Response

Editor:

I would like to thank for your responses to the comments. It would be useful to upload the revised manuscript, including a track changes version and a point-by-point response. Looking forward to your revisions.

We appreciate the decision and have improved the paper following the insightful and constructive comments provided by the reviewers. Please see the point-by-point responses in the following.

Anonymous Referee #1

The authors used the spatial plotting method to identify significant patterns of fully coupled GCMs from ten sets of NMME. Overall, the manuscript is well crafted with clear structures.

We appreciate the positive comments.

Some grammatical errors exist and need careful proofreading. I have some minor comments.

Thank you for the constructive comments. We have improved the paper accordingly and proofread the whole paper.

1. Do you use some bias correction or downscale methods for the NMME forecasts? If yes, suggest to give some details.

Thank you.

"In the analysis, the attention is paid to the retrospective forecasts

$$F_{GCM} = \left[f_{s,l,n,y,x} \right]_{GCM} \tag{1}$$

In Eq. (1), f represents forecast values that are specified by the 5 dimensions; F, which is the set of forecasts, is marked by the GCM that generates the forecasts. It is noted that in NMME, F_{GCM} are raw forecasts generated by GCMs and are not bias-corrected or downscaled." (Page 4, Lines 103 to 107)

2. Suggest to give the difference between the anomaly correlation and the spatial plotting method. Or the advantage of the spatial plotting method against the simple anomaly correlation.

Thank you for the constructive comment. The advantage of spatial clustering is highlighted in the introduction:

"Spatial plotting with latitude and longitude has been extensively used to handle the dimensionality for the verification of GCM forecasts [Kirtman et al., 2014; Hudson et al., 2017; Slater et al., 2017]. The fact that forecasts are commonly generated by GCMs as grid-based data makes spatial plotting a particular tool of choice for verification

[Merryfield et al., 2013; Saha et al., 2014; Jia et al., 2015]. As to anomaly correlation, spatial plotting overcomes tedious eyeball search by grid cell and is effective in locating where there is a good correspondence between forecasts and observations and where the correspondence is not satisfactory [Luo et al., 2013; Saha et al., 2014; Crochemore et al., 2016; Zhao et al., 2018, 2019b]. Similarly, spatial plotting applies to other verification metrics, such as bias and CRPS, and facilitates the examination of forecast attributes [Hersbach, 2000; Gneiting et al., 2007; Kirtman et al., 2014].

The extensive use of spatial plotting underlines the importance of testing the significance of spatial patterns. In spatial statistics, one of the fundamental issues is "are the spatial patterns displayed by the spatial plots significant in some sense and therefore worth interpreting?" [Cliff and Ord, 1981; Anselin, 1995; Getis, 2007]. However, the test of significance is commonly missing in the spatial plotting of GCM forecasts. In other words, verification metrics, such as anomaly correlation, are calculated for each grid cell and then shown as they are. To some extent, the interpretation of predictive performance depends on the color schemes, which are selected subjectively to represent the scale of verification metrics. There is the first law of geography – "everything is related to everything else, but near things are more related than distant things" [Tobler, 1970]. As to spatial plotting, the indication is that when verifying forecasts at one grid cell, attention also needs to be paid to forecasts at surrounding grid cells. For anomaly correlation, a grid cell with high correlation between forecasts and observations can be surrounded by grid cells with similarly high correlation, or by grid cells with low correlation. In the former case, the grid cell is located in a region where the GCM forecasts tend to perform well. But in the latter case, the high correlation can be a suspicious outlier. Moreover, previous studies observed grid cells with negative anomaly correlation, i.e., large (small) values of forecasts correspond to small (high) values of observations [Zhao et al., 2017b, 2018, 2019b]. In such a case, forecasts are cautiously wrong. Therefore, it is critical to characterize the different cases in spatial plotting and test whether the spatial patterns are significant and worth further attention.

In this paper, we are motivated to introduce spatial statistics [e.g., Di Luzio et al., 2008; Lu and Wong, 2008; Woldemeskel et al., 2013] to investigate the spatial plotting of anomaly correlation at the global scale…" (Pages 2 to 3, Lines 50 to 74)

3. Line 102: please check the start updated date of real-time forecasts.

Thank you for the insightful comment. More details on real-time forecasts are added:

"Ten sets of precipitation forecasts, as well as CMAP observations, in the NMME are downloaded from the International Research Institute at the Columbia University (https://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/). Their retrospective forecasts are complete in the period from 1982 to 2010 [Merryfield et al., 2013; Saha et al., 2014; Jia et al., 2015]. In the meantime, their real-time forecasts are updated periodically in a slightly different setting; for example, CFSv2 forecasts are generated since January 2011 using initial conditions of the last 30 days, with 4 runs from each day (https://www.cpc.ncep.noaa.gov/products/CFSv2/CFSv2_body.html)." (Page 4, Lines 98 to 103)

4. Suggest to add the equation of the anomaly correlation calculation and give some details about the climatology for the anomaly.

Thank you for the suggestion. We have added the equation:

"The start time s in Eqs. (1) and (2) comprises year k, i.e., 1982, 1983, …, 2010, and month m, i.e., January, February, …, and December. The predictive performance of GCM forecasts exhibits seasonality [Yuan et al., 2011; Zhao et al., 2017a, 2017b]. Accordingly, in the analysis, forecasts are selected by fixing m while varying k, e.g., pooling forecasts initialised in June 1982, June 1983, …, June 2010. The anomaly correlation is calculated by relating forecasts to the corresponding observations

$$r = \frac{\sum_k \left(rf_k - \overline{rf}\right)\left(ro_k - \overline{ro}\right)}{\sqrt{\sum_k \left(rf_k - \overline{rf}\right)^2} \sqrt{\sum_k \left(ro_k - \overline{ro}\right)^2}} \qquad (2)$$

The above formulation deals with k and omits other dimensions, including m, l, y and x, for the sake of simplicity. In Eq. (3), rf_k (ro_k) is the rank of year k's forecast ensemble mean (observation) in the 29 years' ensemble mean (observations); and $\overline{rf}$ ($\overline{ro}$) is the mean value of rf_k (ro_k). In general, the anomaly correlation characterises how well large (small) values of ensemble mean correspond to large (small) values of observations. Good (poor) correspondence makes r tend towards 1 (–1)." (Page 5, Lines 112 to 119)

5. Suggest to give some explanation for the forecasts of total precipitation in three months.

Thank you. More information is provided:

