107415324.004, 2013.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset, International Journal of Climatology, 34, 623–642, https://doi.org/10.1002/joc.3711, 2014.

IPCC: Technical Sum- mary. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013a.

IPCC: Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013b.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S. C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J. F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L., and Vertenstein, M.: The community earth system model (CESM) large ensemble project : A community resource for studying climate change in the presence of internal climate variability, Bulletin of the American Meteorological Society, 96, 1333–1349, https://doi.org/10.1175/BAMS-D-13-00255.1, 2015.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: World map of the Köppen-Geiger climate classification updated, Meteorologische Zeitschrift, 15, 259–263, https://doi.org/10.1127/0941-2948/2006/0130, 2006.

Landerer, F. W. and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water Resources Research, 48, 1–11, https://doi.org/10.1029/2011WR011453, 2012.

McSweeney, C. F. and Jones, R. G.: No consensus on consensus: The challenge of finding a universal approach to measuring and mapping ensemble consistency in GCM projections, Clim. Change, 119, 617–629, https://doi.org/10.1007/s10584-013-0781-9, 2013.

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., and Houston, T. G.: An overview of the global historical climatology network-daily database, Journal of Atmospheric and Oceanic Technology, 29, 897–910, https://doi.org/10.1175/JTECH-D-11-00103.1, 2012.

Phillips, T. J. and Gleckler, P. J.: Evaluation of continental precipitation in 20th century climate simulations: The utility of multimodel statistics, Water Resources Research, 42, 1–10, https://doi.org/10.1029/2005WR004313, 2006.

Rodell, M., Houser, P., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radarkovich, J., Bosilovich, M., Entin, J., Walker, J., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, American Meteorological Society, pp. 381–394, https://doi.org/10.1175/BAMS-85-3-381, 2004.

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N. W., Clark, D. B., Dankers, R., Eisner, S., Fekete, B. M., Colón-González, F. J., Gosling, S. N., Kim, H., Liu, X., Masaki, Y., Portmann, F. T., Satoh, Y., Stacke, T., Tang, Q., Wada, Y., Wisser, D., Albrecht, T., Frieler, K., Piontek, F., Warszawski, L., and Kabat, P.: Multimodel assessment of water scarcity under climate change, Proceedings of the National Academy of Sciences, 111, 3245–3250, https://doi.org/10.1073/pnas.1222460110, 2014.

Schneider, A., Jost, A., Coulon, C., Silvestre, M., Théry, S., and Ducharne, A.: Global scale river network extraction based on high-resolution topography and constrained by lithology, climate, slope, and observed drainage density, Geophysical Research Letters, pp. 2773–2781, https://doi.org/10.1002/2016GL071844, 2017.

Sun, F., Roderick, M. L., Farquhar, G. D., Lim, W. H., Zhang, Y., Bennett, N., and Roxburgh, S. H.: Partitioning the variance between space and time, Geophysical Research Letters, 37, 1–6, https://doi.org/10.1029/2010GL043323, 2010.

Sun, F., Roderick, M. L., and Farquhar, G. D.: Changes in the variability of global land precipitation, Geophysical Research Letters, 39, 1–6, https://doi.org/10.1029/2012GL053369, 2012.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K. L.: A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons, Reviews of Geophysics, 56, 79–107, https://doi.org/10.1002/2017RG000574, 2018.

Tapiador, F. J., Turk, F. J., Petersen, W., Hou, A. Y., Garcia-Ortega, E., Machado, L. A. T., Angelis, C. F., Salio, P., Kidd, C., Huffman, G. J., and de Castro, M.: Global precipitation measurement: Methods, datasets and applications, Atmospheric Research, 104-105, 70–97, https://doi.org/10.1016/j.atmosres.2011.10.021, 2012.

Tebaldi, C., Arblaster, J. M., and Knutti, R.: Mapping model agreement on future climate projections, Geophys. Res. Lett., 38, 1–5, https://doi.org/10.1029/2011GL049863, 2011.

Willmott, C. J. and Matsuura, K.: Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1900 - 2010), http://climate.udel.edu/data, 2012.

Xie, P., Janowiak, J. E., Arkin, P. A., Adler, R., Gruber, A., Ferraro, R., Huffman, G. J., and Curtis, S.: GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates, Journal of Climate, 16, 2197–2214, https://doi.org/10.1175/2769.1, https://www.esrl.noaa.gov/psd/data/gridded/data.cmap.html, 2003.

Xie, P., Chen, M., Yang, S., Yatagai, A., Hayasaka, T., Fukushima, Y., and Liu, C.: A Gauge-Based Analysis of Daily Precipitation over East Asia, Journal of Hydrometeorology, 8, 607–626, https://doi.org/10.1175/JHM583.1, http://ftp.cpc.ncep.noaa.gov/precip/CPC_UNI_PRCP/, 2007.

Certificate of proofreading service & highlighted changes

Proof-Reading-Service.com
PhD theses, journal papers, books and other professional documents

Proof-Reading-Service.com Ltd, Devonshire Business Centre, Works Road, Letchworth Garden City, Hertfordshire, SG6 1GJ, United Kingdom
Office phone: +44(0)20 31 500 431
E-mail: enquiries@proof-reading-service.com
Internet: http://www.proof-reading-service.com
VAT registration number: 911 4788 21
Company registration number: 8391405

10 January 2020

To whom it may concern,

RE: Proof-Reading-Service.com Editorial Certification

This is to confirm that the document described below has been submitted to Proof-Reading-Service.com for editing and proofreading.

We certify that the editor has corrected the document, ensured consistency of the spelling, grammar and punctuation, and checked the format of the sub-headings, bibliographical references, tables, figures etc. The editor has further checked that the document is formatted according to the style guide supplied by the author. If no style guide was supplied, the editor has corrected the references in accordance with the style that appeared to be prevalent in the document and imposed internal consistency, at least, on the format.

It is up to the author to accept, reject or respond to any changes, corrections, suggestions and recommendations made by the editor. This often involves the need to add or complete bibliographical references and respond to any comments made by the editor, in particular regarding clarification of the text or the need for further information or explanation.

We are one of the largest proofreading and editing services worldwide for research documents, covering all academic areas including Engineering, Medicine, Physical and Biological Sciences, Social Sciences, Economics, Law, Management and the Humanities. All our editors are native English speakers and educated at least to Master's degree level (many hold a PhD) with extensive university and scientific editorial experience.

Document title:	**A new uncertainty estimation with:multiple datasets and implementation for various precipitation products**
Author(s):	**Xudong Zhou; Jan Polcher; Tao Yang; Ching-Sheng Huang**
Format:	**British English**
Style guide:	**Not supplied**

A new uncertainty estimation ~~among~~ with multiple datasets and implementation ~~to~~ for various precipitation products

Xudong Zhou[1,2,3], Jan Polcher[2], Tao Yang[1], and Ching-Sheng Huang[1]

[1]State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Center for Global Change and Water Cycle, Hohai University, Nanjing 210098, China

[2]Laboratoire Météorologie Dynamique du CNRS, IPSL, CNRS, Paris, F 91128, France

[3]Institute of Industrial Science, The University of Tokyo, Tokyo, 〒153-8505, Japan

Abstract.

Ensemble estimates based on multiple datasets are frequently applied once many datasets are available for the same climatic variable. Uncertainty that evaluates the difference ~~among~~ between the ensemble datasets is always provided along with the ensemble mean estimates to show ~~how~~ to what extent the ensemble members are consistent with each other. However, one fundamental flaw of classic uncertainty estimates is that only the uncertainty in one dimension (either the temporal variability or the spatial heterogeneity) can be considered~~while the variation in~~ , whereas the variation along the other dimension is dismissed due to limitations in algorithms for classic uncertainty estimates, resulting in an incomplete assessment of the uncertainties. This study introduces a three-dimensional variance partitioning approach and proposes a new uncertainty estimation (U_e) ~~with integration of~~ that includes the data uncertainties in both spatiotemporal scales. The new methods ~~avoids~~ avoid pre-averaging in either of the spatiotemporal dimensions and as a result, the U_e estimate is around 20% higher than the classic uncertainty metrics. The deviation of U_e from the classic metrics is apparent for regions with strong spatial heterogeneity and where the variations significantly differ in temporal and spatial scales. ~~It demonstrates~~ This shows that classic metrics ~~will~~ reduce the uncertainty estimate through averaging, which means a loss of information in the variations across spatiotemporal scales. Decomposing ~~of U_e formula proves~~ the formula for U_e shows that U_e has integrated four different variations across the ensemble dataset members, while only two of the components are represented in the classic uncertainty estimates. ~~The decomposing analysis~~ This analysis of the decomposition explains the correlation as well as the differences between the newly proposed U_e and the two classic uncertainty metrics. The new approach is implemented and analyzed with multiple precipitation products of different types (e.g., gauge-based products, merged products and GCMs) which contain different sources of uncertainties with different magnitudes. ~~U_e among~~ Among the multiple gauge-based precipitation products, U_e is the smallest~~while U_e~~ , while among other products U_e is generally larger because other uncertainty sources are included and the ~~constrain of observations is~~ constraints of the observations are not as strong as ~~that~~ in gauge-based products. This new three-dimensional approach is flexible in its structure and particularly suitable for a comprehensive assessment of multiple datasets over large regions within any given period.

Copyright statement.

1 Introduction

With the technical ~~development for monitoring the~~ developments in monitoring natural climate variables and the increasing knowledge of the physical mechanisms in the climate system, many institutes have the ability to provide different kinds of climate datasets. ~~Taken the~~ Taking precipitation, which is the dominant variable in the land water cycle, as an example, there are point measurements, such as GHCN-D (global historical climatology network-daily, **?**), gridded products based on gauge measurements and interpolation (e.g., CRU, **?**), products derived from remote sensing (e.g., the Tropical Rainfall Measuring Mission - TRMM), reanalysis datasets (e.g., NCEP) and ~~those~~ estimates from models (e.g., GCMs). These products ~~are~~ have been developed using different original data, technologies ~~or~~ and model settings for various purposes (**????**). As a result, ~~differences exist among~~ there are differences between the various products due to ~~the~~ measurement errors, model biases~~or chaotic noises~~, or chaotic noise. The uncertainty is thus regarded as the deviation of these model results from their real values.

