Reply to the Editor's Decision for Manuscript hess-2015-228
"Uncertainties in calculating precipitation climatology in East Asia"

Topical Editor Decision: Reconsider after major revisions (22 Nov 2015) by Prof. Dominic Mazvimavi

We greatly appreciate the Editor's thorough review and useful comments on our submission. This paper is the first work for assessing the uncertainties in calculating key climatological properties in East Asia due to the discrepancy between the datasets that are commonly used as the reference data (or "observations") in climate analyses and climate model evaluations. The lead author has recently published a similar work of very limited scope for the south Asia region. The major difference of this paper from the previous work, hence its uniqueness, is the systematic examination of the behavior of the uncertainty in calculating multiple statistical properties. This study is the first to report large discrepancy among observation datasets in the trend during the period of rapid warming. We believe that this is an important concern in climate research, especially in climate model evaluations, and is a crucial step in climate projection studies.

We provide the itemized replies to the questions from the Editor below; your comments/questions are depicted in italic and our replies starts with a long right arrow. Thanks again for your thorough review and comments.

Comments to the Author:

One of the Referees raised an issue regarding whether this paper is presenting any new information besides what has already been published in the paper by Kim et al (2015). The authors' response is that they are applying the same methodology to another region. The important issue to ascertain is whether this paper provides new findings or not. The response of the authors was that the study is being applied in another region with large population and farming areas. This response does not suggest that this paper is presenting new material. The authors should clearly demonstrate that this paper is contributing something new besides what they have already published.

$\Longrightarrow$ Unlike in the previous paper, Kim et al. (2015), this one reveals that uncertainties associated with inter-dataset differences tend to be larger for higher order statistical properties. This study examines for the first time the uncertainty in calculating the standard deviation, a widely-used first-order statistical moment, and linear trend against that in calculating the average, the zero-order statistical moment. Examining the uncertainty in assessing the key precipitation characteristics from the current available precipitation data can help interpret future precipitation projections. In East Asia, with huge populations and frequent hydrologic extremes, assessing long-term variations in precipitation has been an important concern. However, the effects of inter-dataset differences on such assessments have not been studied so far. The uncertainty analysis for the East Asia region in this study is also applicable to any other parts of the world.

It will be helpful if the authors briefly explain whether the gridded datasets used differ in their derivation. Were these datasets not derived using the same station observations? Assuming the *answer is yes, then will the differences be due to the gridding procedure?*

$\Longrightarrow$ The gridding procedure must have played an important role in the inter-dataset differences from which the uncertainties examined in this study originate. The analyses datasets examined in this study are based on different sets of observation station data due to the data availability at the time of analysis and specifics of the quality control procedures (e.g., Mitchell and Jones 2005; Yatagai et al. 2012; Pai et al. 2013). In addition, the analysis methodology, essentially the interpolation scheme that vary for different analyses datasets, can have contributed to the inter-dataset differences. Analyses of the effects of different datasets and/or the analysis schemes on the inter-dataset differences found in this study could not be examined due to the lack of the details in the gridding procedure as well as the amount of work needed to perform such an analysis. Not even those groups who generated these gridded analysis datasets have not provided such an analysis. Thus, such analysis of the origins of the inter-dataset differences is beyond the scope of this study.

Taking into account that precipitation tends to have long-term variations, and therefore use of a 25 year record for trend analysis is not advisable, the authors need to justify why this was done. Positive and negative trends may simply be due to the influence of the conditions at both the starting and ending points, i.e. wet/dry periods.

$\Longrightarrow$ This is a valid concern in most trend calculations. The period of the recent three decades examined in this study corresponds to the period of quite steady (near monotonic) and *large increases in the global mean temperature*. The analysis was limited to the 28-year period due to the length of the available data. Examination of the precipitation trend in the period of clear warming trend is a major scientific interest related to the link between the changes in precipitation and temperature. As the authors plan to continue to explore this topic, this work will be an important cornerstone for future works.

It is not clear which data was used to derive Figure 1. Was this done using observational data or one of the gridded datasets?

$\Longrightarrow$ Figure 1 is derived from the gridded datasets used in this study. This is now clarified in the revised manuscript by adding more explanation in the figure caption.

There is no indication regarding whether the significance of trends was statistically evaluated.

$\Longrightarrow$ Statistical significance of the linear trend is presented in the revised manuscript in terms of the $p-$values (in the newly implemented Figs. 3 and 4). This update is incorporated into the revised manuscript.

Some of the conclusions seem obvious or an established fact, e.g., high uncertainty in estimating mean annual precipitation and variability in the sparsely populated areas, deserts, Tibetan plateaus with limited observational data.

$\Longrightarrow$ This is true, but, to our opinion, needs to be added for presenting the overall results. We also include (repeat) this point in order to ascertain that the nations in those regions as well as international communities must pay more attention for gathering and distributing data in these regions. There have been a number of discussions/plans, but the progress has been, of course to our personal opinion, slow.

It is already an established fact that investment in setting up climatological stations is necessary to reduce uncertainty.

$\Longrightarrow$ As stated above, we try to re-ascertain this point to the related nations and communities. We will streamline this statement in the revised manuscript.

Reply to the Comments by Reviewer #3 for Manuscript hess-2015-228 "Uncertainties in calculating precipitation climatology in East Asia"

Reviewer #3 Decision: Accepted subject to minor revisions (16 Sep 2015)

General Comments: *The manuscript presents the uncertainty in calculating the fundamental climatological characteristics of precipitation over the East Asian region for the 28-year period from 1980 to 2007 using statistical comparisons. The manuscript is well written and concisely presents the characteristics of precipitation in East Asia based on five gridded raingauge observational precipitation datasets and assimilation data. However, there are not sufficient discussions about the observational datasets and assimilation data, i.e., locational information of observed data (raingauge station), and how to calculate the multi-dataset ensemble mean. Therefore, I recommend publication after considering the following comments.*

We greatly appreciate the reviewer for careful reading and valuable comments. We have included explanations on how to calculate the multi-dataset ensemble mean; however, the details of station locations from different datasets are not provided to the users. Please find our item-by-item responses to the reviewer's comments below.

Specific Comments:

1. *Page 7768 line 9-17: please, explain what are new methods for scientific findings in this study compared to a recent study of Kim et al. (2015). Is it a kind of case study applying to just different regions? If yes, the authors need to highlight why the study is needed.*

   $\Longrightarrow$ In terms of methodology, this work is largely similar to Kim et al. (2015), but the details are quite different. This study examines for the first time the uncertainty in calculating the standard deviation, a widely-used first-order statistical moment, and linear trend against that in calculating the average, the zero-order statistical moment. The other difference is the region of interest – East Asia, which includes huge populations, large farming areas and frequent hydrologic extremes; thus, assessing long-term variations in precipitation has been an important concern in this region. Examining the uncertainty in assessing the key precipitation characteristics from the current available precipitation data can help interpret future precipitation projections. However, the effects of inter-dataset differences on such assessments have not been studied so far. The uncertainty analysis for the East Asia region in this study is also applicable to any other parts of the world.

2. *Page 7769 line 7-15: More detailed description in the text regarding Table 1 is needed. The authors mainly discuss about the comparisons of statistical variables (mean, standard deviation, and linear trend) between the observational data and ensemble mean data as a reference data. Although the ensemble mean used in the study is constructed using an equal weighting, the availability of observed data could influence the mean. Furthermore, I would strongly suggest that the locational information of selected data is provided; for instance, over-plotting the approximate location of the field observation sites for each data. Page 7772 line 19-20: Please, provide distribution of the observation sites to support this sentence.*

⟹ Locations of actual observation sites included in an analysis often vary according to the quality control procedures. Moreover, the datasets incorporated into this study are provided by other groups. Thus, details like actual station locations are not known to us, and are beyond the scope of this study. The references in Table 1 provide detailed information on the specifics about individual datasets. For example, the station locations used in APHRODITE is available from the reference provided in Table 1 (Yatagai et al. 2012).