“The spatial clustering is performed for the anomaly correlation across the ten sets of forecasts in NMME. In the analysis, the attention is mainly paid to June, July, and August (JJA), which are generally boreal summer and Austral winter. Specifically, the start time of the forecasts is June, and the forecasts at the lead times of 0, 1, and 2 months are aggregated to form the seasonal forecasts. In the meantime, forecasts initialized in September of total precipitation in September, October, and November (SON), forecasts initialized in December of total precipitation in December, (the next) January, and (the next February) (DJF), and forecasts initialized in March of total precipitation in March, April, and May (MAM) are also investigated, with the results presented in the supplementary material.” (Pages 7 to 8, Lines 168 to 174)

6. Are captions 4.1 and 4.2 the same? Please check.

We are very sorry for the typo. The captions have been modified in the revision:

“4.1 Anomaly correlation in JJA” (Page 8, Line 175)

“4.2 Anomaly correlation and its spatial lag in JJA” (Page 10, Line 209)

7. Line 320: The authors showed the spatial extents of clusters vary by season. Suggest to give more explanations/reasons for this.

Thank you very much for the insightful comment. The second reviewer also suggests to explore how the clusters vary by season. In the revision, we have added a new section and illustrated the results of the other three seasons:

“4.5 Frequency of the case HH in SON, DJF, and MAM

Besides JJA, spatial clustering has been performed for the anomaly correlation of GCM seasonal forecasts of total precipitation in SON, DJF, and MAM. Similarly, it is observed that the anomaly correlation varies across the globe (Figures S1, S4, and S7 in the supplementary material), correlates with its spatial lag (Figures S2, S5, and S8), and exhibits significant spatial patterns (Figures S3, S6, and S9). In addition to Figures 4 and 5, the frequency of the case HH is counted for the other three seasons and shown in Figures 6 and 7.

ENSO is one of the most important drivers of global climate [Mason and Goddard, 2001; Saha et al., 2014; Bauer et al., 2015], and the CPC of NOAA has summarized the correlation between ENSO and global precipitation in different seasons (https://www.cpc.ncep.noaa.gov/products/precip/CWlink/ENSO/regressions/geplr.sht ml). In this paper, the results in Figure 6 are associated with the global effects of ENSO. In SON, the CPC shows that ENSO correlates negatively with precipitation in Eastern Australia and Southeast Asia, and positively with precipitation in part of Middle East and East Africa. From the upper part of Figure 6, it is observed that the frequency of the case HH is high in these regions. In DJF, ENSO is shown to correlate positively with precipitation in Southern North America and negatively with precipitation in Northern South America. In these two regions, the frequency of the case HH is high (middle part of Figure 6). In MAM, ENSO is illustrated to correlate negatively with precipitation in part of Southeast Asia, Eastern Brazil, and Eastern Australia. Therein, the frequency of the case HH seem to be high (lower part of Figure 6). Therefore, as previous studies found that GCMs in NMME generate skilful forecasts of ENSO [e.g., Kirtman et al., 2014; Saha et al., 2014; Zhang et al., 2017], Figure 6 suggests that the skill, as is indicated by anomaly correlation, of GCM forecasts in NMME can also be related to ENSO. In Figure 7, the percentage and cumulative percentage of the frequency of the case HH are illustrated for SON, DJF, and MAM. Similar to Figure 5, the results show the complementarity among the ten sets of forecasts.

Besides ENSO, there are other drivers of global climate. For example, North Atlantic Oscillation (NAO) and Arctic Oscillation (AO) extensively affect the climate in Europe, Asia, and North America [Hurrell et al., 2001; Ambaum et al., 2002]. Several sea surface temperature indices of the Atlantic and Indian Oceans and ENSO jointly impact the climate in Africa [Rowell, 2013]. As can be observed from Figures 4, 5, 6, and 7, there is still substantial room for improvement of seasonal precipitation forecasts for large parts of Europe, Asia, and Africa. The overall neutrally skilful precipitation forecasts in these regions can possibly be due to that GCM formulations of other climate drivers are not as effective as the formulations of ENSO. In the meantime, the difficulty of global climate forecasting due to spatially-temporally varying teleconnections between regional precipitation and global climate drivers is noted [Merryfield et al., 2013; Saha et al., 2014; Jia et al., 2015; Hudson et al., 2017; Kushnir et al., 2019].