However, the real values are difficult to measure and the uncertainties are difficult ~~to be removed~~ remove from the datasets. Thus, using ensembles consisting of multiple datasets to generate a weighted average ~~becomes~~ has become very popular in climate-related ~~researches~~research. The ensemble means of multiple datasets are considered ~~as~~ more reliable estimates than a single dataset. For example, IPCC uses 42 CMIP5 (Coupled Model Intercomparison Project Phase 5) models to show historical temperature changes and 39 CMIP5 models to average future temperature projections in an RCP 8.5 scenario (Figure SPM.7 in **?**). **?** use nine global hydrological models to evaluate ~~the~~ global water scarcity under climate change. GLDAS (Global Land Data Assimilation System) involves four different land surface models (**?**) and GRACE (Gravity Recovery and Climate Experiment) provides ~~estimations~~ estimates from three independent institutes (**?**). Using multiple datasets reduces the dependence on a single dataset and eliminates the random variations associated to biases or ~~noises~~ noise in each single model ~~estimates~~estimate.

Along with the ensemble means, uncertainty information is recommended to be presented because the ~~uncertainty level decides~~ level of uncertainty determines the reliability of the ensemble results. In general, uncertainties can be quantified as the range of maximum and minimum values (i.e., $V_{max} - V_{min}$), the value difference at different quantiles (e.g., $V_{5\%} - V_{95\%}$), the consistency of models (ratio of models following a certain pattern to the total number of models), the variation (σ^2) or the standard deviation (σ) among multiple model estimations. These metrics describe the differences ~~of multiple model estimations from~~ between multiple model estimates in different aspects. Among the metrics, the standard deviation (σ) is the most used because it has the same magnitude **Author query: Perhaps you mean 'units', not 'magnitude'. Anyway, it is not clear what one would mean by 'magnitude' of a dataset: it it meant anything, it would mean the number of samples, which is obviously irrelevant here.** as the original dataset. Moreover, it is less sensitive to ~~the~~ extreme samples and to the number of datasets used for the investigation. The ratio of the standard deviation (σ) to the mean value (μ), the so-called coefficient of variance (CV), representing the dispersion or spread of the distribution of various ensemble members (**?**), is a ~~unit-less~~ unitless value which also shows the degree of uncertainty efficiently.

Depending on the purpose of the data evaluation, the uncertainty ~~among~~ between the datasets can be displayed or visualized ~~over~~ in space to show the spatial heterogeneity. For example, the predicted future temperature increase has a higher significance in the northern high-latitudes among different models than in the middle-latitudes (Box TS.6 Figure 1 in **?**). ~~The other~~ Another typical implementation is to evaluate the ~~uncertainty evolution along the temporal scale~~evolution of the uncertainty over time. In general, the ~~uncertainty range~~ range of the uncertainty decreases in the historical period over time because more observations ~~are~~ have been accessible recently. ~~While,~~ But the uncertainty increases in future projections because of the increasing spread of model ~~estimations~~ estimates (Figure SPM.7 in **?**), indicating a decreasing of consistency but increasing variation among various datasets.

The two kinds of ways can easily show the spatial distribution or the temporal evolution of the uncertainty. But ~~the~~ a short-coming is apparent, as the variation ~~of~~ along one dimension (time or space) has to be collapsed to generate the mean values ~~)~~ when we attempt to assess the uncertainty for the other dimension (space or time). For example, the averaging over a specific region to obtain the spatial mean is estimated at each time step before obtaining the temporal evolution of the model uncertainty (red flowcharts in Figure 1). ~~On the contrary, the~~ In contrast, averaging over a certain ~~period to~~ temporal period to obtain the temporal mean is necessary ~~at~~ for each grid cell when estimating the spatial variations of model uncertainties (blue flowcharts in Figure 1). ~~Though, the averaging~~The averaging, in either dimension, means a loss of ~~the information in data variation~~information about the variation **Author query: Perhaps ‘variability’ would be a better word than ‘variation’.** in the data. Any changes in the variation ~~but with the mean value to be remained~~ that leaves the mean values unchanged will not be propagated to the global uncertainty estimation. ~~This results in~~ The result of this is that the variations ~~among datasets not being~~ between datasets is not fully considered when estimating the uncertainties. In other words, neither of the uncertainty estimates can represent the ~~full differences among~~ whole of the differences between multiple datasets. The uncertainty can be underestimated~~and their similarity be overestimated. However, current studies have~~, and the similarity of the datasets thus overestimated. Indeed, the current literature has not paid attention to the ~~ignorance~~ ignoring of variation after averaging as well as its influence on the ~~uncertainty assessment~~ assessment of the uncertainty.

Figure 1. The two classic uncertainty assessments in the current ~~researches as~~ literature: the temporal evolution of the model uncertainty (red) and the spatial distribution of the model uncertainty (blue). ~~Either~~ Each of these estimations of the uncertainty ~~estimates~~ has ~~to do the averaging in~~ average over one of the dimensions~~in~~, either space or time, ~~and it~~ which will lead to ~~the loss of~~ losing information ~~in~~ about the corresponding dimension.

The total variation among multiple datasets is contributed to by the spatial heterogeneity, temporal variability and the model uncertainties. To some degree, the model uncertainty is similar to other dimensions as a variation ~~in~~ along a third dimension (ensemble dimension). The key to ~~evaluate~~ evaluating the model uncertainty is to decompose the variation caused by ~~dataset differences~~ differences between the datasets from the other two contributors. ~~Though variation decomposition with method~~ Although decomposing the variation by means of ANalysis Of VAriance (ANOVA) is often seen in hydro-metrological studies,

~~it~~ this is designed to separate the process uncertainties generated in different model processes that ~~propagated~~ propagate to the final variation. For example, **?** ~~separated~~ decomposed the uncertainties of Regional Climate Models (RCM) ~~to~~ into four sources of ~~uncertainties (~~uncertainty: sampling uncertainty, model uncertainty, radiative uncertainty and boundary uncertainty~~)~~. **?** decomposed the uncertainty in river streamflow projections to uncertainties from climate models, statistical post-processing schemes and hydrological models. These implementations differ from the ~~scope~~ purpose of the present study because they fail to separate the uncertainties from the spatiotemporal variations because spatiotemporal averaging ~~has been~~ was already applied in the estimation process. **??** ~~in~~ for the first time decomposed the total variation ~~to~~ into temporal variation and spatial heterogeneity. They concluded that the variations ~~in space~~ along the spatial **Author query: If you use 'space' as an adjective, it should be 'spatial'. Same for 'time', whose adjective is 'temporal'. Thus 'spatial dimension' is more correct than 'space dimension' since the latter is a kind of Germanic piling up of nouns, whereas to modify 'dimension' you need an adjective, not another noun. So you need 'temporal' or 'spatial'. It is true that in English one sometimes uses a noun as if it were an adjective (e.g. 'space travel'), but much less so than in German, and to say 'space dimension' sounds too Germanic.** dimension contributed more to the total variation ~~compared to~~ than did the temporal variabilities. However, their method is only valid for ~~the~~ one single dataset and is thus not able to evaluate the uncertainties if multiple datasets describe the same variable. But ~~the~~ a generalized method should be based on Sun's work, as one more dimension can be added for a specific analysis ~~for~~ of the uncertainties.

In ~~this~~ the present study, we aim to introduce a new approach ~~for~~ to estimating uncertainty among multiple datasets. The new uncertainty metric should avoid any averaging ~~in~~ over time or space~~dimension~~, so that all information ~~across the~~ along each of these two dimensions can be maintained for ~~uncertainty assessment~~ the assessment of the uncertainty. Multiple precipitation products will be used to display the results and explain the ~~peculiarity~~ specifics of the new ~~methodology. In section~~ method. In Section 2, the detailed ~~methodology~~ method of the three-dimensional variance partitioning approach is introduced. The characteristics of multiple precipitation datasets and estimations of two other classic uncertainty metrics are shown in ~~section~~ Section 3. The results ~~classicof~~ of the new approach for precipitation products are discussed in terms of the types of precipitation datasets in ~~section~~ Section 4. The differences between the new uncertainty estimation and two selected classic metrics used in uncertainty analysis are analyzed and discussed in ~~section~~ Section 5. ~~The discussion and conclusion are followed in the end of this article.~~ A discussion and some conclusions follow in Section 6.

2 ~~Methodology~~ Method and datasets

2.1 Mathematical Derivation

Multiple datasets recording the same climatic variable should be reorganized ~~to~~ into a three dimensional database~~. The database consists in three dimensions as~~, using the dimensions (1) **time** with a regular time interval (e.g. monthly or annual), (2) **space** with regular spatial units, ~~while~~ with all the grids ~~are~~ re-organized into one dimension from the original ~~latitude-longitude~~ latitude–longitude grids, and (3) **ensemble** as the third dimension ~~with~~ describing the different ensemble datasets. Thus, the dataset array can be ~~reformed as~~ re-organized to be

$$\mathbf{Z} = [z_{ijk}] \tag{1}$$

with the i-th time step ($i = 1, 2, \ldots, m$), j-th grid ($j = 1, 2, \ldots, n$), and k-th ensemble member or ensemble model ($k = 1, 2, \ldots, l$).