3. *Pages 7769 line 16-19: What is a definition of fine-resolution* $(0.25 \times 0.25?)$ *and coarse* $(2.5 \times 2.5?)$ *in this study?*

   ⟹ Within this context of this paper, fine or coarse resolution is mentioned in a relative sense not with rigorous definitions.

4. *Page 7769 line 16-19: Please, clarify what is "the same conclusions" when examined uncertainties of the coarse resolution GPCP data (Adler et al., 2003).*

   ⟹ "The same conclusions" implies the "conclusions" that we can obtain using the original five datasets only and that reported in the Summary and Discussions section. This sentence will be clarified in the revised paper.

5. Page *7770 line 10-12: Please, explain how to calculate the reference data (the multi-dataset ensemble mean). Before the comparison of five gridded precipitation datasets using Taylor diagram, the authors need to provide clearly how to grid observed data used in the current study.*

   ⟹ The multi-dataset ensemble was calculated by averaging all observational datasets included in the analysis using equal weights. The equal weighting is based on the fact that accuracy of individual datasets cannot be determined objectively, thus there is no ground to apply unequal weighting. This is now explained in the revised manuscript.

6. *Page 7772 line 14-16 and line 24-25, and Page 7773 line 1-2: Is SNR 5 a critical value to determine uncertainty? Please, explain more and provide some references.*

   ⟹ There is no established threshold SNR value. However, we may use some subjective guidance to interpret the SNR values. If SNR < 1, the signal is smaller than noise, a clear case that the signal is not reliable. SNR > 5 may indicate that the spread amongst the multiple datasets is small enough so that we can take the multi-data ensemble as the representative value for the included datasets.

6-1. *Page 7773 line 25-26: To draw a meaningful conclusion in trend analysis, authors need to show statistically significant trends over the regions before discussing uncertainty of the trends.*

   ⟹ We have provided the $p-$value plots in the revised manuscript. Please see the figures below (Figs. R1 and R2), which are newly implemented as Figs. 3 and 4 in the revised manuscript. The original Fig. 3 is now renamed as Fig. 5.

7. *Page 7781: Please, modify confusing color bar and different scales of each variable in Figure 1. Particularly for Fig. 1c and 1f the displayed color bars are difficult to distinguish between positive and negative trends (e.g.,* -0.003 *to* 0.000 *in Fig. 1c,* 0 *to* 0.0015 *in Fig. 1f). In order to compare the figures properly, the range of color bars should be synchronized.*

⟹ We wish we could modify the color bar, but the color plots are from an existing package. The authors do not have control to adjust color scales.

Technical Corrections:

Page 7769 line 22: Please replace “properties” with “property”. ⇒ Done.

Page 7774 line 5: You may want to replace “liner” with “linear”. ⇒ Done.

Page 7774 line 27: You may want to change “rage” to “range”. ⇒ We have removed “rage” from the sentence.

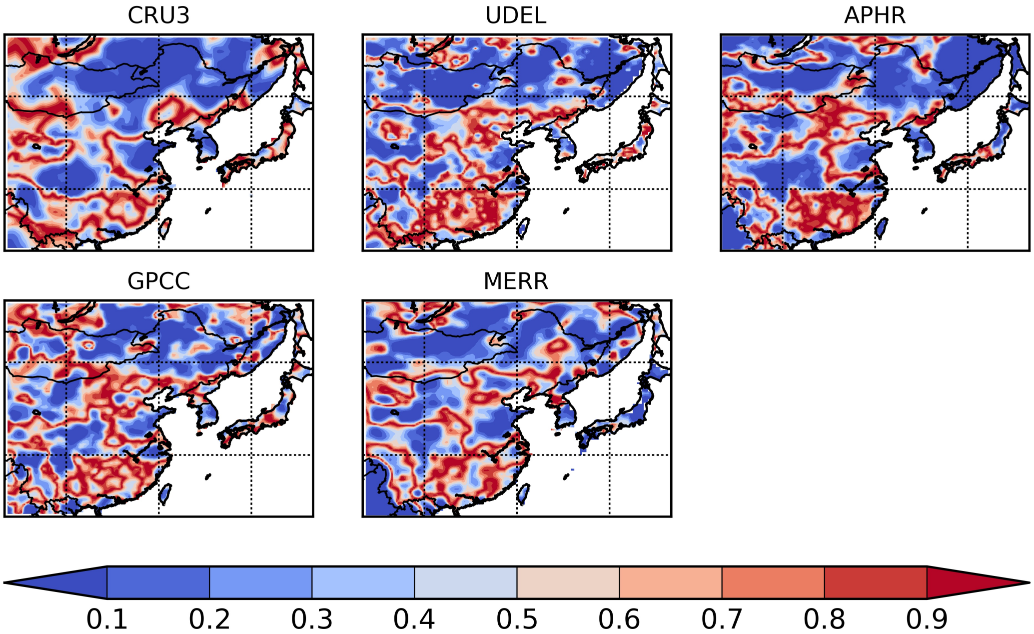

Fig. R1. The $p-$values in calculating the linear trend of the annual-mean precipitation from each dataset.

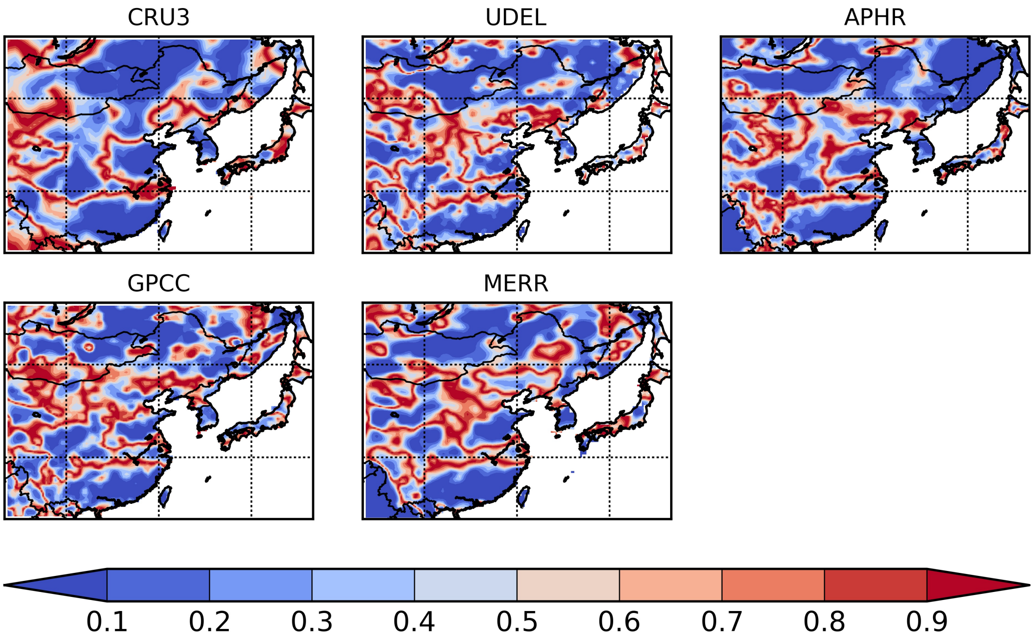

Fig. R2. Same as in Fig. R1, but for the summer-mean precipitation trend.

Reply to the Comments by Reviewer #2 for Manuscript hess-2015-228
"Uncertainties in calculating precipitation climatology in East Asia"

Reviewer #2 Decision: Accepted subject to minor revisions (15 Sep 2015) by Lydia Gachahi

Comments to the Authors: *The few technical and scientific comments based on my review of the paper are highlighted in my reviewed copy which I will attach. Please also note the supplement to this comment: http://www.hydrol-earth-syst-sci-discuss.net/12/C3656/2015/hessd-12-C3656-2015-supplement.pdf*

We greatly appreciate the reviewer for careful reading and valuable comments which were inserted in the separate pdf file. Please find our item-by-item responses to the reviewers comments below.