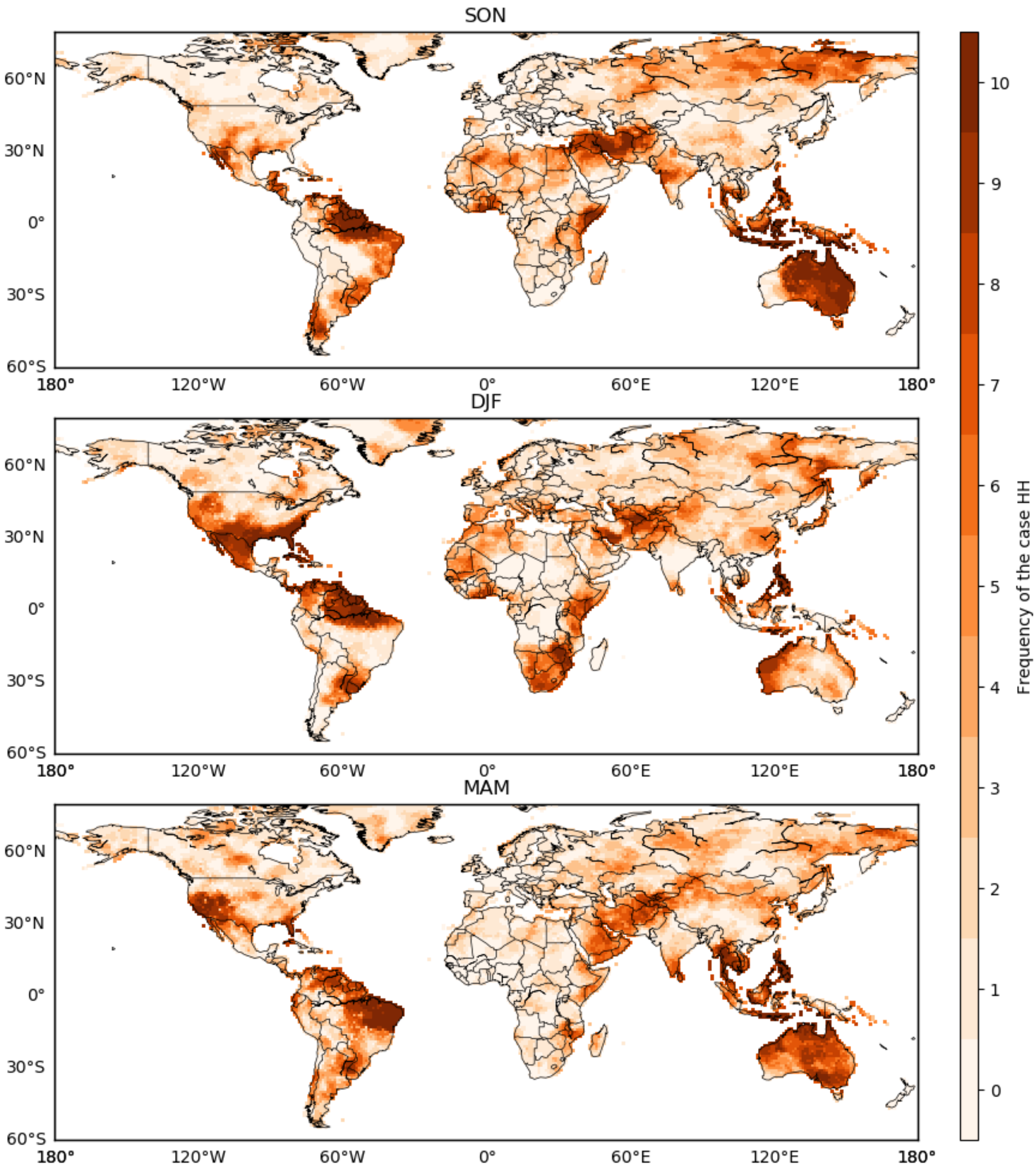


Figure 6: As for Figure 4, but for SON, DJF, and MAM

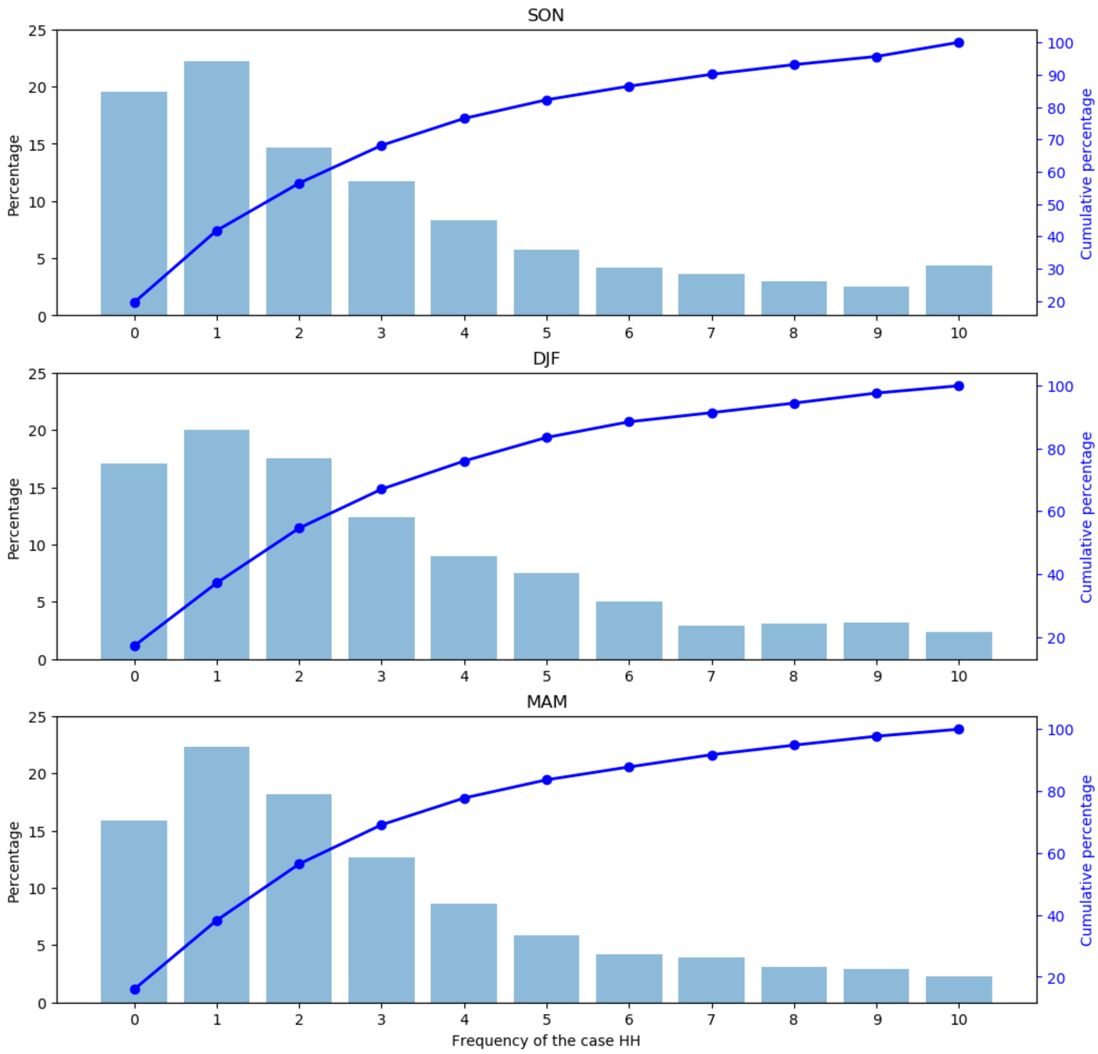


Figure 7: As for Figure 5, but for SON, DJF, and MAM”

(Pages 16 to 19, Lines 311 to 345)

Anonymous Referee #2

This study examines the spatial patterns in skill of seasonal scale dynamical precipitation forecasts from the NMME models. By examining the spatial clusters of skill the study concludes that in general the skill pattern are spatially coherent – in other words regions with higher/positive [lower/negative] skills are typically surrounded by the pixels with higher/positive [lower/negative] skills. The study also finds the spatial pattern of the forecasts from the same climate center to be similar and that use of diverse models helps improve the regions with higher skill because of complementary skills.

This is a very useful study, as the authors argue that screening of forecast skill based on spatial pattern can help identify the regions with coherent high skill and hence the regions where the forecasts are likely to be most useful for decision-making. In addition to this implication, I think spatial pattern of regions with high skills can help with the attribution of skill. For example, presumably, the regions with known ENSO teleconnection would show spatial coherence in high skill.

Thank you very much for the positive comments. We are encouraged by the comments and shall explore more in this area.