We define the three dimensions ~~as~~ to be time, space and ensemble dimension, and the means for these three dimensions ~~as~~ to be the temporal mean, spatial mean and ensemble mean~~, respectively~~. The corresponding variances are ~~named time variance, space variance~~referred to as the temporal variance, spatial variance, and ensemble variance~~, respectively. The~~. We also define the grand mean (μ), grand variance (σ^2) and the total sum of squares (SST) (or total variation) across the entire database~~.~~:

$$\mu = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{l} z_{ijk}/(mnl) \tag{2}$$

$$\sigma^2 = \frac{SST}{mnl} \tag{3}$$

$$SST = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{l} (z_{ijk} - \mu)^2. \tag{4}$$

The total variation ~~is contributed by the variation in~~ receive contributions from the variations along all three dimensions (Eq. 4). It can be reformulated as an ~~express of variations in~~ expression in terms of the variations along each of the three different dimensions. For instance, the derivation of the total variation can start from the third ensemble dimension. For a specific k^{th} ensemble member, the grand mean is formulated as $\mu_{ts}[k] = \sum_{i=1}^{m} \sum_{j=1}^{n} z_{ijk}/(mn)$, leading to the total ~~squares rewritten as~~ sum of squares being rewritten as

$$SST = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{l} (z_{ijk} - \mu_{ts}[k] + \mu_{ts}[k] - \mu)^2. \tag{5}$$

The SST can be further expanded and rearranged as

$$\begin{aligned} SST = \quad & \textstyle\sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{l} (z_{ijk} - \mu_{ts}[k])^2 \\ & + 2 \times \textstyle\sum_{k=1}^{l} (\mu_{ts}[k] - \mu) \underbrace{\left[\sum_{i=1}^{m} \sum_{j=1}^{n} (z_{ijk} - \mu_{ts}[k]) \right]}_{=0} \\ & + \underbrace{\left[\sum_{i=1}^{m} \sum_{j=1}^{n} \right]}_{=mn} \textstyle\sum_{k=1}^{l} (\mu_{ts}[k] - \mu)^2 \end{aligned} \tag{6}$$

$$SST = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{l}(z_{ijk}-\mu_{ts}[k])^2 + mn\sum_{k=1}^{l}(\mu_{ts}[k]-\mu)^2 \tag{7}$$

$$SST = mn\sum_{k=1}^{l}\sigma_{ts}^2[k] + mnl\sigma^2(\mu_{ts}), \tag{8}$$

~~Where~~ where $\sigma^2(\mu_{ts})$ is the variation of the grand mean for each ensemble member ~~.~~ and **Author query: In modern English technical writing, one must not begin a sentence with a mathematical symbol because of capitalisation issues. In the 1920s, people were not aware of this problem, but now they are.** $\sigma_{ts}^2[k]$ is the grand variance in ~~space and time dimension~~ the spatial and temporal dimensions for the ensemble member k. Moreover, $\sigma_{ts}^2[k]$ can be split using the mean of the spatial variation at each time step $\overline{\sigma_s^2[k,:]}$ and the variation of the spatial mean $\sigma^2(\mu_s[k,:])$, denoted as in Eq. (9) with its derivation ~~followed as Eq.~~ given in Eqs (10)~~to Eq.~~ ~~-~~(17).

$$\sigma_{ts}^2[k] = \overline{\sigma_s^2[k,:]} + \sigma^2(\mu_s[k,:]). \tag{9}$$

For a specific dataset k, the grand mean $\mu_{ts}[k]$ ~~through~~ at the spatiotemporal scale is

$$\mu_{ts}[k] = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} z_{ijk}. \tag{10}$$

The total ~~square for difference~~ sum of squares of the differences from the grand mean of this ensemble member is

$$SST[k] = \sum_{i=1}^{m}\sum_{j=1}^{n}(z_{ijk}-\mu_{ts}[k])^2 \tag{11}$$

and the grand variance σ_{ts}^2 is

$$\sigma_{ts}^2[k] = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(z_{ijk}-\mu_{ts}[k])^2. \tag{12}$$

The derivation can start from either the ~~space~~ spatial dimension or the temporal dimension. If the derivation starts from the ~~space~~ spatial dimension, Eq. (11) can be rewritten by incorporating the spatial mean of each time step $\mu_s[k,i] = \sum_{j=1}^{l} z_{ijk}/n$

$$SST[k] = \sum_{i=1}^{m}\sum_{j=1}^{n}(z_{ijk}-\mu_s[k,i]+\mu_s[k,i]-\mu_{ts}[k])^2. \tag{13}$$

~~It~~ This can be expanded and then rearranged as

$$
\begin{aligned}
SST[k] = & \sum_{i=1}^{m}\sum_{j=1}^{n}(Z_{ijk}-\mu_s[k,i])^2 \\
& +2\times\sum_{i=1}^{m}(\mu_s[k,i]-\mu_{ts}[k])\times\underbrace{\left[\sum_{j=1}^{n}(Z_{ijk}-\mu_s[k,i])\right]}_{=0} \\
& +\underbrace{\left[\sum_{j=1}^{n}\right]}_{=n}\sum_{i=1}^{m}(\mu_s[k,i]-\mu_{ts}[k])^2
\end{aligned}
\tag{14}
$$

$$
SST[k] = \sum_{i=1}^{m}\sum_{j=1}^{n}(Z_{ijk}-\mu_s[k,i])^2 + n\sum_{i=1}^{m}(\mu_s[k,i]-\mu_{ts}[k])^2 \tag{15}
$$

$$
\begin{aligned}
SST[k] = & \, n\sum_{i=1}^{m}\sigma_s^2[k,i] + nm\sigma^2(\mu_s[k,:]) \\
= & \, nm\overline{\sigma_s^2[k,:]} + mn\sigma^2(\mu_s[k,:])
\end{aligned}
\tag{16}
$$

The grand variance of this specific dataset is Eq. 17 (identical to Eq. 9).

$$
\sigma_{ts}^2[k] = \frac{SST[k]}{mn} = \overline{\sigma_s^2[k,:]} + \sigma^2(\mu_s[k,:]). \tag{17}
$$

Here, $\overline{\sigma_s^2[k,:]}$ is the mean of the spatial variation at each time step and $\sigma^2(\mu_s[k,:])$ is the variation of the spatial mean.

Or if we started the derivation from the time dimension, the grand variance can be split using the average of the temporal variation from all regions $\overline{\sigma_t^2[:,k]}$ and the ~~space~~ spatial variation of the temporal mean $\sigma^2(\mu_t[:,k])$:

$$
\sigma_{ts}^2[k] = \overline{\sigma_t^2[:,k]} + \sigma^2(\mu_t[:,k]). \tag{18}
$$

With Eq. (9) or Eq. (17) and Eq. (18), we ~~can have~~ obtain

$$
\sigma_{ts}^2[k] = \frac{1}{2}\left\{[\sigma^2(\mu_t[:,k]) + \overline{\sigma_s^2[k,:]}] + [\sigma^2(\mu_s[k,:]) + \overline{\sigma_t^2[:,k]}]\right\}. \tag{19}
$$

Author query: In English technical writing, equations are considered part of the text and must obey the laws of English grammar. This mostly concerns punctuation. The equals sign, if any, is considered the verb. The grammar of the equation affects what punctuation the preceding word receives as well as the punctuation that comes at the end of the equation. See, e.g., the style manual of the *Physical Review*, but this is a general rule. I have had to supply the punctuation for every single one of your equation displays. Substituting Eq. (19) into Eq. (8) results in

$$
\begin{aligned}
SST = & \quad \tfrac{mn}{2}\textstyle\sum_{k=1}^{l}[\sigma^2(\mu_t[:,k]) + \overline{\sigma_s^2[k,:]}] \\
& + \tfrac{mn}{2}\textstyle\sum_{k=1}^{l}[\sigma^2(\mu_s[k,:]) + \overline{\sigma_t^2[:,k]}] + mnl\sigma^2(\mu_{ts}).
\end{aligned}
\tag{20}
$$

The first term on the right-hand side of Eq. (20) can be transformed to ~~:~~

$$\frac{mn}{2}\sum_{k=1}^{l}[\sigma^2(\mu_t[:,k]) + \overline{\sigma_s^2[k,:]}] = mnl\left[\frac{\overline{\sigma_{s\_t}^2}+\overline{\sigma_s^2}}{2}\right], \tag{21}$$

where $\overline{\sigma^2_{s\_t}}$ is the mean value across ensemble members of the spatial variation of the temporal mean, and $\overline{\sigma_s^2}$ represents the grand mean of σ_s^2, which is the grand variance across ~~time~~ the temporal and ensemble dimensions. Eq. (20) then becomes ~~:~~

$$SST = mnl\left[\frac{\overline{\sigma_{s\_t}^2}+\overline{\sigma_s^2}}{2}\right] + mnl\left[\frac{\overline{\sigma_{t\_s}^2}+\overline{\sigma_t^2}}{2}\right] + mnl\sigma_e^2(\mu_{ts}), \tag{22}$$

where $\overline{\sigma^2_{t\_s}}$ ~~is~~ denotes the mean value across ensemble members of the ~~time~~ temporal variation of the spatial mean, $\overline{\sigma_t^2}$ ~~represents~~ denotes the grand mean of σ_t^2, the grand variance across space and ensemble dimensions~~.~~, and $\sigma_e^2(\mu_{ts})$ ~~represents~~ denotes the variation across ensemble members of the spatial-temporal means ~~(μ_{ts})~~.

Similarly, the global derivation of SST can start from any of the other two dimensions (i.e., space or time). ~~And the derivation is~~ This derivation can then be formulated as

$$SST = mnl\left[\frac{\overline{\sigma_{s\_e}^2}+\overline{\sigma_s^2}}{2}\right] + mnl\left[\frac{\overline{\sigma_{e\_s}^2}+\overline{\sigma_e^2}}{2}\right] + mnl\sigma_t^2(\mu_{se}) \tag{23}$$

$$SST = mnl\left[\frac{\overline{\sigma_{e\_t}^2}+\overline{\sigma_e^2}}{2}\right] + mnl\left[\frac{\overline{\sigma_{t\_e}^2}+\overline{\sigma_t^2}}{2}\right] + mnl\sigma_s^2(\mu_{et}), \tag{24}$$

~~Where~~ where each variable is defined in ~~the~~ Appendix A. Averaging these three expressions of SST defined in Eqs ~~.~~ **Author query: You asked for British English, which does not put a point after some abbreviations. The abbreviation 'Eq.' does receive a point, but the abbreviation 'Eqs' does not. In American English, all abbreviations, including 'Eqs.', receive points after them.** (22)~~-~~–(24) leads to

$$\begin{aligned} SST = \quad & \tfrac{mnl}{3}[\tfrac{\overline{\sigma_{t\_s}^2}+\overline{\sigma_{t\_e}^2}}{2} + \overline{\sigma_t^2} + \sigma_t^2(\mu_{se})] \\ & + \tfrac{mnl}{3}[\tfrac{\overline{\sigma_{s\_t}^2}+\overline{\sigma_{s\_e}^2}}{2} + \overline{\sigma_s^2} + \sigma_s^2(\mu_{et})] \\ & + \tfrac{mnl}{3}[\tfrac{\overline{\sigma_{e\_t}^2}+\overline{\sigma_{e\_s}^2}}{2} + \overline{\sigma_e^2} + \sigma_e^2(\mu_{ts})]. \end{aligned} \tag{25}$$

With the total ~~degree of freedom (~~number of degrees of freedom being $m \times n \times l$~~)~~, the grand variance is expressed as

$$\begin{aligned} \sigma^2 = \quad & \underbrace{\frac{1}{3}[\frac{\overline{\sigma_{t\_s}^2}+\overline{\sigma_{t\_e}^2}}{2} + \overline{\sigma_t^2} + \sigma_t^2(\mu_{se})]}_{V_t} \\ & + \underbrace{\frac{1}{3}[\frac{\overline{\sigma_{s\_t}^2}+\overline{\sigma_{s\_e}^2}}{2} + \overline{\sigma_s^2} + \sigma_s^2(\mu_{et})]}_{V_s} \\ & + \underbrace{\frac{1}{3}[\frac{\overline{\sigma_{e\_t}^2}+\overline{\sigma_{e\_s}^2}}{2} + \overline{\sigma_e^2} + \sigma_e^2(\mu_{ts})]}_{V_e}, \end{aligned} \tag{26}$$

where V_t, V_s and V_e ~~represent the time, space~~ denote the temporal, spatial and ensemble variances, respectively. ~~Moreover, an~~ An illustration of the present approach is shown in Figure 2 to facilitate the understanding of the partitioning results. The original database, consisting of multiple datasets~~is organized as three dimension~~, is re-organized into three dimensions (grey in the ~~center~~centre). Zones with different colors represent different processes of the original database from different dimensions (~~please~~ see the details in the caption of Figure 2 and Appendix A).