(1) Line 3: "in-situ rain gauge observations" ⇒ "in-situ rain gauge observations and data assimilations"

(2) Line 7: "multiple" ⇒ "two"

(3) Line 74: "entire" ⇒ "entire year"

(4) Line 87: "properties" ⇒ "property"

(5) Line 92: *Explain how the the multi-dataset ensemble was obtained.* ⇒ The multi-dataset ensemble was obtained by simple averaging of all datasets included in the analysis. The equal weighting is based on the fact that accuracy of individual datasets cannot be determined objectively, thus there is no ground to apply varying weights. This point is addressed in the revised manuscript.

(6) Line 96: "occurs" ⇒ "occur"

(7) Line 98-100: *Rephrase the statement.* ⇒ Rephrased by removing "This is not a concern in this study" (the first sentence). It now reads "Because it is practically impossible to rank the selected observational datasets in terms of their accuracy, the ensemble is constructed using an equal weighting."

(8) Line 145: *What is the significant or critical values of SNR?* ⇒ There is no established threshold SNR value to distinguish "good" from "bad". However, we may use some subjective guidance to interpret the SNR values. If SNR<1, the signal is smaller than noise, a clear case that the signal is not reliable. SNR>5 may indicate that the spread amongst the multiple datasets may be small enough so that we can take the multi-data ensemble as the representative value for the included datasets.

(9) Line 190: "liner" ⇒ "linear"

(10) Line 246: *Not clear in "It also suggests that ⋯"* ⇒ We made it clearer by rewriting the last two sentences starting from "Hence, without major investments ⋯" as "Remote sensing of precipitation will play important roles in monitoring precipitation over these regions of sparse

observations in addition to the investments for installing and maintaining additional surface observing stations.”.

Reply to the Comments by Reviewer #1 for Manuscript hess-2015-228 "Uncertainties in calculating precipitation climatology in East Asia"

Reviewer #1 Decision: Accepted subject to minor revisions (11 Sep 2015) by Hector Chikoore

We greatly appreciate the referee for thorough reading of the manuscript and useful comments. Below are our replies to the comments from the referee and actions to address referee's concerns.

General Comments: *The manuscript deals with a comparison of five gridded precipitation datasets based on raingauge observations in East Asia. Means, interannual variability and long-term linear trends are investigated for the 28 year period from 1980-2007. The manuscript is well written and presents an interesting analysis of precipitation in a region of complex geography. The major finding is that the datasets are in agreement on means and inter-annual variability whilst large uncertainty is found with linear trends. Uncertainties are larger over the drier and mountainous area of the domain perhaps due to sparse observations.* **(1)** *However, there seems to be little reference to the datasets in the discussion and summary, instead focusing on the means, interannual variability and linear trends. I would expect that some datasets agree more than with others.*

(1) Presenting only the multi-data ensemble characteristics in this paper is intentional because we try to avoid discussing accuracy of individual datasets. We treated all datasets as if they are of the same accuracy or uncertainty because there is no ground to argue accuracy of a specific dataset. We think this is the right way to address uncertainties in deriving precipitation climatology from available multiple observational data in the region.

Specific Comments: **(2)** *The datasets are described and cited in Table 1. Three datasets are of the same horizontal resolution* (0.5×0.5) *whilst the other two have different resolutions. This may not be a sound method for analysis of inter-dataset variability and may affect the calculation of the ensemble reference.* **(3)** *But perhaps the starting point should be to show a graph with annual cycles of the five rainfall datasets to show the spread or agreement. Rainfall trends are less certain in many regions of the world and even in IPCC projections.* **(4)** *I would suggest to analyze the datasets for trends which are not necessarily linear, for example polynomial trends. In some regions the polynomial trends are more significant than linear trends. Of course, a longer time-series depicts trends better.*

(2) This is an important point indeed. In real world, observational data comes in various resolutions and discretizations. In fact, datasets of the same horizontal resolution can be defined in different grid structures. Because of this, all datasets are interpolated onto a common grid so that we can compare all datasets at the same locations. The spatial interpolation procedure can affect the characteristics of spatial variability of the interpolated data. This is an important concern in deriving the characteristics of horizontal variability, e.g., spatial power spectra, but is not expected to have serious effects on deriving temporal variability of the interpolated data. Because all of the properties we concern in this study (temporal means, standard deviations, trends) are related with the temporal variability, we expect the differences in the horizontal resolutions and subsequent spatial interpolation have minimal impacts on the reported results. This point is now addressed in the revised manuscript.

(3) This will be very useful if we concern the variability at a single point or averages over a large area. This study concerns the spatial variations of selected properties of temporal variability. Hence, examining the annual cycle at all data points within the domain will be impractical.

(4) Very interesting suggestion of its own virtue. Nonlinear fitting is also straightforward as well. In this study, however, the record length of 28 years may not be sufficient to derive nonlinear trends. In addition, we do not have any prior information on the plausible shape of the nonlinear trend that can be of interest.

Technical Corrections: **(5)** *Few typos and language errors can be found in the manuscript. Page 7767, line 4, 'climate' should read 'climates'; Page 7767, line 26, delete 'an'; Page 7769, line 22, replace 'properties'; Page 7774, line 1-4, consider rewording; Page 7775, line 2, replace 'is' with 'are'; Page 7775, line 10-12 is a repetition of line 5-6; Page 7776, line 5, replace 'is' with 'are'.*

(5) We have incorporated all these suggestions in the revised manuscript. Thanks again for careful reading.

Manuscript prepared for Hydrol. Earth Syst. Sci.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 3 January 2016

Uncertainties in calculating precipitation climatology in East Asia

J. Kim[1,2] and S. K. Park[3,4,5,6]

[1]Joint Institute for Regional Earth System Science and Engineering, UCLA, Los Angeles, CA, USA
[2]Department of Atmospheric and Oceanic Sciences, UCLA, Los Angeles, CA, USA
[3]Department of Environmental Science and Engineering, Ewha Womans University, Seoul, Korea
[4]Department of Atmospheric Science and Engineering, Ewha Womans University, Seoul, Korea
[5]Severe Storm Research Center, Ewha Womans University, Seoul, Korea
[6]Center for Climate/Environment Change Prediction Research, Ewha Womans University, Seoul, Korea

Correspondence to: S. K. Park (spark@ewha.ac.kr)

Abstract. This study examines the uncertainty in calculating the fundamental climatological characteristics of precipitation in the East Asia region from multiple fine-resolution gridded analysis datasets based on in-situ rain gauge observations and data assimilations. Five observation-based gridded precipitation datasets are used to derive the long-term means, standard deviations in lieu of interannual variability and linear trends over the 28-year period from 1980 to 2007. Both the annual and summer (June–July–August) mean precipitation is examined. The agreement amongst these precipitation datasets are examined using ~~multiple~~ two metrics including the signal-to-noise ratio (SNR) defined as the ratio between long-term means and the corresponding standard deviations, and Taylor diagrams which allows examinations of the pattern correlation, the standard deviation, and the centered root mean square error. It is found that the five gauge-based precipitation analysis datasets agree well in the long-term mean and interannual variability in most of the East Asia region including eastern China, Manchuria, South Korea, and Japan, which are densely populated and have fairly high density observation networks. The regions of large inter-dataset variations include Tibetan Plateau, Mongolia, northern Indo-China, and North Korea. The regions of large uncertainties are typically lightly populated and are characterized by severe terrain and/or extreme high elevations. Unlike the long-term mean and interannual variability, agreements between datasets in the linear trend is weak, both for the annual and summer mean values. In most of the East Asia region, the SNR for the linear trend is below 0.5, i.e., the inter-dataset variability exceeds the multi-data ensemble mean. The uncertainty in the spatial distribution of long-term means among these datasets occurs both in the spatial pattern and variability, but the uncertainty for the interannual variability and time trend is much larger in the variability than in the pattern correlation. Thus, care must be taken in using

long-term trends calculated from gridded precipitation analysis data for climate studies over the East Asia region.