I think this study is certainly suitable for publication however I would like to suggest some additional analysis in the hopes of further improving this manuscript.

We are grateful to you for the constructive comments and have improved the paper accordingly.

(1) The figure 4 and 5 are great as they summarize several relevant information on the skill for the JJA season, I think it would be good to add similar figures for other seasons here in the main manuscript rather than in supplementary material. As of now the main manuscript has only 5 figures so the manuscript certainly has space for it.

Thank you very much for the insightful comment. We have added a new section and illustrated the results of the other three seasons:

"4.5 Frequency of the case HH in SON, DJF, and MAM

Besides JJA, spatial clustering has been performed for the anomaly correlation of GCM seasonal forecasts of total precipitation in SON, DJF, and MAM. Similarly, it is observed that the anomaly correlation varies across the globe (Figures S1, S4, and S7 in the supplementary material), correlates with its spatial lag (Figures S2, S5, and S8), and exhibits significant spatial patterns (Figures S3, S6, and S9). In addition to Figures 4 and 5, the frequency of the case HH is counted for the other three seasons and shown in Figures 6 and 7.

ENSO is one of the most important drivers of global climate [Mason and Goddard, 2001; Saha et al., 2014; Bauer et al., 2015], and the CPC of NOAA has summarized the correlation between ENSO and global precipitation in different seasons (https://www.cpc.ncep.noaa.gov/products/precip/CWlink/ENSO/regressions/geplr.shtml). In this paper, the results in Figure 6 are associated with the global effects of ENSO. In SON, the CPC shows that ENSO correlates negatively with precipitation in Eastern Australia and Southeast Asia, and positively with precipitation in part of Middle East and East Africa. From the upper part of Figure 6, it is observed that the frequency of the case HH is high in these regions. In DJF, ENSO is shown to correlate positively with precipitation in Southern North America and negatively with precipitation in Northern South America. In these two regions, the frequency of the case HH is high (middle part of Figure 6). In MAM, ENSO is illustrated to correlate negatively with precipitation in part of Southeast Asia, Eastern Brazil, and Eastern Australia. Therein, the frequency of the case HH seem to be high (lower part of Figure 6). Therefore, as previous studies found that GCMs in NMME generate skilful forecasts of ENSO [e.g., Kirtman et al., 2014; Saha et al., 2014; Zhang et al., 2017], Figure 6 suggests that the skill, as is indicated by anomaly correlation, of GCM forecasts in NMME can also be related to ENSO. In Figure 7, the percentage and cumulative percentage of the frequency of the case HH are illustrated for SON, DJF, and MAM. Similar to Figure 5, the results show the complementarity among the ten sets of forecasts.

Besides ENSO, there are other drivers of global climate. For example, North Atlantic Oscillation (NAO) and Arctic Oscillation (AO) extensively affect the climate in Europe, Asia, and North America [Hurrell et al., 2001; Ambaum et al., 2002]. Several sea surface temperature indices of the Atlantic and Indian Oceans and ENSO jointly impact the climate in Africa [Rowell, 2013]. As can be observed from Figures 4, 5, 6, and 7, there is still substantial room for improvement of seasonal precipitation forecasts for large parts of Europe, Asia, and Africa. The overall neutrally skilful precipitation forecasts in these regions can possibly be due to that GCM formulations of other climate drivers are not as effective as the formulations of ENSO. In the meantime, the difficulty of global climate forecasting due to spatially-temporally varying teleconnections between regional precipitation and global climate drivers is noted [Merryfield et al., 2013; Saha et al., 2014; Jia et al., 2015; Hudson et al., 2017; Kushnir et al., 2019].