Note that ~~ensemble variance (~~the ensemble variance V_e ~~)~~ in Eq. (26) is a combination of several variations across the ensemble members. The four components are the variations of temporal and spatial values ($\overline{\sigma_e^2}$, zone B3), temporal mean ($\overline{\sigma_{e\_t}^2}$, zone C5), spatial mean ($\overline{\sigma_{e\_s}^2}$, zone C6) and the grand variance of the spatiotemporal mean for a single ensemble member ($\sigma_e^2(\mu_{ts})$, zone F3). Similarly, the other variances only rely on the variances in the corresponding dimension, which shows the independence ~~in~~ of the three dimensions. ~~It also illustrates~~ This also is an illustration of the fact that the uncertainty across ensemble members is similar to the temporal variation ~~or~~ and spatial heterogeneity.

Figure 2. ~~The illustration of~~ Partitioning the ~~partitioning time-space-ensemble~~ temporal-spatial-ensemble variance~~method~~. The original database is ~~reorganized~~ re-organized into three dimensions~~of~~: time, space and ensemble. Zones with different ~~colors~~ colours represent different processes based on the original database through different dimensions. The ~~denotations~~ labels of the zones are listed ~~in~~ on the right~~, while the~~; detailed definitions ~~of these denotations~~ can be found in Appendix A. The grand variance is ~~defined as~~ σ^2 and the grand mean ~~as~~ is μ. The subscripts t, s, and e ~~represent~~ indicate dimensions of time, space and ensemble, respectively. In Zone A~~(, μ_x) indicates~~ shows the ~~means value~~ mean values across the x~~dimension~~-dimension (x=t, s or e); ~~zone~~ in Zone B~~(, σ_x^2)~~ indicates the variation across the x~~dimension~~-dimension; ~~zone~~ in Zone C~~(, $\sigma_{x\_y}^2$)~~ indicates the variation across the x~~dimension~~-dimension of μ_y (y=t, s or e); ~~zone~~ in Zone D ~~(μ_{xy})~~ indicates the means across the x- and y~~dimensions~~-dimensions; ~~zone~~ in Zone E~~(, σ_{xy}^2)~~ indicates the variation across the x- and y~~dimensions~~-dimensions; ~~zone~~ in Zone F~~(, $\sigma_x^2(\mu_{yz})$)~~ indicates the variation across the x~~dimension~~-dimension of the means across the y- and z~~dimensions~~-dimensions (z=t, s or e).

2.2 ~~Metrics definition~~ Definitions of the metrics for model uncertainty

Although the total variation is ~~contributed by~~ a result of contributions from the spatial heterogeneity, temporal variability, and the uncertainties across different datasets~~. We~~, we mainly focus on the variance in the ensemble dimension because the spatial or temporal variation is natural for ~~the~~ climatic variables. The uncertainty among ensemble members is normalized as the ratio of the square root of the ensemble variance (V_e) to the grand mean value of the datasets (μ).

$$U_e = \sqrt{V_e}/\mu \tag{27}$$

Two classic metrics are also introduced for comparison. For each basic spatial unit (~~grid cell in this study~~ in the present study this means a grid cell), we can estimate the temporal mean of the target variable in each ensemble dataset as $\mu_t[j,k]$, $j = 1,...,n$ represents the ~~space~~ spatial unit, and $k = 1,...,l$ represents the ~~number of datasets~~index of the dataset**Author query: You seem to have not succeeded in expressing your intended meaning. It seems that l is the number of datasets, with each dataset is indicated by an index, the index running from 1 to l.** . Then we can estimate the variations across different ensemble datasets of the mean values as $\sigma^2(\mu_t[j,:])$ (expressed as $\sigma^2_{e\_t}[j]$ in this study). The spatial distribution of the $\sigma^2_{e\_t}$ shows the magnitude of the model uncertainty over space and its root $\sigma_{e\_t}[j]$ is the model deviation at each ~~space~~ spatial unit. The ~~estimation~~ estimate of this model deviation over the entire region can be expressed as ~~:~~

$$N.s.std = \sqrt{\overline{\sigma^2_{e\_t}}}/\mu = \frac{1}{\mu}\sqrt{\frac{1}{n}\sum_{j=1}^{n}\sigma^2_{e\_t}[j]}. \tag{28}$$

For each spatial unit, $\sigma^2_{e\_t}[j]$ $(j = 1,...,n)$ ~~has different values for each spatial unit and the values for~~ can take a different value. The values for all the grid cells are averaged to obtain $\overline{\sigma^2_{e\_t}}$, which shows the general magnitude of the ensemble variation over space. The quantity $N.s.std$ is normalized as the ratio of the square root of the averaged variations $\sqrt{\overline{\sigma^2_{e\_t}}}$ to the grand mean of all the datasets μ.

Similarly, the model uncertainty can be normalized as the ratio of the square root of the averaged ensemble variation but at different time steps $\overline{\sigma^2_{e\_s}}$ to the entire means~~(Eq. 29).~~ :

$$N.t.std = \sqrt{\overline{\sigma^2_{e\_s}}}/\mu = \frac{1}{\mu}\sqrt{\frac{1}{m}\sum_{i=1}^{m}\sigma^2_{e\_s}[i]}, \tag{29}$$

where ~~the~~ $\sigma^2_{e\_s}[i]$ $(i = 1,...,m)$ is the variation across different datasets of the spatial means of each product at each time unit $\mu_s[i,k],(i = 1,...,m,k = 1,...,l)$.

The two uncertainty estimates (Eqs ~~.~~28 and 29) correspond to the two classic metrics presented in the Introduction. We will compare ~~the~~ U_e with these two classic metrics ($N.t.std$ and $N.s.std$) to show their relations and differences.

2.3 Study area and data description

~~The China mainland is~~ Mainland China has been selected as the study area because of its large area and different ~~climate types~~ types of climate (?). Ten different subregions ~~are further~~ have been defined to facilitate the comparisons and analysis ~~on~~ of the strong spatial variations. The subregions are ~~listed as the~~ (1) Songhua River Basin, (2) Liao River Basin, (3) Hai River Basin, (4) Yellow River Basin, (5) Huai River Basin, (6) Yangtze River Basin, (7) Southeast China, (8) South China, (9) Southwest China, (10) Northwest China~~in~~, see Figure 3. **Author query: There are several issues with 'the southern China', which is certainly incorrect. The important issue is that one cannot write 'the southern China' although one can write 'the southern part of China' or 'southern China' without the 'the'. The less important issue is that southern China necessarily includes southeastern China and southwestern China. More usual, therefore, is 'South China', 'Southwest China', etc., where the meaning is whatever is generally understood, but South China does not have to logically include Southwest China unless that is what is generally understood. For example, in the U.S., 'the South' does not include 'the Southwest', for traditional and historical reasons, whereas southern U.S. does necessarily include the Southwest. Now, this issue is less important, and you can go back to what you were writing provided you omit your 'the' (or include 'part of'). The reason I have made this change consistently is because it would be easier for you to revert these changes than to implement my suggestions on your own.** The entire Chinese mainland is numbered as the 11~~st~~th region. Most of the subregions are natural river basins~~, and~~ ; this definition is more ~~proper for water resources~~ appropriate for water resource analysis than definitions using ~~longitude-latitude grids or that are~~ longitude–latitude grids or those based on administrative regions.

Figure 3. Ten subregions are ~~identified~~ defined in this study. These subregions are mainly ~~divided as the~~ river basins (~~regions 1-8~~Regions 1–8)~~and~~ , but 9 ~~as the southwestern~~ is Southwest China and 10 ~~as the northwestern~~ is Northwest China. ~~The~~ Region 11 ~~represents~~ is the ~~whole~~ entirety of the Chinese mainland.

Precipitation is one of the ~~sensitive climatic variables to the~~ climatic variables sensitive to large-scale atmospheric cycles and the local topography. Thirteen different precipitation datasets from various sources ~~are~~ have been collected for comparison (Table 1). These datasets ~~are~~ have been categorized into three groups according to the ~~methodologies~~ methods they used for generating the products, ~~i.e.~~ namely, gauge-based products, merged products and General Circulation Models (GCMs). The gauge-based products (~~i.e.~~ namely, CMA, GPCC, CRU, CPC and UDEL) use ~~observed data~~ data observed from global precipitation gauges. ~~While the density of~~ The density of the ground observation gauges, the representativeness of the gauges, and the interpolation algorithms for converting the gauge observations to ~~gridded dataset vary~~ a gridded dataset differ from product to product. ~~CMA (stands for~~ The CMA (China Meteorological Administration) dataset ~~uses the densest~~ has the densest distribution of gauges and probably has the best quality to capture the spatiotemporal variations of the precipitation over the study area. ~~But CMA~~ The CMA dataset is excluded when estimating the uncertainty among the gauge-based products~~but~~ ; it is chosen as the reference ~~datasets~~ dataset for comparison.

Among the merged precipitation products, the CMAP, GPCP and MSWEP use different sources of precipitation data (~~e.g.~~ namely, gauge observations, satellite remote sensing, and atmospheric model re-analysis). These different precipitation sources are averaged using different weights. Thus, the differences ~~among~~ between the three merged products are associated with the precipitation sources and the weight of the gauge observations. ERA-Interim is a re-analysis product~~, while~~ ; it uses near-real-time assimilation with data from global observations (**?**). Thus, the forecasting model is constrained by the observations and forced to follow the real system to some degree. Because of ~~the usage~~ its use of observations, ERA-interim ~~is also belonging to the~~ also belongs to the category of merged products.