1 Introduction

Long-term means, standard deviations in lieu of interannual variability, and trends calculated from observed data are among the fundamental fields in representing the characteristics of regional ~~climate~~climates. These climatological properties play crucial roles in defining climatological norms, occurrence of extreme events, detection of climate change, and projecting future climate variations and change as well as their impacts (Giorgi et al., 1994; Groisman et al., 2001; Kim, 2005). For example, reliability of the climate change detection is examined by comparing the long-term means and trends calculated from observations against those simulated in climate model sensitivity experiments (e.g., IPCC, 2001, 2007). In addition, the changes in key local hydrological fields such as precipitation are frequently measured relative to their climatological means. Thus, calculating reliable values of these properties is a critical step in climate research for identifying regional climate characteristics, through quantification of their changes due to external and/or internal forcings such as emissions of anthropogenic greenhouse gases, and the impacts of such changes on regionally important sectors.

Gridded representations of observed data on the basis of a variety of instruments, locations, platforms, retrieval algorithms and analysis schemes are widely employed in climate research with various goals (Legates and Willmott, 1990; Mitchell and Jones, 2005; Shige et al., 2006; Schneider et al., 2014). Typically, only a limited number of such datasets were available, and most climate studies employed a single dataset which includes features needed for their analyses. Recently, a number of researchers and institutions have introduced newly developed observation-based gridded analysis datasets of global or regional coverage with fine spatial resolutions (Legates and Willmott, 1990; Adler et al., 2003; Mitchell and Jones, 2005; Shige et al., 2006; Yatagai et al., 2012; Pai et al., 2013; Schneider et al., 2014). These newly introduced analysis datasets provide precipitation and/or surface air temperatures over ~~an~~ extended periods of multiple decades at spatial resolutions of $0.5°$ or finer, which are substantial improvements from previous generation datasets that are typically at much coarser horizontal resolutions, for example, the $2.5°$ resolution GEWEX Global Precipitation Climatology Project (Adler et al., 2003). These recent fine-scale datasets allow us to better examine the regional precipitation and temperature climatology and to perform more reliable evaluations of today's high-resolution climate simulations, especially over the regions of complex terrain, that are important for climate-change impact assessments and climate model evaluations (Kim et al., 2013). These new datasets also introduce uncertainties in calculating regional climate characteristics because of the differences amongst them. Based on these concerns, two recent studies by Prakash et al. (2014) and Kim et al. (2015) examined uncertainty in calculating precipitation climatology over India and its surrounding regions using multiple precipitation analysis datasets. These two studies have revealed independently that there exist substantial amounts of differences amongst today's gridded precipitation datasets resulting in uncertainties in the calculated precipitation climatology and that the uncertainty or the spread amongst multiple datasets vary according to regions as well as seasons. Kim et al. (2015) further revealed that uncertainties in the calculated precipitation climatology defined relative to their climatological means are generally larger in the dry regions and/or local dry seasons. These two studies strongly suggest that uncertainty due to the differences between various datasets needs to be examined and quantified in all climate studies because the absolute accuracy of individual datasets cannot be quantified in practice.

~~This studyexamines~~ In this study, we investigate the uncertainty in calculating fundamental properties in representing regional climate characteristics of precipitation over the far east Asian region due to the differences amongst today's fine-resolution gridded datasets based on analyses of observed data. ~~The~~ This study examines for the first time the uncertainty in calculating the standard deviation, a widely-used first-order statistical moment, and linear trend against that in calculating the average, the zero-order statistical moment. Examining the uncertainty in assessing the key precipitation characteristics from the current available precipitation data can help interpret future precipitation projections. In East Asia, with huge populations and frequent hydrologic extremes, assessing long-term variations in precipitation has been an important concern. However, the effects of inter-dataset differences on such assessments have not been studied so far. The uncertainty analysis for the East Asia region in this study is also applicable to any other parts of the world. The methodology and data are presented in Sect. 2, and results are given in Sect. 3. Section 4 summarizes and discusses the implications of the findings in this study.

2 Methodology and data

In this study, spatial variations in the long-term means, interannual variabilities, and linear trends over the region of interest are examined in terms of inter-dataset variability measured using signal-to-noise ratio (SNR) and the similarity with reference data.

Five gridded precipitation datasets are used to estimate the uncertainty in constructing regional climate characteristics over East Asia for the entire year and for the summer season (June–July–August). Only the datasets that cover more than 25 years are selected for analysis for reliable calculations of the temporal variability in lieu of interannual variability and linear trends. The period of the recent three decades examined in this study corresponds to the period of quite steady (near monotonic) and large increases in the global mean temperature. The analysis was limited to the 28-year period (1980 ∼ 2007) due to the length of the available data. Examination of the precipitation trend in the period of clear warming trend is a major scientific interest related to the link between the changes in precipitation and temperature.

Based on the selection criterion, five high-resolution gridded datasets are selected, including the Climate Research Unit of the University of East Anglia (CRU), University of Delaware (UDEL), Global Precipitation Climatology Center (GPCC), the Asian Precipitation − Highly Resolved Observational Data Integration Towards Evaluation of water resources (APHRODITE), and the Modern Era Retrospective-analysis for Research and Applications (MERRA) land, that are either based on rain gauge data or assimilations. These datasets and references are summarized in Table 1. We also examined uncertainties including the coarse resolution Global Precipitation Climatology project (GPCP) data (Adler et al., 2003) to get essentially the same conclusions ~~;~~ that are obtained with the original five datasets only; thus, the results including the GPCP data are not presented here to focus on fine-resolution datasets.

Note that there are some factors leading to differences among the datasets − e.g., the horizontal and/or vertical resolutions, the gridding procedure, the analysis methods, etc. Such inter-dataset differences may be an unavoidable source of uncertainty in this study. As seen in Table 1, observational data are available in various resolutions and discretizations. In fact, datasets of the same horizontal resolution can be defined in different grid structures. The gridding procedure might also be different for different dataset. The analysis datasets are usually based on different sets of station (observational) data, depending on the data availability at the time of analysis and specifics of the quality control procedures (e.g., Mitchell and Jones, 2005; Yatagai et al., 2012; Pai et al., 2013) . Furthermore, the analysis methodology, essentially the interpolation scheme that varies for different analysis datasets, can contribute to the inter-dataset differences. However, assessing the effects of different datasets and/or the analysis schemes on the inter-dataset differences used here is beyond the scope of this study.

To alleviate the uncertainty related to the inter-dataset differences, we have interpolated all datasets onto a common grid so that we can compare all datasets at the same locations. The spatial interpolation procedure can affect the characteristics of spatial variability of the interpolated data. This can be an important concern in deriving the characteristics of horizontal variability, e.g., spatial power spectra, but is not expected to have serious effects on deriving temporal variability of the interpolated data. Because all of the properties we concern in this study are related to the temporal variability (e.g., temporal means, standard deviations, and trends), we expect the differences in the horizontal resolutions and subsequent spatial interpolation have minimal impacts on the results. We have also created a multi-dataset ensemble by simple averaging of all observational datasets included in the analysis, using equal weights. The equal weighting is employed because the accuracy of individual datasets cannot be determined objectively.

Uncertainties in representing precipitation climatology due to the spread amongst today's observational data are examined in terms of the SNR. The SNR has been a key ~~properties~~ property in a number of climate studies in which the uncertainty of climate signals are estimated against noises stemming from various sources (e.g., Giorgi and Mearns, 2002; Covey et al., 2003; Meehl et al., 2005; Tebaldi and Knutti, 2007; Duan and Phillips, 2010). In climate and weather forecast research based on ensembles of multiple model or observation datasets, the SNR has been used to measure the reliability of the multi-dataset ensemble mean against the spread of the datasets in the ensemble. Within this context, the signal and noise are defined as the associated mean and standard deviation, respectively, of multiple datasets. The definition of "noise" can be complicated when the data reliability varies among datasets and the weighting factor in constructing multi-dataset ensemble can vary for different dataset (Duan and Phillips, 2010). Such complications in calculating "noise" frequently ~~occurs~~ occur in climate projections where outputs from various models of varying performance are used to construct an ensemble mean using the variable weighting (e.g., Giorgi and Mearns, 2002). ~~This is not a concern in this study because~~ Because it is practically impossible to rank the selected observational datasets in terms of their accuracy~~; hence~~ , the ensemble is constructed using an equal weighting.