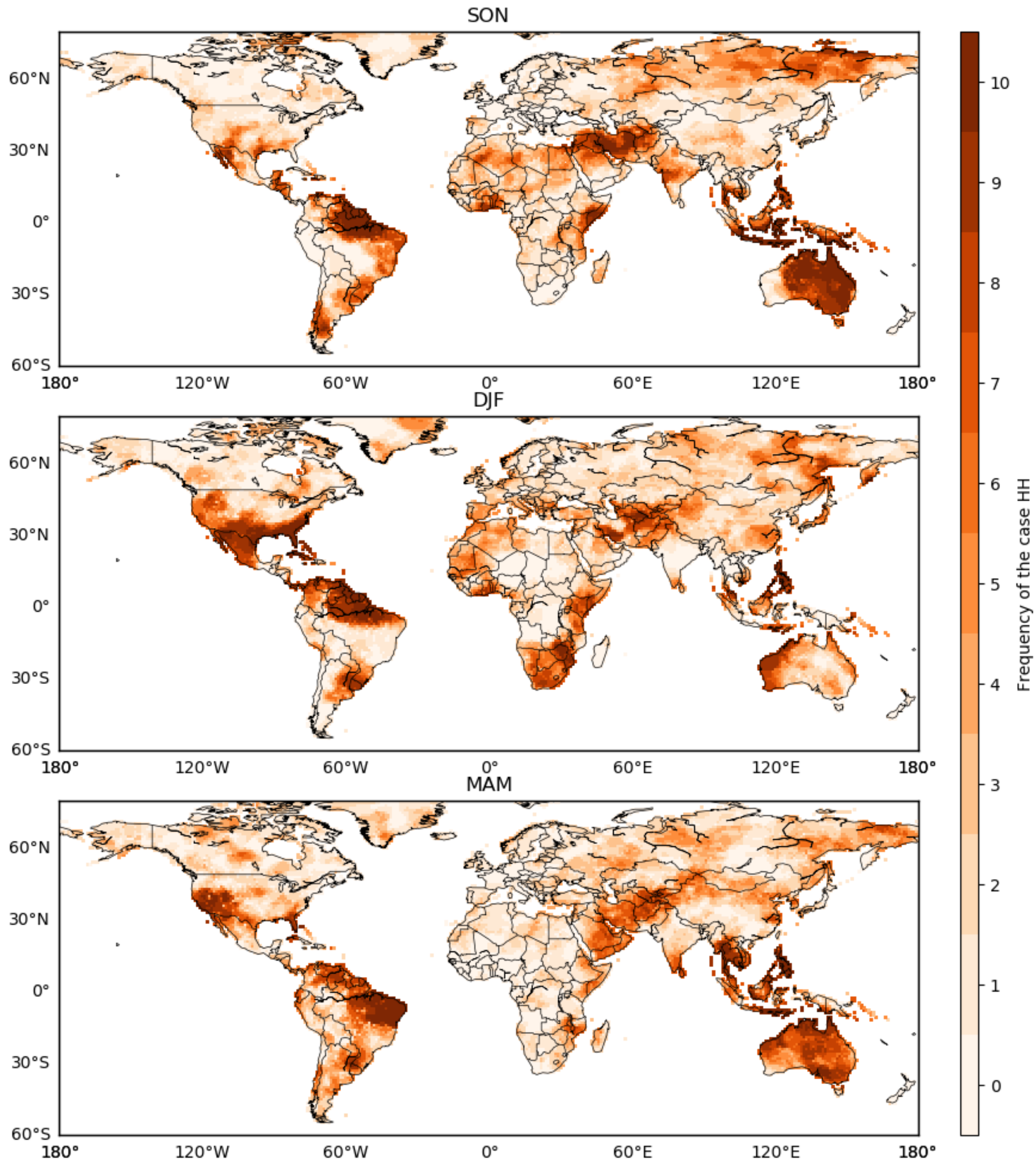


Figure 6: As for Figure 4, but for SON, DJF, and MAM

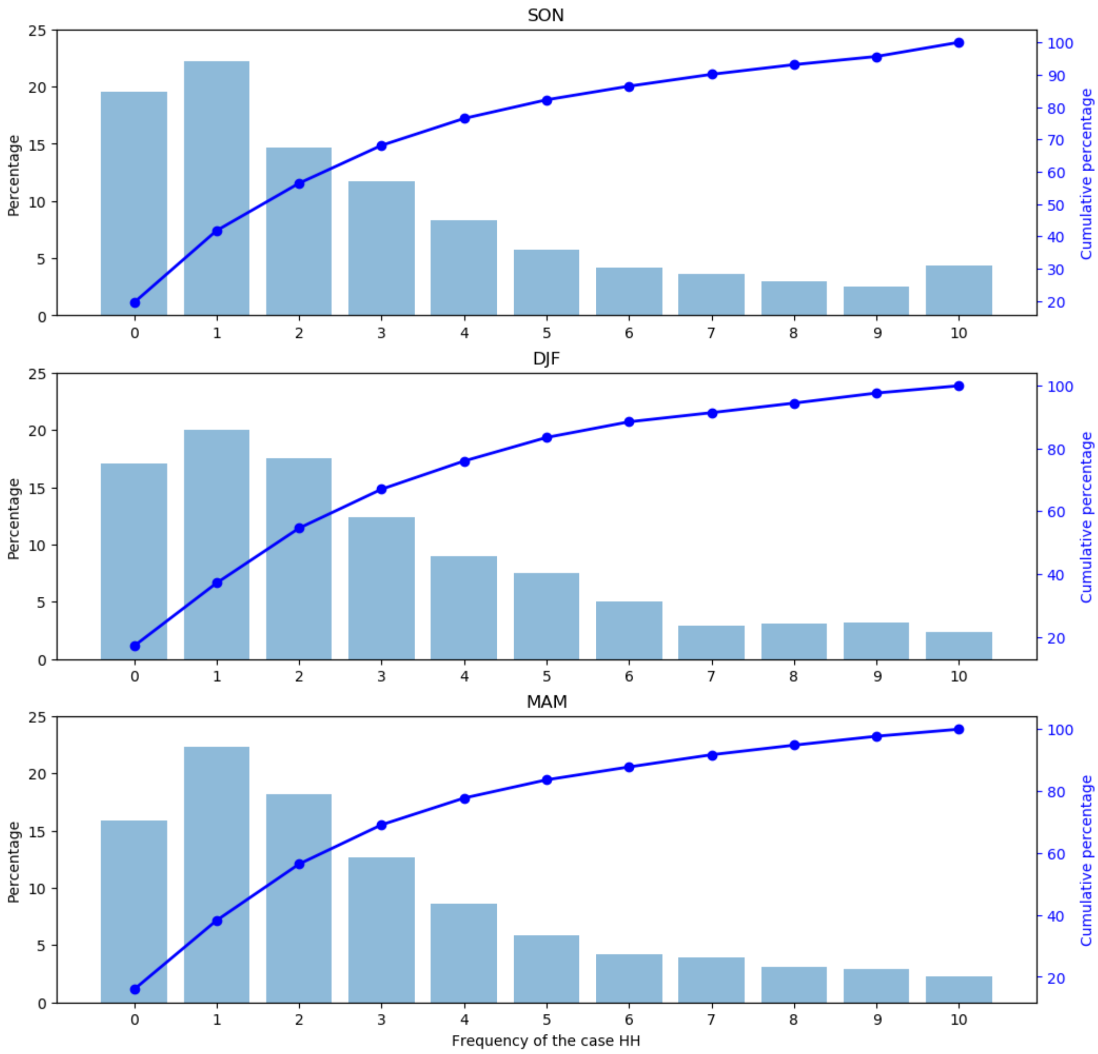


Figure 7: As for Figure 5, but for SON, DJF, and MAM”

(Pages 16 to 19, Lines 311 to 345)

(2) It would also be instructive to examine how figure 4 and 5 change with increase in the lead-time, as in general, for decision-making applications forecasts are most useful at higher lead times. It would be very interesting to see how spatially coherent high skill regions changes at higher lead times and how complementary the forecasts from different models are in increasing the overall multimodel skill. It would also be interesting to look into the attribution of the high skill spatially coherent regions. Of course there are several sources of skill but at the least, I would suggest the authors to contrast Figure 4 (for at least DJF and JJA seasons) with ENSO and precipitation teleconnection maps (maps showing correlation between the two).

Thank you. The analysis is for seasonal forecasts that add monthly forecasts at three lead times. The CPC of NOAA has investigated the global effects of ENSO and posted the plots of correlation on its website. We have contrasted the results of seasonal forecasts to the plots on CPC's website:

"ENSO is one of the most important drivers of global climate [Mason and Goddard, 2001; Saha et al., 2014; Bauer et al., 2015], and the CPC of NOAA has summarised the correlation between ENSO and global precipitation in different seasons (https://www.cpc.ncep.noaa.gov/products/precip/CWlink/ENSO/regressions/geplr.shtml). In this paper, the results in Figure 6 are associated with the global effects of ENSO. In SON, the CPC shows that ENSO correlates negatively with precipitation in Eastern Australia and Southeast Asia, and positively with precipitation in part of Middle East and East Africa. From the upper part of Figure 6, it is observed that the frequency of the case HH is high in these regions. In DJF, ENSO is shown to correlate positively with precipitation in Southern North America and negatively with precipitation in Northern South America. In these two regions, the frequency of the case HH is high (middle part of Figure 6). In MAM, ENSO is illustrated to correlate negatively with precipitation in part of Southeast Asia, Eastern Brazil, and Eastern Australia. Therein, the frequency of the case HH seem to be high (lower part of Figure 6). Therefore, as previous studies found that GCMs in NMME generate skilful forecasts of ENSO [e.g., Kirtman et al., 2014; Saha et al., 2014; Zhang et al., 2017], Figure 6 suggests that the skill, as is indicated by anomaly correlation, of GCM forecasts in NMME can also be related to ENSO. In Figure 7, the percentage and cumulative percentage of the frequency of the case HH are illustrated for SON, DJF, and MAM. Similar to Figure 5, the results show the complementarity among the ten sets of forecasts." (Page 17, Lines 317 to 330)

(3) The results of this study, I think can be used to also highlight the minimum number of models required for majority of the skill, which could be very useful for numerically expensive operational applications when impacts models (such as hydrologic or crop yield models) are driven by dynamical forecasts.

Thank you. The performances of the GCM forecasts vary from region to region. This result suggests that different sets of forecasts can have advantages/disadvantages in different regions:

"While an inter-comparison of the ten sets of GCM forecasts in terms of anomaly correlation is presented in Figure 1, the anomaly correlation exhibits considerable spatial variability that hinders the analysis across the different sets of forecasts. As a result, it is none too easy to identify regions where the forecasts persistently exhibit promising predictive performance.

The first row of Figure 1 is for the forecasts generated by two Canadian GCMs. Although CanCM3 and CanCM4 share the ocean components and have slightly different atmospheric components [Merryfield et al., 2013], their anomaly correlation shows differences. For example, in Asia and Africa, the clusters of red pixels do not seem to overlap but differ instead; and in Australia, the anomaly correlation is high in Southeast and part of Western Australia for CanCM3 while it is high in East Australia for CanCM4. These results are in accordance with a previous finding that CanCM3 and CanCM4 tend to complement each other [Merryfield et al., 2013]. The second row of Figure 1 shows the performance of two sets of forecasts by COLA-RSMAS GCMs. Complementary performance is no longer seen. Instead, CCSM4 forecasts show higher anomaly correlation and largely outperform CCSM3 forecasts in North and South America, Africa, and Australia. The outperformance can be attributed to the developments in ocean, atmospheric, and land components and the new coupling infrastructure of CCSM4 [Gent et al., 2011].

The third and fourth rows of Figure 1 are for the forecasts produced by four GFDL GCMs. In the third row, CM2p1 and CM2p1-aer04 forecasts seem to exhibit similar anomaly correlation, which tends to be high in Northeast South America, Western Africa, and Southeast Australia. In the fourth row, CM2p5-FLOR-A06 and CM2p5-FLOR-B01 forecasts show similarly high anomaly correlation in Northeast and Southeast South America, Northeast Australia and part of West Australia. On the other hand, the anomaly correlation differs from the CM2p1/CM2p1-aer04 forecasts to the CM2p5-FLOR-A06/CM2p5-FLOR-B01 forecasts. Jia et al. [2015] illustrated that CM2p5-FLOR GCMs have higher-resolution atmospheric and land components but coarser-resolution ocean components than CM2p1 GCMs. It is likely the changes in the setting of GCMs that lead to the difference in predictive performance. The fifth row of Figure 1 is for NCAR-CESM1 and NCEP-CFSv2 forecasts. Compared to CESM1 forecasts, CFSv2 forecasts tend to exhibit similar anomaly correlation in South America and show higher anomaly correlation in Asia, Africa and Australia." (Page 8, Lines 181 to 204)

(4) Lastly, I can't help but wonder how the results of HH regions over Africa would vary if a different precipitation dataset such as the Climate Hazards Center Infrared Precipitation with Stations (CHIRPS, https://chc.ucsb.edu/data/chirps) dataset used. CHIRPS ingests a lot more local in-situ observations than other precipitation datasets largely based on global in-situ databases (e.g. GHCN), and hence tends to be better

than other global datasets (see: https://www.nature.com/articles/sdata201566)

Thank you very much for the insightful comment. We agree with you on the importance of testing the results with another precipitation dataset. The CHIRPS dataset is of different spatial coverage and spatial-temporal resolution. In the revision, we have applied the Global Precipitation Climatology Centre (GPCC)'s monthly precipitation dataset to the test:

"To test whether the spatial patterns are robust, the observations of precipitation are also sourced from the Global Precipitation Climatology Centre (GPCC) [Becker et al., 2011; Schamm et al., 2014]. The anomaly correlation is re-calculated, and the spatial clustering is re-conducted. The results of GPCC precipitation, which are shown in Figures S10 to 25 in the supplementary material, are overall similar to the results of CMAP precipitation. In particular, as to the two datasets of precipitation observations, the spatial distributions of the case HH resemble in JJA (Figures 4 and S10) and also in SON, DJF, and MAM (Figures 6 and S12). This outcome highlights the existence of significant spatial patterns and confirms that the spatial clustering can serve as an effective tool to yield insights into the predictive performance of GCM forecasts." (Page 20, Lines 355 to 363)

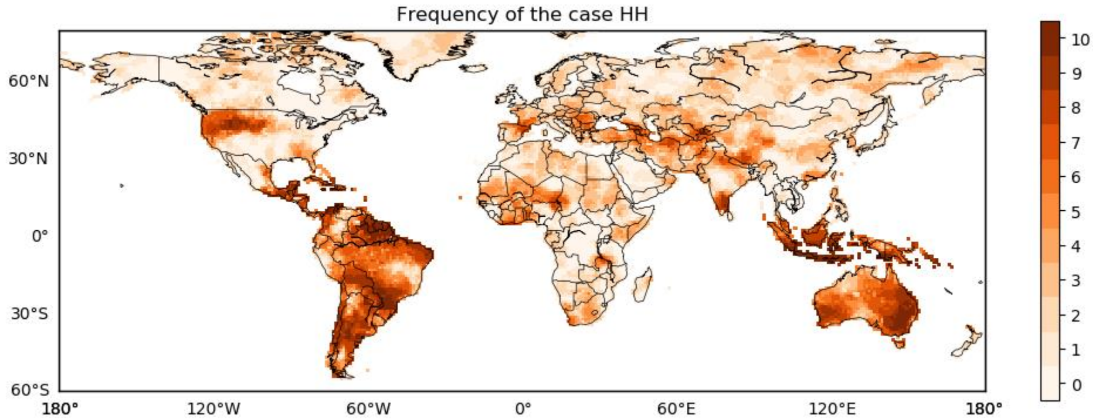


Figure 4: The spatial distribution of the frequency of the case HH across the globe for the ten sets of GCM forecasts of the total precipitation in JJA