GCM precipitation is a pure model estimation because observations are not used to constrain the simulations. The implemented physical and numerical processes will affect the accuracy of the model results. The lack of constraints on the GCMs will cause them ~~not following~~ to not follow the actual synoptic variability and explore other trajectories in the solution space. **?** repeatedly ~~run~~ ran the same GCM with a very small shift in the initial conditions. But the small difference leads to a spread in the model outputs after a number of running time steps (see Figure 2 in **?**). Therefore, the uncertainty in GCMs can be attributed to the differences in the model structures, parameter settings~~as well as~~ , and the initial conditions as well. There are more than 20 kinds of different GCMs~~, while only four of them are randomly chosento keep the~~ ; only 4 of them have been chosen, randomly, to maintain the same number of datasets ~~as~~ using the gauge-based products ~~and~~ as those using merged products.

All the products of the three precipitation types~~including CMA~~ , including CMA, are in gridded format. Although they differ in their original spatial resolution, all products ~~are interpolated to~~ have been interpolated to a 0.5^{o} spatial resolution to unify the spatial units. Annual average values are summed ~~up~~ based on their original time steps (daily or monthly) and the overlap time span of all the datasets is from 1979 to 2005 for all products.

3 Characteristics of precipitation and model quantified uncertainties with classic metrics

3.1 Spatial patterns of annual precipitation

The long-term annual mean precipitation (~~1979-2005~~1979–2005) obtained by averaging the precipitation from multiple datasets in the corresponding precipitation group is mapped in Figure 4. The annual mean precipitation obtained from the CMA dataset is 589.8 $\mathrm{mm\ yr^{-1}}$ (1.6 $\mathrm{mm\ day^{-1}}$) over the entire ~~mainland China~~Chinese mainland. The gauge-based precipitation has the least bias (-4.1$\mathrm{mm\ yr^{-1}}$, -0.7% in percentage) compared to the CMA precipitation. ~~Precipitation~~ The precipitation in the merged products and GCMs is larger than that of the CMA by 63.1 and 232.0 $\mathrm{mm\ yr^{-1}}$ (with the bias ~~as~~ equal to +10.7% and +39.3%), respectively.

The spatial pattern of the annual precipitation shows a decreasing gradient from ~~the southeastern~~ Southeast China (>1600 $\mathrm{mm\ yr^{-1}}$) to ~~the northwestern~~ Northwest China (<400 $\mathrm{mm\ yr^{-1}}$) in CMA and all other three precipitation groups. ~~While, they have different ability~~ They have different abilities to display the spatial gradient of the precipitation in some ~~details~~detail. For instance, some areas have abrupt precipitation changes rather than ~~following~~ follow the general gradient in CMA. This is probably caused by the sudden changes in topography (e.g., the northern Tienshan Mountain, the Qilian Mountains), ~~while it~~ which is not captured in the gauge-based products because some of the key gauges are not included in the production of the gauge-based products. The abrupt changes can be somehow represented by merged products and GCMs because the local variation due to topographic changes can be observed by other methods or by model algorithms. The precipitation in the merged products and the GCMs is higher than ~~CMA in Himalayas~~ that of CMA in the Himalayas, and particularly the GCMs show higher precipitation in the ~~northern~~ North Tibet Plateau as well as the southern part of the Hengduan Mountains. These differences show the general characteristics of the three types of precipitation products.

Figure 4. The annual precipitation over a long-term period (~~1979-2005~~1979–2005) for ~~the~~ each group of ~~the~~ precipitation datasets. (a) Annual precipitation of CMA dataset, (b) ensemble means of the annual precipitation over the precipitation products in gauge-based precipitation excluding CMA, (c) ensemble mean of the annual precipitation of all merged products, (d) ensemble means of the annual precipitation of all GCMs. The observations in Taiwan are not released in the CMA dataset.

Table 1. The precipitation datasets used in this study. Three different precipitation groups ~~are~~ have been identified according to the way the precipitation dataset is generated.

No.	Type	Name	Long name	Institute	Reference
1		CMA	China Meteorological Administration dataset	China Meteorological Administration	
2		GPCC	Global Precipitation Climatology Centre	the World Climate Research Programme (WCRP) and to the Global Climate Observing System (GCOS)	**?**
3	Gauge-based	CRU TS	Climatic Research Unit Time-Series	Climatic Research Unit (CRU) / Ian Harris, Phil Jones	**?**
4		CPC	CPC Global Unified Gauge-Based Analysis of Daily Precipitation	NCEP/Climate Prediction Center	**?**
5		UDEL	University of Delaware Air Temperature & Precipitation Global (land) precipitation and temperature	University of Delaware	**?**
6		CMAP	CPC Merged Analysis of Precipitation	NOAA CPC	**?**
7	Merged Products	GPCP	Global Precipitation Climatology Project	GSFC (NASA)	**?**
8		MSWEP	Multi-Source Weighted-Ensemble Precipitation	Princeton University, Princeton, NJ, USA	**?**
9		ERA-I	ERA-Interim	European Centre for Medium-Range Weather Forecasts	**?**
10		HadCM3	Hedley Centre Coupled Model Version 3	Met Office Hadley Centre, UK	
11	GCMs	IPSL-CM5A-LR		Insitute Pierre Simon Laplace, Paris, France	
12		CMCC-CM		Cetro **Author query: Should this be 'Centro' ? maybe 'cetro' is a kind of tree or something.** Euro-Mediterraneo per I Cambiamenti	
13		MIROC5		AORI, Chiba, Japan, NIES, Ibaraki, Japan, JAMSTEC, Kanagawa, Japan	

3.2 Spatial distribution of model uncertainties

In addition to the precipitation differences in its long-term annual means, differences can be found ~~among~~ between datasets within the same precipitation group. The spatial distribution of the model uncertainty for each precipitation group, which is expressed as the ensemble deviation of the annual precipitation from different precipitation products, is mapped in Figure 5.

Figure 5. The spatial distribution of model uncertainty in annual precipitation among different ensemble products. The uncertainty is expressed as the standard deviation of the annual precipitation across ensemble precipitation products of a specific group. The left panels are ~~uncertainty in~~ the values ~~and~~ of the uncertainty. The right panels are the ~~ratio~~ ratios of ensemble deviation to the ensemble means of the datasets in the corresponding group.

Among the datasets based on gauge observations, the ensemble deviation value is small in most of the land area of China (<50 $\mathrm{mm\ yr^{-1}}$, Figure 5-a). Although the deviation is higher in the south of China (50-100 $\mathrm{mm\ yr^{-1}}$), the area is not continuous in space. The highest deviation occurs along the Himalayas, indicating a high variation among the observed datasets. Regarding the merged precipitation products, the deviation shows high values (>200 $\mathrm{mm\ yr^{-1}}$, Figure 5-c) in ~~the southwestern~~ Southwest China (e.g., the Tibet Plateau, Yunnan Province, Guangxi Province). Moderate deviation is found in ~~the northeastern China, northern China and southeastern~~ Northeast China, North China and Southeast China. The deviation of precipitation has a correlation ~~between~~ with the topology, which indicates that the performance of the technologies used for the merged products are subject to the topologies as well. Compared to the gauge-based and merged products, the deviation among the selected GCMs has the highest value (>400 $\mathrm{mm\ yr^{-1}}$, Figure 5-e) in ~~the southern~~ South China, indicating a significant model uncertainty of the annual precipitation between different GCMs.

The ratio of the ensemble deviation to the mean value, which shows the model uncertainty with no ~~unit~~units, is very low in ~~the eastern~~ East China (<10%, Figure 5-b). ~~While, it~~ It is higher in ~~the western~~ West China especially in the Himalayas and the ~~northern~~ North Tibet Plateau. Similar to that of the gauge-based products, the uncertainty in the merged products has higher values in the ~~west than that in the east~~ West than in the East of China (Figure 5-d). The area with ~~the~~ a deviation ratio less than 10% is mainly distributed in ~~the southeastern~~ Southeast China and is apparently smaller than that of the gauge-based products, showing a decreasing similarity among different merged products. The area with a moderate deviation ratio (10%~~-40~~–40%) increases compared to that of the gauge-based products, and the area is mostly in ~~the middle~~ central and western China. The uncertainty estimated in the GCMs shows similar patterns in ~~western~~ West China to that of the merged products but with higher magnitudes in ~~the eastern~~ East China (Figure 5-f). Only the area in the ~~northeastern~~ Northeast and part of ~~the middle~~ central China features small uncertainty, less than 10%, and the deviation ratio rises significantly in ~~the southern~~ South China (e.g., the Pearl River basin), which corresponds to the high standard ~~deviation value~~ deviations in the GCMs shown in Figure 5-e.

The magnitude of the ensemble deviation demonstrates the model uncertainty among the different products in ~~a~~ the same precipitation group and ~~it~~ shows the ability ~~of precipitation estimation with different methodologies~~to estimate the precipitation with different methods. For all products, the ensemble deviation is relatively larger where the precipitation is higher, especially along the mountains and the subtropical regions. The deviation ratio is higher in ~~the northwestern China~~Northwest China, where the precipitation is among the lowest in China. Particularly for the gauge-based products, ~~the higher ratio occurs~~ higher ratios occur where the gauge density is low and the orographic effect is apparent (e.g., the Tibet Plateau and other mountainous area). For the merged products and the GCMs, the deviation ratio increases especially in ~~the southeastern~~ Southeast China, showing decreasing similarities among different precipitation products. Because the deviation ratio has taken into account both the variation and the means (which may have a systematic bias), the deviation ratio is better than the absolute ensemble deviation ~~to represent~~ at representing the uncertainty, and it is the most commonly used in ~~the~~ geographic studies.

3.3 Temporal evolution of model uncertainties

Figure 5 shows the spatial distribution of the ensemble deviation among different precipitation products. However, the temporal evolution of the deviation~~that shows the ability of products performance across~~ , which shows the performance of product over time and its changes, are not captured because the temporal variation has been averaged in order to estimate the spatial ensemble deviation in Figure 5. In this subsection, we examine the temporal evolution of the uncertainties in regional annual precipitation among different ensemble products. The analysis is based on the ten subregions defined in Figure 3 and the entire Chinese mainland.