The similarity between individual datasets and the reference data defined as the multi-dataset ensemble is measured in terms of the pattern correlation and the standard deviation of individual datasets relative to the reference datasets. Measurements of these two properties are presented using Taylor diagrams (Taylor, 2001; Gleckler et al., 2008). Taylor diagram was first introduced by Taylor (2001) to provide a way to intuitively present two properties simultaneously; the correlation coefficient of a dataset with the reference data are presented in the azimuth angle (the angle for perfect agreement is zero) and the relative magnitude of the standard deviation of a dataset with respect to that of the reference data is expressed as the radial distance (e.g., see Fig. ~~3a~~5a). Thus, the radial distance of 1 and the azimuthal angle of 0° implies that a sample data has the same pattern and variability as the reference data. In addition, the distance between the point $(0°,\ 1.0)$ and a data point in this diagram corresponds to the centered root mean square error (RMSE). This diagram has become one of the most widely used methodologies in climate studies for presenting the evaluations of multiple models and/or variables or intercomparison of multiple datasets (IPCC, 2001; Taylor, 2001; Duffy et al., 2006; Gleckler et al., 2008; Kim et al., 2013, 2015).

3 Results

3.1 Regional climatology

Figure 1 presents the three basic characteristics of the annual and summer (June–July–August) precipitation climatology over East Asia – long-term means, interannual variability and trends, calculated from the ensemble mean of the multiple datasets in Table 1. The annual mean precipitation in the region is characterized by the wet regions in southeastern China and Japan (Fig. 1a). Precipitation over the Korean Peninsula is characterized by maxima in the southwestern and central regions and a rapid decrease towards the northwestern part of the peninsula bordering with Manchuria. The driest region covers southern Mongolia, the Gobi desert, and northern Tibetan Plateau. Interannual variability of the annual mean precipitation (Fig. 1b) also shows similar distribution as the annual means. Linear trend of the annual precipitation varies substantially according to geography (Fig. 1c). The most notable features include the positive trend in the driest region, including southern Mongolia, the Gobi desert and northern Tibetan Plateau, and the negative trend along the wet Yangtze River. Strong positive trends are also found in much of the Korean Peninsula, the coastal region of northern China to the west of the Shandong Peninsula, most of southern China, and eastern Japan. Decreasing precipitation trends also occur in the region between 45 and $50°$ N and extending from central Mongolia to far-eastern Russia. The summer rainfall climatology (Fig. 1d–f) resembles the annual mean climatology but with larger magnitudes. This shows that the precipitation climatology over the East Asia region is primarily determined by the summer rainfall.

3.2 Uncertainties in precipitation climatology

The climatology presented in Fig. 1 varies for different datasets. This is inevitable because each dataset utilizes different raw data, data quality control, and analysis methodology (Xie and Arkin, 1995). Because it is practically impossible to determine which dataset is more accurate, assessing the reliability of climatological properties calculated from various datasets as well as the expected range of uncertainty due to the diversity of these datasets is crucial in calculating regional climatology (Kim et al., 2015). In this section, the range of uncertainty in the three precipitation characteristics is measured in terms of the SNR and the agreement between individual datasets and the multi-data ensemble mean in terms of the spatial pattern correlation and the magnitude of spatial variability following the methodology of Kim et al. (2015)~~.~~, using the Taylor diagram.

The SNR is calculated as the ratio between the multi-data ensemble mean and the inter-dataset variability, i.e., a measure of the magnitude of the multi-dataset ensemble mean relative to that of the inter-dataset variations. Thus, as SNR increases, these datasets agree more closely with each other. ~~The spatial pattern correlation and the magnitude of spatial variability are examined using the Taylor diagram~~There is no established threshold value of SNR to distinguish “good” from “bad”. However, we may use some subjective guidance to interpret the SNR values. For instance, if SNR $<$ 1 the signal is smaller than the noise, and it becomes a clear case that the signal is not reliable. The case with SNR $>$ 5 may indicate that the spread amongst the multiple datasets may be small enough so that we can take the multi-data ensemble as the representative value for the included datasets.

The SNRs for the annual mean precipitation (Fig. 2a) and its interannual variability (Fig. 2b) over the 25-year period exceeds 5 in most of the study domain. Hence, the five datasets examined in this study agree well in terms of the annual mean precipitation and its interannual variability in the East Asia region. The regions of small SNR, i.e., showing poor agreements amongst the selected datasets, are located in the western part of the domain that include eastern Tibetan Plateau, the Gobi desert, and northern Indochina bordering with China. It is notable that the station density is relatively low in these regions. The SNR for the interannual variability is generally smaller than that for the mean; thus, uncertainty in calculating the interannual variability is larger than in calculating the mean climatology. Unlike the annual mean and its interannual variability, the SNR for the linear tendency of the annual precipitation (Fig. 2c) is generally below 5 in most regions. Thus, long-term annual precipitation trend in the region is highly uncertain except in a few small areas.

~~Figure 2d–f shows~~ Figures 2d$-$f show the SNR for the summer mean precipitation. Overall, the reliability of the three characteristics of the summer precipitation calculated from these five datasets is similar to that of the annual precipitation. The SNRs for the summer precipitation climatology is somewhat smaller than those for the annual precipitation climatology, but still largely exceed 5 in about the same region as for the annual precipitation. For the interannual variability (Fig. 2b vs. Fig. 2e) and linear trend (Fig. 2c vs. Fig. 2f), the five datasets agree more closely for the summer mean values than for the annual mean values. It is noteworthy that the positive tendency of the summer rainfall in southern China (Fig. 1f) is highly reliable as all five datasets agree closely (i.e., relatively smaller inter-dataset variations compared with the multi-dataset ensemble mean).

~~Figure~~ To evaluate the statistical significance of trends, we have plotted the $p-$values from each dataset in calculating the linear trend of the annual-mean precipitation and the summer-mean precipitation (see Figs. 3 and 4, respectively). The regions of large SNR correspond to the regions of small $p-$values in calculating the linear trend. This suggests that some of the uncertainty in the multi-dataset ensemble may be inherited from the uncertainty in calculating the trend from individual datasets. Still, a significant portion of the region of small $p-$values shows small SNR values. Thus, inter-dataset differences are the main cause of the uncertainty in calculating long-term trends.

Figure 5 measures the spatial variations in the three climatological properties represented by the five observational datasets using the Taylor diagrams and the simple multi-dataset ensemble as the reference. In these diagrams, the areas encompassed by the red polylines may be regarded as the range of uncertainty (see Kim et al., 2015). Thus, as the area is smaller, the uncertainty due to the differences between the examined datasets is smaller. The spread in the azimuthal and radial direction indicates the spread in the spatial pattern and in the magnitude of spatial variability, respectively. Similarly as from Fig. 2, the uncertainties in the spatial variations of the annual and summer mean precipitation and their interannual variability are much less than the uncertainty in the spatial variations of the linear trend. The distances from the reference data at the point indicated by a star (i.e., the reference point with both standardized deviation and correlation being equal to 1.0) to individual datasets for the means (~~Fig. 3a and d~~Figs. 5a and 5d) are similar to those for their interannual variability (~~Fig. 3b and e~~Figs. 5b and 5e), indicating similar level of spread amongst these datasets in representing these two properties of the precipitation climatology in the region. Regarding the linear trend (~~Fig. 3c and f~~Figs. 5c and 5f), compared to the means and their interannual variabilities, the distances between the reference point and individual datasets are much larger. This is another indication of the larger uncertainties in the linear trend represented by these datasets.