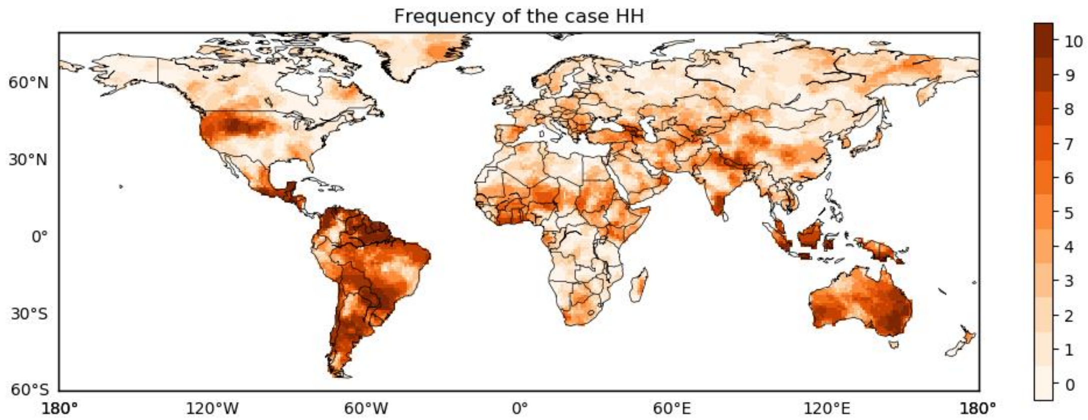


Figure S10: As for Figure 4, but for GPCC precipitation in JJA

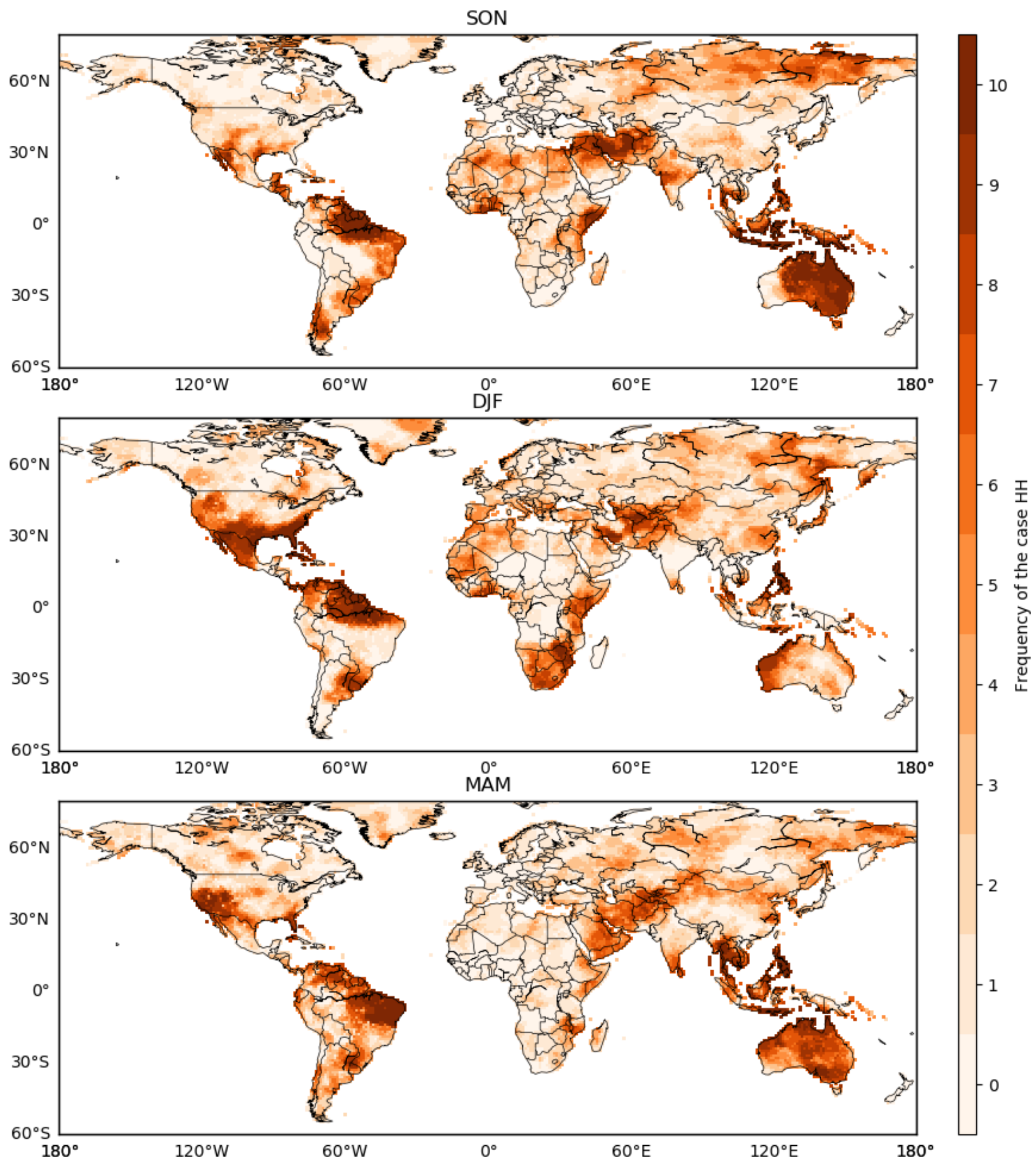


Figure 6: As for Figure 4, but for SON, DJF, and MAM

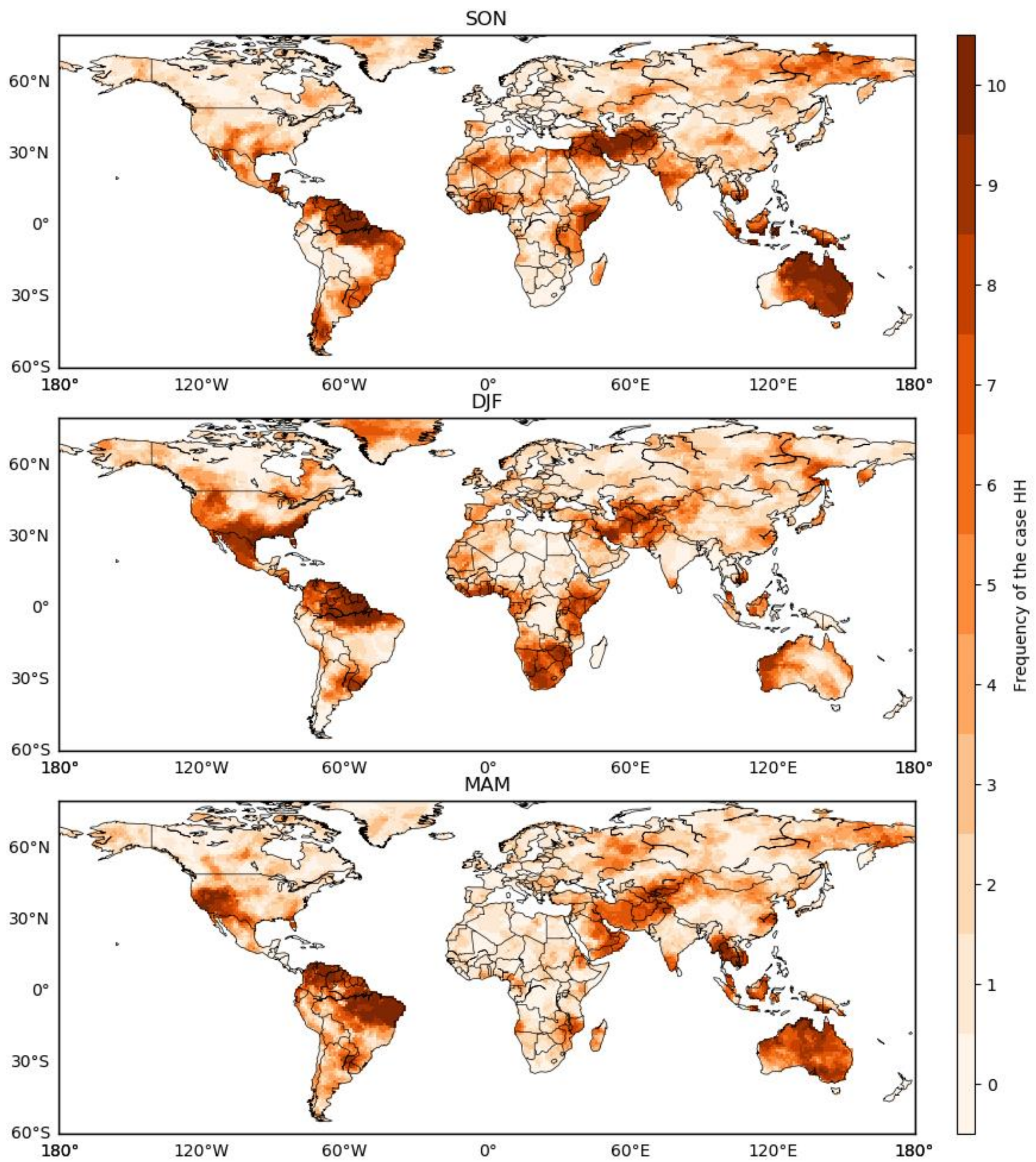


Figure S12: As for Figure 6, but for GPCC precipitation in SON, DJF, and MAM

In the meantime, the importance of tests with more datasets is noted in the discussion:

"Given that the global and local Moran's I are flexible and easy to compute, they are ready to be extended in future analysis to other datasets of forecasts, such as forecasts generated by GCMs in Europe and Asia or by regional climate models (RCMs) [Alfieri et al., 2013; Bellprat et al., 2019; Kushnir et al., 2019]. Also, the forecasts can be verified using global and regional datasets of precipitation [Funk et al., 2015; Zhao et al., 2017a, 2017b]. A more extensive investigation would contribute to better understanding of the predictive performance and illustrate the advantages of different sets of forecasts. Of particular interest is to explore which forecasts achieve promising predictive performance in large parts of Europe, Asia, and Africa. In the meantime, it is meaningful to account for the dynamics of global climate and investigate the model physics that leads to the improved performance.” (Page 20, Lines 368 to 375)

Minor comments:

(1) Abstract: Please briefly explain global and local Moran's in the abstract as well.

Thank you.

“The global Moran’s I associates anomaly correlation at neighbouring grid cells to one another and indicates that at the global scale anomaly correlation at one grid cell relates significantly and positively to anomaly correlation at surrounding grid cells. The local Moran’s I links anomaly correlation at one grid cell with its spatial lag and reveals clusters of grid cells with high, neutral, and low anomaly correlation.” (Page 1, Lines 14 to 17)