Figure 6. The temporal evolution of the model uncertainty. The uncertainty is expressed as the normalized ensemble deviation of annual precipitation across ensemble datasets in each precipitation group for specific subregions. The value on the top right of each panel is the annual regional precipitation estimated in CMA dataset (~~1979-2015~~1979–2015). The annual precipitation is normalized as the ratio to the CMA long-term annual precipitation. The solid curve represents the ensemble mean of precipitation in each precipitation data group over the subregion. The width of the shaded area represents the standard deviation of the annual precipitation in each year among the datasets within that group (divided by the annual precipitation of the corresponding group). The shaded area ~~distributes~~ is equally distributed in the two sides of the ensemble mean values for the corresponding precipitation group.

The annual precipitation of each precipitation group has been normalized as the ratio to the long-term annual ~~means of~~ mean of the CMA in each subregion (black line in Figure 6). The magnitude of the annual precipitation in the gauge-based products (the blue solid line) is similar to that of CMA except in ~~the southwestern~~ Southwest China (Figure 6-i) for the overestimation along the Himalayas (Figure 4-a,b). The precipitation in the merged products (the green solid line) is higher in ~~the southwestern and northwestern~~ Southwest and Northwest China, in accordance with Figure 4-c. The annual precipitation of the GCMs (the red solid line) is apparently higher than that of the gauge-based products and merged products for almost all regions, which agrees with the spatial patterns in Figure 4-d.

The ensemble deviation across time scale is shown in the shaded area in Figure 6. It is estimated as the deviation of regional annual precipitation among different products in ~~a~~ the same group at a specific time step for each subregion. The deviation in normalized ~~for facilitating the~~ to facilitate comparisons between different subregions. High deviations are found in ~~the southwestern~~ Southwest China (Figure 6-i) in all three precipitation groups because of the large differences along the Himalayas. The deviations among the gauge-based products and the merged products in other regions are small and getting smaller with time. ~~It~~ This is mainly because more observations are ~~integrated and technologies improve~~ included and technologies have improved with time to control the ~~data quality~~ quality of the data. A large deviation is found in the merged products in ~~10-northwest~~ 10-Northwest China (Figure 6-j) and ~~the~~ 4-Yellow River Basin (Figure 6-d), where a dry climate dominates and the annual precipitation is among the lowest. The model deviation of GCMs varies ~~among~~ between regions as it is ~~smallest in the~~ at its smallest in 1-Songhua River Basin (Figure 6-a) and ~~the~~ 6-Yangtze River Basin (Figure 6-f), while it is among the highest in ~~the 8-south China and the west~~ 8-South China and West China (9,10), agreeing with the deviation maps in Figure 5.

Despite ~~of the~~ their mean values and ~~the magnitude of~~ magnitudes**Author query: It seems there was a word or symbol missing here?** , the temporal evolution of the gauge-based products and merged products agree well with ~~that~~ those of the CMA dataset, while the temporal evolution of ~~GCMs members~~ the members of the category of GCMs is weaker and not well correlated with that of the CMA. The main reason is that GCMs are not constrained in their synoptic variability and the sequence of ~~the~~ wet and dry years can be very different from that of the observations. A smoother result is thus obtained when we build the ensemble mean from the GCMs. ~~On the contrary to the~~ Unlike the weak variation in GCMs, the gauge-based and merged products have a strong co-variance and the ensemble mean preserves this co-variance.

For the entire ~~China~~ Chinese mainland (Figure 6-k), the ensemble deviation remains stable in different precipitation groups. In contrast, the annual precipitation spans the strongest spatial heterogeneity in the mainland compared to those divided by subregions (Figure 4). However, the spatial variation has been collapsed because the regional precipitation has to be obtained before the temporal analysis. It is therefore interesting to evaluate how the uncertainty changes when the variations ~~in~~ along both the time dimension and ~~in the space~~ the spatial dimension are considered in the precipitation datasets.

3.4 Variations ~~in~~ along the ~~time~~ temporal and ~~space~~ spatial dimensions

Previous subsections provide the deviation analysis in either temporal scale or the spatial scale. However, the two are seldom compared with each other. Herein, the standard deviation of the temporal and spatial variations in the precipitation datasets are compared in Figure 7 in ten subregions and the China mainland for different precipitation groups. The gauge-based products provide similar annual regional precipitation to CMA over the China mainland and the ten specific subregions except for the region ~~7-southeast~~ 7-Southeast China (Figure 7-g) and region ~~9-southwest~~ 9-Southwest China (Figure 7-i). While the merged products provide larger precipitation estimations for most of the regions. It might indicate the degraded ability of remote sensing, one of important data sources in the merged products, to estimate the precipitation amount in storms as the storms mainly contribute to the total precipitation for the two subregions. The regional precipitation in GCMs is even larger except in the region ~~8-south~~ 8-South China (Figure 7-h). These results indicate the degraded ability of merged products and GCMs in reproducing the total value of the annual precipitation.

Regarding the temporal and spatial deviations, ~~the~~ regions 9, 10 and 11 have the largest spatial standard deviation (~~in~~ as a ratio to the mean), indicating the strongest spatial heterogeneity over the regions. ~~The 7-southeast~~ Regions 7-Southeast China and the 3-Hai River have the smallest variations, either because of their small area or because ~~of~~ the homogeneity in these subregions is high. However, the spatial deviation in most of the subregions is larger than the temporal deviation. The ratio of the temporal deviation to the spatial deviation is among the smallest in the subregions 9, 10 and 11 (k=0.1, 0.12 and

Figure 7. The spatial standard deviation (horizontal) and temporal standard deviation (vertical) of the annual precipitation across ensemble datasets in each of the different precipitation groups for each subregion. The ~~P~~ P value in the ~~left~~ bottom left is the annual precipitation of CMA. The cross ~~center~~ centre represents the long-term means of the regional annual precipitation in ratio to the CMA mean value. The horizontal error bar represents the spatial standard deviation (spatial variation of the long-term annual precipitation at all the grids). The vertical error bar represents the temporal standard deviation (temporal variations of region-averaged annual precipitation in different years).

0.05, respectively. k is the ratio of the temporal deviation to the spatial deviation), showing an apparent difference between the ~~variation in~~ variations along the two dimensions. While, the difference between ~~variation in~~ the variations along the two dimensions is small in the 3-Hai River basin (k=1.15) and ~~7-southeast~~ 7-Southeast China (k=0.90), mainly due to the relatively strong variability of the annual precipitation in different years.

In addition to the differences ~~among~~ between regions, the variations in different precipitation groups also vary in magnitude. Excluding the CMA dataset, which consists of only one single product, the total variation (the sum of the spatial and temporal ~~variation~~variations) in the gauge-based products ~~are~~ is higher than that of the other two groups. ~~The~~ This difference demonstrates that the gauge-based products may have the largest ~~variation over space and the correlation among~~ spatial variation, and the correlations between the different gauge-based products are high~~so that the~~, so that this variation is preserved when ~~doing~~ passing to the ensemble. ~~On the contrary~~In contrast, the GCMs have the smallest variations, either because the precipitation estimated in the GCMs are more ~~homogenous over space than that~~ spatially homogenous than those of other precipitation products, or because the precipitation ~~estimation~~ estimations in different GCMs are not consistent in time or space since there are no ~~constrains in~~ constraints on the GCM simulation. The inconsistent precipitation patterns will be further eliminated when ~~doing ensemble mean with~~ carrying out an ensemble averaging over multiple datasets.

4 Variances in precipitation products

4.1 Variances in three dimensions

In the ~~above~~ preceding section, we introduced the spatial and temporal characteristics of the annual precipitation. The ~~precipitation variations in~~ variations in the precipitation in two dimensions among different precipitation products in ~~a~~ the same precipitation group ~~are~~ were estimated by two classic methods. In this section, we will present the uncertainty results estimated by the newly proposed ~~variance approach~~ approach to the variance. As introduced in the ~~methodology~~ methods section, the input annual precipitation to the approach is re-organized into three dimensions~~as~~: (1) **time**, 27 years from 1979 to 2005, (2) **space**, ~~the number of~~ 0.5° grids in a specific region and (3) **ensemble**, the number of ~~the models in a same~~ models in each precipitation group (four models ~~in all~~ for each of the three groups). Note that the estimated variance is for a specific subregion because it is an analysis based on regions and a long-term scale.

The grand variance (V, total value of the variance for all three dimensions) and its three components (i.e., variance in time ~~-~~V_t, space ~~-~~V_s and ensemble dimension ~~-~~V_e) for all the subregions ~~are~~ is mapped in Figure 8. The grand variance is similar in space in the precipitation groups of the gauge-based products and the merged products (Figure 8-a,b,c), while the grand variance in the GCMs is larger and is ~~approximating~~ approximately twice the V in the other two groups in regions ~~9-south China and 10-southwest~~ 9-South China and 10-Southwest China. The differences are mainly constituted by the ~~space~~ spatial variance and ensemble variance (Figure 8-i,l).

Figure 8. ~~The maps~~ Maps of the estimated grand variance (~~V~~V) and variances in different dimensions (V_t, V_s, V_e) across the ensemble datasets in each of the three different precipitation groups.

The ~~time variance (~~temporal variance V_t ~~)~~ is the smallest among all three ~~variance proportions~~ variances, and it has very little differences in ~~the northern~~ North China (Figure 8-d,e,f). ~~V_t~~ But it is higher in the gauge-based products than ~~that~~ in the merged products and GCMs in regions ~~8-southeast China and 9-south~~ 8-Southeast China and 9-South China, indicating a relatively strong temporal variation in the annual precipitation series, in accordance with the larger uncertainty ranges shown in Figure 6-h,i. Similar patterns of the ~~space variance (~~spatial variance V_s ~~)~~ are found in the gauge-based products and merged products (Figure 8-g,h). ~~The~~ Regions 7-Southeast River basin and ~~9-southwest~~ 9-Southwest China have the largest V_s because the precipitation significantly varies in space in these two subregions~~. V_s~~; it is higher in GCM precipitation especially in ~~the 9-southwest~~ 9-Southwest China, indicating the strong spatial heterogeneity in the GCM models over the Himalayas (Figure 8-i). The ensemble variance ~~(~~V_e ~~)~~ is relatively small in most regions for gauge-based products (Figure 8-j), indicating that the model variation among datasets in the observation group is small. ~~Similar~~ A similarly small V_e is found in the northern regions among the merged products as well as in the GCMs for the regions in ~~the northern~~ North China, while the intra-ensemble variations are large in the GCMs, especially in the ~~south especially the 9-southwest China and 8-south~~ South, especially 9-Southwest China and 8-South China (Figure 8-k,l).