One interesting feature in the examination of the uncertainties in the spatial variability in Fig. ~~3~~ 5 is that the spreads in these datasets occur in both the spatial pattern and the magnitude for the annual and summer mean values; however, these datasets show more consistency in the spatial pattern than in the variability. ~~Figure 3b and e shows~~ Figures 5b and 5e show that the five datasets show similar spatial correlations with the reference data and that the predominant spread among these datasets are in the radial direction, i.e., the magnitude of the spatial variability. This feature is more pronounced for the linear trend (~~Fig. 3c and f~~Figs. 5c and 5f) which shows nearly ~~liner~~ linear distribution of the data points in radial directions, i.e., much smaller spread in the azimuthal direction (pattern correlations) than in the radial direction (magnitude of variability relative to the reference data).

4 Summary and discussions

The uncertainties in three fundamental climatological characteristics of the precipitation over East Asia due to the differences among available fine-scale observation-based gridded analysis datasets have been examined using the metrics selected for objectively measuring the spread of these properties calculated from individual datasets. The three climatological characteristics include the means, interannual variabilities, and linear trends in the annual and summer mean precipitation, which are key fundamental climatological characteristics widely used in studies for examining regional climate characteristics and model evaluations. The spread or the magnitude of disagreements amongst the selected datasets are measured using the signal-to-noise ratio (SNR) and examined visually using the Taylor diagrams which allows simultaneous evaluations of three properties – pattern correlation, standard deviations and the centered mean square errors between multiple datasets and a reference dataset.

The SNR values calculated from the five selected precipitation datasets show that the mean climatology of the annual and summer mean precipitation values and their interannual variability are highly reliable in much of East Asia except in southern Mongolia, the Gobi desert, and the Tibetan Plateau – the regions of sparse population and complex terrain. Precipitation measurements in regions of dry climate and complex terrain require high ~~rage~~ density networks (e.g., Kim et al., 2015). Unlike the climatological mean values and interannual variability, linear trends calculated over the 28-year period ~~is~~ are highly uncertain except in a few limited areas. It is striking that reliable estimations of the temporal trend of the annual mean precipitation (Fig. 2c) is very low compared to that for the means and the variability (Fig. 2a and b, respectively). Reliable calculation of linear trends is only possible over the southern China region for the summer mean precipitation. Thus extra caution must be taken when analyzing precipitation trends over the East Asian region.

The uncertainty characteristics also vary according to the climatological properties. Figures 1 and 2 discussed above show that the reliability of calculating temporal variabilities is much lower than that of time mean values, especially for linear trends~~which shows that reliable estimates of trends~~

~~may be obtained only in the southern China coastal regions during summertime~~. In addition, the spatial pattern and variability of the calculated linear trend (Fig. ~~3c~~5c) show much larger spread (i.e., uncertainty) among these datasets compared to the annual means (Fig. ~~3a~~5a) and interannual variability (Fig. ~~3b~~5b). The consistency in the spatial pattern between individual datasets and the reference data measured in terms of the correlation are near or over 0.95 for the temporal means and variability whilst it barely exceeds 0.8 for the linear trend. The range of spatial variability measured in terms of the standardized deviation (the ratio between the standard deviation of a datasets and the reference dataset) for the linear trend is over 0.5 which is more than twice of the range of the means and the variabilities. It is also observed that uncertainties in the spatial distribution of the annual and summer mean precipitation (Fig. ~~3a and d~~5a and 5d, respectively) occur in both the spatial pattern and the magnitude of variability. For the interannual variability and linear trends, the spread in the standardized deviation (i.e., the magnitude of variability) is much larger than that in the spatial pattern. These may suggest that all of these datasets are affected by some common factors in determining the characteristics of these datasets. For example, the station datasets included in each analysis dataset may provide high consistency in the spatial distribution pattern, but different analysis schemes may lead to a larger spread in the magnitude of their variability because of different basis functions employed in different interpolation schemes (e.g., Xie and Arkin, 1995; Prakash et al., 2014). This is just a hypothesis and needs close examinations in future studies.

The ~~uncertainties~~ uncertainty in calculating precipitation climatology in the regions including southern Mongolia, the Gobi desert, and the Tibetan Plateau is of a special concern. These regions can respond sensitively to climate change because of disproportionally larger impacts of global warming on high elevation regions and snow-ice processes (e.g., IPCC, 2007; Waliser et al., 2011). Because of rapid variations in the spatial precipitation distributions according to terrain during storms, accurate measurement of precipitation in the regions of extreme terrain requires high gauge network (Xie and Arkin, 1995). The sparse population density in these regions may require higher cost to build and maintain additional gauges to reduce the uncertainties. ~~Hence, without major investments in observations, the uncertainties in calculating precipitation climatology for these regions may remain large in the future. It also suggests that remote~~ Remote sensing of precipitation will play important roles in monitoring precipitation over these regions of sparse observations in addition to the investments for installing and maintaining additional surface observing stations.

Acknowledgements. This work is supported by the National Research Foundation of Korea grant (No. 2009-0083527) funded by the Korean government (MSIP) and the NSF ExArch 1125798 project. The APHRODITE data was obtained from the link http://www.chikyu.ac.jp/precip/.

References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., and Nelkin, E.: The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–Present), J. Hydrometeorol., 4, 1147–1167, 2003.

Covey, C., AchutaRao, K., Cubasch, U., Jones, P., Lambert, S., Mann, M., Phillips, T., and Taylor, K.: An overview of results from the Coupled Model Intercomparison Project, Global Planet. Change, 37, 103–133, 2003.

Duan, Q. and Phillips, T.: Beysian estimation of local signal and noise in multimodel simulations of climate change, J. Geophys. Res., 115, D18123, doi:10.1029/2009JD013654, 2010.

Duffy, P., Aritt, R. W., Coquard, J., Gutowski, W., Han, J., Iorio, J., Kim, J., Leung, L.-R., Roads, J., and Zeledon, E.: Simulations of present and future climates in the western United States with four nested regional climate models, J. Climate, 19, 873–895, 2006.

Giorgi, F. and Mearns, L.: Calculation of average, uncertainty range and reliability of regional climate changes from AOGCMs via the Reliability Ensemble Averaging (REA) method, J. Climate, 15, 1141–1158, 2002.

Giorgi, F., Brodeur, C., and Bates, G.: Regional climate change scenarios over the United States produced with a nested regional climate model: spatial and seasonal characteristics, J. Climate, 7, 375–399, 1994.

Gleckler, P., Taylor, K., and Doutriaux, C.: Performance metrics for climate models, J. Geophys. Res., 113, D06104, doi:10.1029/2007JD008972, 2008.

Groisman, P., Knight, R., and Karl, T.: Heavy precipitation and high streamflow in the contiguous United States: trends in the twentieth century, B. Am. Meteorol. Soc., 82, 219–246, 2001.

IPCC: Climate Change 2001: The Scientific Basis, edited by: Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van den Linden, P. J., Dai, X., and Johnson, C. A., IPCC, Cambridge University Press, Cambridge, UK, and New York, 881 pp., 2001.

IPCC: Climate Change 2007: Synthesis Report, edited by: Allali, A., Bojariu, R., Diaz, S., Elgizouli, I., Griggs, D., Hawkins, D., Hohmeyer, O., Jallow, B., Kajfez-Bogataj, L., Leary, N., Lee, H., and Wratt, D., IPCC, Geneva, 73 pp., 2007.

Kim, J.: A projection of the effects of the climate change induced by increased CO_2 on extreme hydrologic events in the western U.S., Climatic Change, 68, 153–168, 2005.