~~In conclusion,~~ One can conclude that the grand variance and individual variance for each of the three different dimensions are generally larger in the precipitation group consisting of GCMs. The variations for the gauge-based products and merged products are similar in values and spatial distribution. However, in addition to the variances, the deviation defined as the ratio of the square root of the variance to the mean (e.g., U, U_t, U_s, U_e) contains extra information ~~of~~ about the regional means, and will be discussed in the following section.

4.2 Deviations in three dimensions

In contrast to the spatial gradient of the ~~variance magnitude distributed in~~ magnitude of the variance distributed over the ten subregions (Figure 8), the larger **Author query: Do you mean ‘largest’ ?** values of the total deviation ($U = \sqrt{V}/\mu$) occurs in the ~~northwest, but~~ Northwest, but a lower value generally occurs in ~~the southern~~ South China (Figure 9). The decreasing tendency of ~~precipitation magnitude~~ magnitude of the precipitation from the southeast to the northwest results in ~~the shift of~~ a shift of the spatial gradient compared to Figure 4. The total deviation ~~(~~U ~~)~~ is the highest in ~~the northwest~~ Northwest China (U=0.89, Figure 9-a,b,c) for all three precipitation groups~~. U~~, but is relatively small in the northeastern 1-Songhua River

(U=0.27) and 8-South China (U=0.29) for the gauge-based products. Subregion 6-Yangtze River has a relatively lower U in the merged products and GCMs in the ~~east~~ eastern part of China.

Figure 9. ~~The maps~~ Maps of deviations (U, U_t, U_s, U_e) estimated as the ratio of the square root of the corresponding variances (i.e., V, V_t, V_s, V_e) to the regional mean (μ) among the ensemble datasets in each of the three different precipitation groups. ~~Among which~~Of these, ~~the~~ U_e is considered ~~as~~ to be the model uncertainty.

~~The deviation in time and space dimension are inherent~~Deviations along the temporal and spatial dimensions are inherent, as they show the temporal evolution and spatial heterogeneity of the precipitation products. ~~Results show that the~~ The results show that U_t is small and contributes very little to the total U, indicating the weak fluctuation of annual precipitation compared to the spatial heterogeneity (Figure 9-d,e,f). The smallest ~~U_t value~~ value of U_t for the GCMs is in accordance with the weakest temporal variations in Figure 6. The deviation in ~~space~~ the spatial dimension (U_s) contributes the most to the total deviation, especially in ~~the northwestern~~ Northwest China (U_s=0.77 for the gauge-based products, Figure 9-g). The high U_s indicates the strong spatial heterogeneity of precipitation in the region, demonstrating that the ability to describe the precipitation ~~significant varies~~ varies significantly in different places in the subregions. However, because the spatial variations characterized ~~by GCMs in the northwestern China is~~ **Author query: I do not think 'characterized' is the right word here. The truth is that this word is very often misused by scientists and engineers. I think you should use either 'found' or 'determined' or 'obtained'.** by the GCMs in Northwest China are less significant than with the other two groups, the value of U_s for region ~~10-southwest~~ 10-Southwest China (=0.51) is smaller than that of the gauge-based and merged products.

The variations ~~in time and space~~ along the temporal and spatial dimensions show the natural precipitation patterns but the deviation of the values among multiple products (U_e) ~~show~~ shows the ability to consistently represent the spatiotemporal patterns. Therefore, U_e ~~therefore shows~~ indicates the uncertainty of the precipitation products among ensemble members of the same group. For the gauge-based products, ~~the~~ U_e is smaller than 0.1 for regions in ~~the eastern~~ East China, indicating that the model ~~variation~~ variations are relatively small compared to the annual means. The ~~U_e value~~ value of U_e is higher for ~~the 9-southwest~~ 9-Southwest China (=0.30) and ~~10-northwest~~ 10-Northwest China (=0.37), showing large variations even in the gauge-based products. For the merged products, U_e is similar to that of the gauge-based products in ~~the western~~ West China (=0.36), while it is larger in the ~~east especially for the~~ East, especially for 6-Yangtze River and 4-Yellow River (more than two times larger than U_e of the gauge-based products).

For the GCM precipitation, U_e increases compared to the other two groups in the eastern subregions, corresponding to the higher spatial model uncertainty in GCMs over the eastern regions shown in Figure 5. ~~While, it decreases in 10-northwest~~ It decreases in 10-Northwest China (U_e=0.25) and a possible reason for this is that the spatial homogeneity of the variations in ~~the region 10-northwest~~ 10-Northwest China (Figure 5-f) is stronger than that of the other groups (Figure 5-b,d,f). In the GCMs, the highest U_e occurs in ~~the southwestern China~~Southwest China, where both the means and the variations are higher (Figure 4 and 5). ~~In conclusion, the~~ One can conclude that U_e is linked with the magnitude of the model uncertainties in Figure

5 and Figure 6, indicating that ~~the U_e~~ it is to some degree correlated ~~to~~ with the classic metrics, as higher U_e covers the grid cells or regions with higher model uncertainty.

5 ~~Uncertainty and~~ Comparison of the uncertainty U_e with the classic metrics~~comparison~~

5.1 Deviation from the classic uncertainty metrics

In this section, we will compare the uncertainty ~~(U_e)~~ among ensemble members estimated by the three-dimensional partitioning approach with the two classic metrics (defined as $N.s.std$ in Eq. 28 and $N.t.std$ in Eq.29), to explain how these three metrics are related and differ with each other. As shown in Figure 10, U_e is correlated ~~to both the~~ with both $N.s.std$ and $N.t.std$. The correlation is stronger when U_e is smaller than 0.2, where the regions from 1 to 8 are generally included for all three precipitation groups. But U_e is in general larger than ~~the~~ $N.s.std$ and $N.t.std$ for the products. ~~The~~ This deviation is because the variation ~~of one dimension have~~ along one dimension has been collapsed when calculating the deviation ~~in~~ along the other dimension. For ~~the~~ subregions 9, 10 and 11, $N.s.std$ and $N.t.std$ deviate the most from the 1:1 line of U_e. Taking subregion ~~9-southwest~~ 9-Southwest China in the gauge-based products as an example, the temporal variance is 62.4 $\mathrm{mm\ yr^{-1}}$ while the spatial variance is 571.8 $\mathrm{mm\ yr^{-1}}$ (Figure 7-i). The difference between $N.s.std$ and U_e is 0.058 (=0.297-0.239, deviation ratio is 24.3%) when the temporal variation is collapsed. ~~While, the~~ The difference between $N.t.std$ and U_e is 0.126 (=0.297-0.171, deviation ratio is 73.4%) when the spatial variation is collapsed. The deviation is significantly larger than that between ~~the~~ U_e and $N.s.std$, showing that the collapse will induce a deviation ~~which relates~~ related to the magnitude of the collapsed dimension.

Figure 10. The relation of ~~the~~ U_e to the two classic metrics ~~as~~ (a) the normalized spatial standard deviation ~~-~~ $N.s.std$ and (b) the normalized temporal standard deviation ~~-~~ $N.t.std$. The two metrics are estimated with ~~eqs~~ Eqs 28 and 29 among the ensemble datasets in each of the three different precipitation groups.

These subregions (9, 10, 11) feature strong spatial heterogeneities (Figure 7-i,j,k) in the annual mean precipitation (Figure 4). The averaging process before estimating the classic metrics will cause a significant ~~smooth~~ smoothing of the datasets when the spatial heterogeneity among the datasets is very strong, because the spatial variation is significantly higher than ~~temporal variations~~ the temporal variation, as shown in Figure 7. The estimation of $N.t.std$~~which needs the averaging in spatial dimension~~, which needs an averaging over the spatial dimension, will lose more information than that in the time dimension. The deviation between $N.t.std$ and U_e (Figure 10-b) is larger than that between $N.s.std$ and U_e (Figure 10-a). The priority of the precipitation types also changes~~from the~~ , from model dominated (the model uncertainty in GCMs are larger than the other) to ~~the~~ region dominated (~~uncertainty in~~ the uncertainties in the specific regions 9, 10, and 11, are larger than in the other regions no matter ~~in~~ which precipitation data is used). This indicates that ~~difference of~~ the difference in model variation over space can be reflected in the new uncertainty U_e.

Each classic metric has its physical ~~meanings as the~~ meaning: $N.s.std$ represents the uncertainties over space and $N.t.std$ represents the uncertainties across time. The comparison of U_e with each of them demonstrates the metric performance on the same physical meaning. It is possible to compare U_e with a combination of the two classic metrics, but the combination ~~can~~ could be far more complex than a simple sum of the two classic metrics. However, ~~the~~ a qualitative comparison is accessible because U_e has a linear correlation with either of them. ~~The correlation will also remain~~ This correlation will persist, and occur between U_e and a combination of the two classic metrics by summing ~~up them~~ them up with certain weights.

5.2 Decomposition of the ensemble uncertainty

We now decompose the ensemble variance to ~~verify~~ determine the reason for the deviation of U_e from ~~the~~ $N.s.std$ and $N.t.std$. As shown in Eq. (26), the ensemble variance ~~(V_e) is formulated as~~ is expressed by

$$V_e = \frac{1}{3}\left[\frac{\overline{\sigma^2_{e\_t}} + \overline{\sigma^2_{e\_s}}}{2} + \overline{\sigma^2_e} + \sigma^2_e(\mu_{ts})\right]. \tag{30}$$

~~It~~ This combines four components which stand for the variation of different estimates across the ensemble dimension (i.e., the variance of original temporal and spatial values - $\overline{\sigma^2_e}$, of the temporal mean - $\overline{\sigma^2_{e\_t}}$, of the spatial mean - $\overline{\sigma^2_{e\_s}}$ and of the grand mean - $\sigma^2_e(\mu_{ts})$). Among ~~which, the~~ these, $\overline{\sigma^2_{e\_t}}$ is the mean of the ~~square of~~ squares of the spatial deviation in Figure 5-a,c,e for all grids in a specific region and $\overline{\sigma^2_{e\_s}}$ is the mean of the ~~square~~ squares of the temporal deviation in Figure 6 for each time step in a specific region. These two components are closely related to the two classic metrics $N.s.std$ (Eq. 28) and $N.t.std$ (Eq. 29), respectively.