Kim, J., Waliser, D., Mattmann, C., Mearns, L., Goodale, C., Hart, A., Crichton, D., McGinnis, S., Lee, H., Loikith, P., and Boustani, M.: Evaluation of the surface climatology over the conterminous United States in the North American Regional Climate Change Assessment Program hindcast experiment using a Regional Climate Model Evaluation System, J. Climate, 26, 5698–5715, 2013.

Kim, J., Sanjay, J., Mattmann, C., Boustani, M., Ramarao, M. V. S., Krishnan, R., and Waliser, D.: Uncertainties in estimating spatial and interannual variations in precipitation climatology in the India-Tibet region from multiple gridded precipitation datasets, Int. J. Climatol.,~~doi:, online first, ~~, 35, 4557–4573, 2015.

Legates, D. and Willmott, C.: Mean seasonal and spatial variability in gauge-corrected, global precipitation, Int. J. Climatol., 10, 111–127, 1990.

Meehl, G., Covey, C., McAvaney, B., Latif, M., and Stouffer, R.: Overview of the Coupled Model Intercomparison Project, B. Am. Meteorol. Soc., 86, 89–93, 2005.

Mitchell, T. and Jones, P.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, Int. J. Climatol., 25, 693–712, 2005.

Pai, D., Sridhar, S., Rajeevan, M., Sreejith, O., Satbhai, N., and Mukhopadyay, B.: Development and analysis of a new high spatial resolution ($0.25^\circ \times 0.25^\circ$) long period (1901–2010) daily gridded rainfall dataset over India, Research Report No. 1/2013, National Climate Centre, India Metrorological Department, Pune, India, 63 pp., 2013.

Prakash, S., Mitra, A., Momin, I., Rajagopal, E., Basu, S., Collins, M., Turner, A., Rao, K., and Ashok, K.: Seasonal intercomparison of observational rainfall datasets over India during the southwest monsoon season, Int. J. Climatol., 35, 2326–2338, doi:10.1002/joc.4129, 2014.

Reichle, R., Koster, R., De Lannoy, G., Forman, B., Liu, Q., Mahanama, S., and Toure, A.: Assessment and enhancement of MERRA land surface hydrology estimates, J. Climate, 24, 6322–6338, doi:10.1175/JCLI-D-10-05033.1, 2011.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., and Rudolf, B.: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, Theor. Appl. Climatol., 115, 15–40, 2014.

Shige, S., Sasaki, H., Okamoto, K., and Iguchi, T.: Validation of rainfall estimates from the TRMM precipitation radar and microwave imager using a radiative transfer model: 1. Comparison of the version-5 and -6 products, Geophys. Res. Lett., 33, L13803, doi:10.1029/2006GL026350, 2006.

Taylor, K.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106, 7183–7192, 2001.

Tebaldi, C. and Knutti, R.: The use of the multi-model ensemble in probabilistic climate projections, Philos. Trans. R. Soc. A, 365, 2053–2075, 2007.

Waliser, D., Kim, J., Xue, Y., Chao, Y., Eldering, A., Fovell, R., Hall, A., Li, Q., Liou, K. N., McWilliams, J., Kapnick, S., Vasic, R., De Sales, F., and Yu, Y.: Simulating cold season snowpack: impacts of snow albedo and multi-layer snow physics, Climatic Change, 109, S95–S117, 2011.

Xie, P. and Arkin, P.: An intercomparison of gauge observations and satellite estimates of monthly precipitation, J. Appl. Meteorol., 34, 1143–1160, 1995.

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., and Kitoh, A.: APHRODITE: constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges, B. Am. Meteorol. Soc., 93, 1401–1415, 2012.

Table 1. The precipitation datasets employed in this study.

Dataset name	Source	Resolution	References
CRU	Rain gage	$0.5^\circ \times 0.5^\circ$	Mitchell and Jones (2005)
UDEL	Rain gage	$0.5^\circ \times 0.5^\circ$	Legates and Willmott (1990)
APHR	Rain gage	$0.25^\circ \times 0.25^\circ$	Yatagai et al. (2012)
GPCC	Rain gage	$0.5^\circ \times 0.5^\circ$	Schneider et al. (2014)
MERRA-Land	Assimilation	$2/3^\circ \times 0.5^\circ$	Reichle et al. (2011)

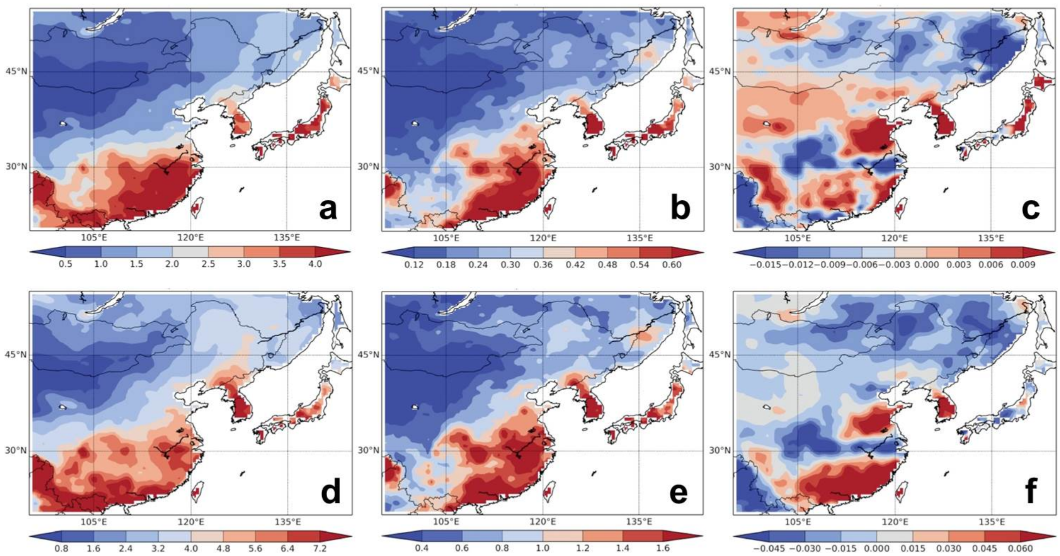

Figure 1. The climatological properties of the annual (upper panels) and summer (lower panels) precipitation for the period 1980–2007 over East Asia: **(a, d)** the mean climatology, **(b, e)** the standard deviation, and **(c, f)** the linear trend of precipitation. These properties are derived from the ensemble of the corresponding properties calculated from the datasets in Table 1.

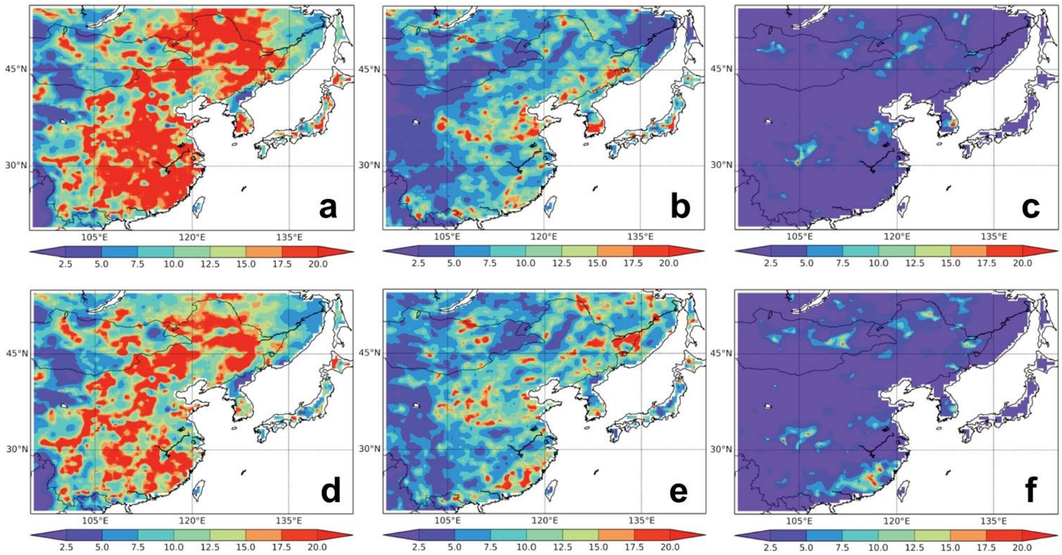

Figure 2. The signal-to-noise ratio (SNR) for the properties shown in Fig. 1, calculated from the corresponding properties of the five precipitation analysis datasets ~~.~~in Table 1.