Figure 11. The ~~proportion~~ proportions of the four components in Eq. (30) to ~~the~~ V_e among the ensemble datasets in each of the three different precipitation groups~~,~~: (a) gauge-based products, (b) merged products and (c) GCMs. The contribution is normalized so that ~~the~~ their sum ~~of them~~ is 1.0 for each region. Among the four components, ~~the~~ $\overline{\sigma^2_{e\_t}}$ and $\overline{\sigma^2_{e\_s}}$ are associated with the two classic ~~metric~~ metrics $N.s.std$ and $N.t.std$, respectively.

By decomposing Eq. (30), the contributions of the four components to the ensemble variance ~~(V_e)~~ are shown in Figure 11. For all three precipitation groups, $\overline{\sigma^2_e}$ is the dominant component simply because all the information on variations among the original datasets is retained in the uncertainty estimation. ~~While, the~~ The other three components ~~are estimations after~~ result from estimations after an averaging is performed~~in~~ , either over time, space, or the full spatiotemporal dimensions, which means a loss of information. The contribution of $\overline{\sigma^2_{e\_t}}$ and $\overline{\sigma^2_{e\_s}}$ is approximating 0.15 for regions from 1 to 8. ~~While~~ But $\overline{\sigma^2_{e\_t}}$ increases for ~~the region~~ regions 9, 10 and 11, indicating that ~~the spatial heterogeneity is significant for~~ there is significant spatial heterogeneity in these regions. ~~On the contrary~~In contrast, $\overline{\sigma^2_{e\_s}}$ decreases because the spatial averaging has collapsed the spatial variations. The very small contribution of $\overline{\sigma^2_{e\_s}}$ related to $N.t.std$ is the cause for larger deviations between $N.t.std$ and U_e in these subregions (Figure 10-b).

Although ~~all the components~~ any component can be used as ~~metrics~~ a metric for evaluating the variations among multiple datasets, there are limitations for each of the variations. For the variation of the temporal mean $\overline{\sigma^2_{e\_t}}$ and spatial mean $\overline{\sigma^2_{e\_s}}$, the collapse of a dimension has ignored part of the information. Moreover, the variation of the grand mean $\sigma_e^2(\mu_{ts})$ has ignored both the temporal variability and spatial heterogeneity, which further decreases its applicability ~~in uncertainty assessment~~ to the assessment of uncertainty. The variation $\overline{\sigma_e^2}$ is estimated based on the original data without averaging, and thus it represents the most information. However, it does not ~~account for~~ take into account the systematic uncertainty (bias in the mean values) which is expressed ~~as~~ by $\sigma_e^2(\mu_{ts})$.

Therefore, ~~neither~~ none of the single ~~component~~ components is able to represent the others. ~~Integrated~~ The integrated metric V_e is therefore a solution ~~to indicate~~ that represents all metrics to different degrees. What is interesting is that the variability of the proportions of $\overline{\sigma_{e\_t}^2}$ and $\overline{\sigma_{e\_s}^2}$ (or $\overline{\sigma_e^2}$ and $\sigma_e^2(\mu_{ts})$) are opposite and the sum of their proportions is stable, around 0.3 (or 0.7). This indicates a complementary relation between the two pairs of elements ($\overline{\sigma_{e\_t}^2}$ & $\overline{\sigma_{e\_s}^2}$; $\overline{\sigma_e^2}$ & $\sigma_e^2(\mu_{ts})$). On the other ~~word~~hand, some of the information is ignored in one of the components but ~~remained~~ remains in the other one within the same pair. ~~And therefore~~Therefore, the variation ~~in~~ along the time dimension and that ~~in the space~~ along the spatial dimension should be considered together~~as~~, as is done in the estimation of the ensemble variance ~~(~~V_e~~)~~. The normalized uncertainty ~~(~~U_e ~~)~~ derived from the integrated variation ~~(V_e), which has better ability to demonstrate the uncertainties compared to~~ , which is better able to determine the uncertainties than are the classic metrics, should be ~~a properer choice for the~~ the more proper choice for an uncertainty analysis.

5.3 ~~Metrics differences~~ Differences between the metrics in value and proportion

Figure 10 shows that ~~the~~ U_e is generally higher than the uncertainty identified by the two classic metrics (i.e., $N.s.std$ and $N.t.std$). Figure 12 then ~~summaries~~ summarizes the magnitude of the ~~deviation from~~ deviations of the classic metrics ~~to~~ from the new uncertainty ~~identified by~~ U_e. We can ~~find~~ see that the two classic metrics generally underestimate the uncertainty by around 0.03 (Figure 12-a). The variation of the underestimation of $N.t.std$ is larger than that of ~~the~~ $N.s.std$, showing a larger deviation between ~~the~~ U_e ~~with~~ and $N.t.std$. ~~Applying~~ Employing the new uncertainty metric will increase the ~~estimation of~~ estimated uncertainty by around ~~20-40%~~20%–40% for half of the cases~~compared to the~~, when compared to $N.s.std$ (Figure 12-b). For nearly 25% of the cases, the new U_e increases the ~~estimation of~~ estimated uncertainty by more than 50%. In ~~the~~ extreme cases, U_e is ~~larger than twice the~~ more than double $N.t.std$ (Figure 12-b). The results show that the widely applied uncertainty ~~estimated by~~ estimates from the two classic metrics have underestimated the uncertainty among different models / datasets. ~~The~~ Such an underestimation may especially occur for the temporal assessment of the uncertainties ($N.t.std$), which is very commonly seen in scientific reports and articles to illustrate the temporal evolution of the variables of interest.

Figure 12. The changes in (a) value and (b) percentage when using U_e as the new uncertainty metric compared to classic metrics $N.s.std$ (Eq. 28) and $N.t.std$ (Eq. 29).

6 Discussion and ~~Conclusion~~Conclusions

6.1 Features and applicability of the approach

The total variation of the database which consists of multiple datasets is contributed by the spatiotemporal variations as well as the uncertainties among ensemble datasets. While the uncertainty assessment with current approaches (e.g., eqs. 28 and 29) needs either the temporal variability or the spatial heterogeneity to be averaged which means a loss of information. The ~~proposed~~ variance partitioning approach proposed in this study works in three dimensions. It uses all the information ~~across the time and the space~~ over both the temporal and the spatial dimensions among the multiple ensemble members. It avoids the collapse of variation ~~in any dimension~~along any dimension, and thus the proposed uncertainty estimate U_e provides a more accurate ~~uncertainty estimation.~~ estimate of the uncertainty. The estimate U_e is especially suitable for ~~the~~ an overall assessment among multiple datasets over a certain period and over a specific space. ~~Although the compensation is that the~~ The trade-off is that U_e cannot provide the temporal evolution or spatial heterogeneity for users' consideration, even though in many cases we would like to know the general performance of the ensemble models based on a global single estimate.

The results of ~~the~~ this partitioning approach can be affected by the choice of the time step intervals. For example, the ~~time variance or time variance proportion~~ temporal variance or proportion of temporal variance will significantly increase if the time interval is chosen ~~as~~ to be one month. The inter-annual variability of precipitation will result in higher V_t. The changes depend on how significant the inter-annual variability is compared to the intra-annual variations. Moreover, only changes in the temporal variation (~~remain~~ the average values ~~but increase or reduce the variation magnitude~~remain but the magnitudes of the variation increase or decrease) can be captured ~~in the~~ by U_e. But $N.s.std$ will ~~keep~~ remain the same because the temporal variability has been neglected in the averaging process. ~~The case will be the same for~~ It is the same with $N.t.std$ if different spatial ~~resolution~~ resolutions of the measurements ~~is~~ are used.

The proposed approach has a flexible structure that can deal with different problems~~from~~, from a global scale to regional studies. The ~~time~~ temporal dimension can also ~~spans~~ span from daily, monthly, annual to decadal ~~analysis~~ analyses with different scopes. The ensemble dimension is applicable from ~~2~~ two members (i.e., model evaluation between simulations and observations) to any number of multi-models (consensus evaluation, **??**). The present approach is also applicable to any variables that are organized in ~~the~~ three dimensions, such as climatic variables (e.g., temperature, evaporation), hydrological variables (e.g., soil moisture, runoff) or environmental variables (e.g., drought index). Based on these advantages, ~~the~~ this three-dimensional partitioning approach can be widely applied in ~~the~~ hydro-climatic analysis.

6.2 ~~Conclusion~~Conclusions

A new three-dimensional partitioning approach ~~is~~ has been proposed in this ~~study~~ paper to assess the model uncertainties among multiple ensemble datasets. The new uncertainty metric (U_e) is estimated with an overall consideration of temporal and spatial variations as well as the differences among the ensemble products. ~~Results show~~ The results have shown that U_e is generally larger than the classic uncertainty metrics $N.s.std$ and $N.t.std$, which require a collapse of ~~variation in either of the time or space~~ the variation along either the temporal or spatial dimension. The deviation occurs where the spatial variations are significant but being averaged in the $N.t.std$ estimation. The ~~decomposing~~ decomposition of the total variance ~~(~~V_e ~~)~~ shows the complementary relation ~~of~~ between the two classic metrics, and therefore the new uncertainty U_e (derived from V_e) is a more comprehensive ~~estimation of~~ estimate of the uncertainty among multiple ensemble products.

Thirteen precipitation datasets generated by different ~~methodologies are~~ methods have been categorized into three groups (~~i.e.~~namely, gauge-based products, merged products and GCMs) and the model uncertainty in the ensemble products ~~is~~ has been analyzed with the new approach and with the two classic uncertainty metrics for each precipitation group. ~~The GCMs are identified with the largest model uncertainty with the classic metrics~~Using the classic metrics, in most regions, ~~while the new estimation~~ the GCMs have been indicated as having the largest model uncertainty. But the new estimator U_e indicates that the largest model uncertainty occurs in specific regions no matter ~~in~~ which precipitation group is considered. The impact of spatial heterogeneity on the model uncertainty has been represented well in the new uncertainty metric (U_e). In addition to the theoretical analysis of ~~U_e components~~ the components of U_e, the overall model uncertainty ~~(~~U_e ~~)~~ can be used as a new uncertainty estimate which involves more information and should receive more attention in the ~~uncertainty assessment field~~field of uncertainty assessment.