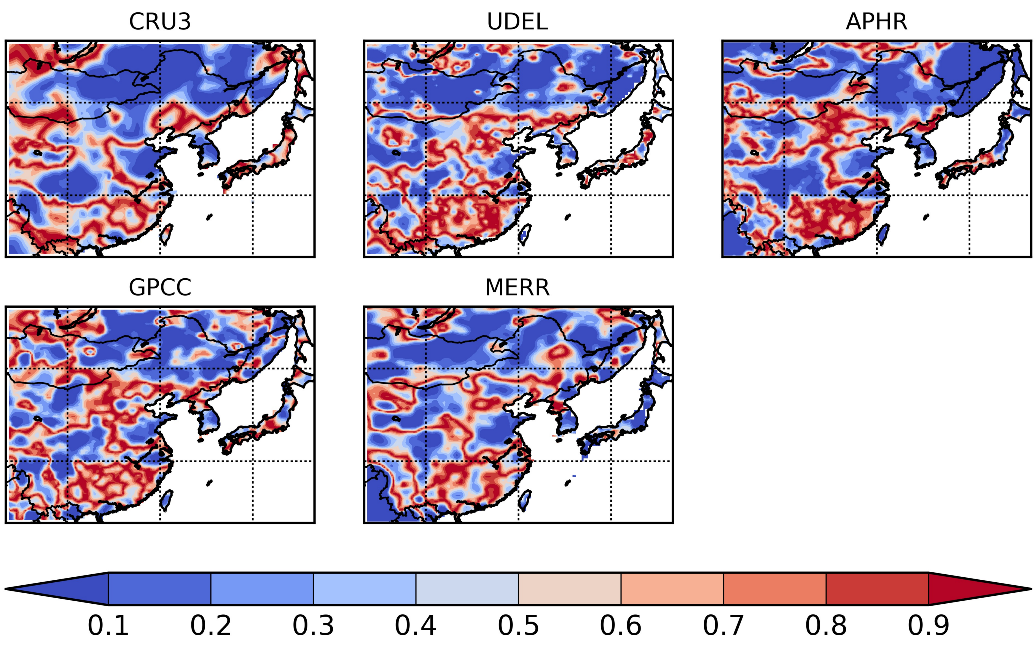

Figure 3. The $p-$values in calculating the linear trend of the annual-mean precipitation from each dataset.

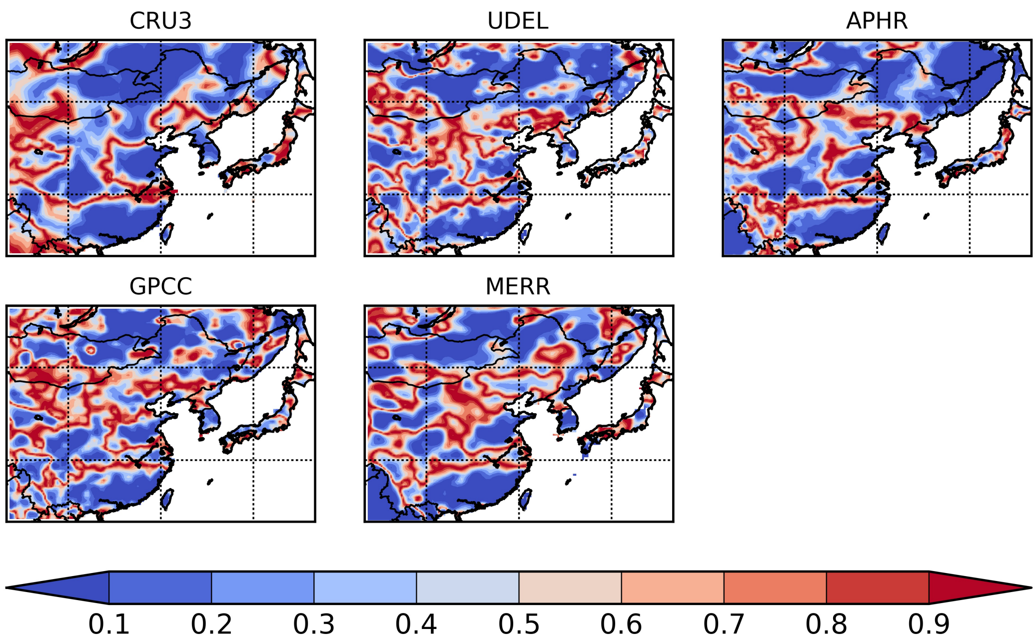

Figure 4. Same as in Fig. 3, but for the summer-mean precipitation trend.

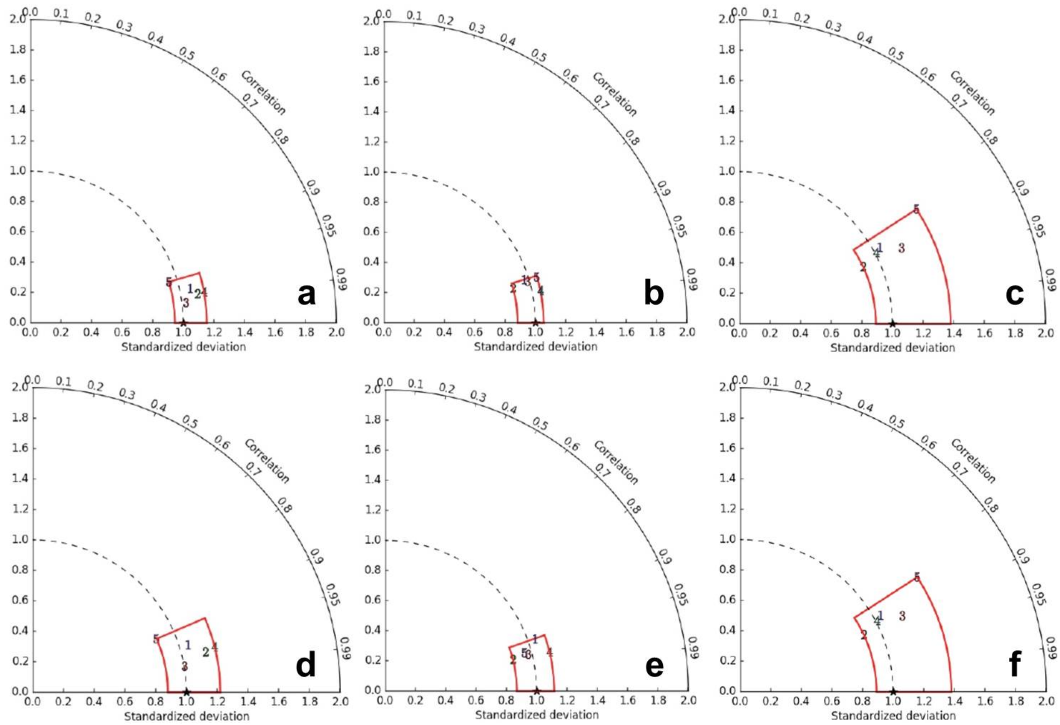

Figure 5. The spread amongst the five precipitation datasets in representing the spatial variability of the three climatological properties of the annual (upper panels) and summer (lower panels) precipitation over East Asia: **(a, d)** the mean, **(b, e)** the interannual variability, and **(c, f)** the trends of precipitation. They are presented in terms of their spatial pattern correlations (the azimuthal direction), the standardized deviation, and the standard deviation of individual datasets normalized by that of the reference data (the radial direction). The area within the red polyline represent the range of spread amongst these datasets.