(2) Figure 2 is missing a color bar.

Thank you.

“Figure 2: Scatter plots of anomaly correlation at one grid cell against the corresponding spatial lag, i.e., spatially-weighted and -averaged anomaly correlation at surrounding grid cells. The density of points is estimated by kernel density function and shown by the viridis heatmap, with yellow (blue) colour indicating high (low) density” (Page 11, Lines 227 to 229)

(3) Figure 3: Please provide legends explaining the abbreviations HH, HL etc.

Thank you.

“Figure 3: Classification of grid cells across the globe into five cases based on spatial clustering of anomaly correlation. The case HH is marked in orange, the case HL in red, the case NS in grey, the case LH in green, and the case LL in blue. H and L are respectively short for high and low; the case HH (HL, LH, and LL) indicates that a grid cell with high (high, low, and low) anomaly correlation is surrounded by grid cells with high (low, high, and low) anomaly correlation. NS is short for not significant; the case NS means that the anomaly correlation at a grid cell or surrounding grid cells is neutral” (Page 14, Lines 269 to 273)

Significant spatial patterns from the GCM seasonal forecasts of global precipitation

Tongtiegang Zhao[1], Wei Zhang[2], Yongyong Zhang[3], Zhiyong Liu[1], and Xiaohong Chen[1]

[1]Center of Water Resources and Environment, Sun Yat-Sen University, Guangzhou 510275, China
[2]IIHR-Hydroscience & Engineering, University of Iowa, Iowa City 52242, United States
[3]Key Laboratory of Water Cycle and Related Land Surface Processes, Chinese Academy of Sciences, Beijing 100101, China

Correspondence to: Tongtiegang Zhao (zhaottg@mail.sysu.edu.cn)

Abstract: Fully-coupled global climate models (GCMs) generate a vast amount of high-dimensional forecast data of the global climate; therefore, interpreting and understanding the predictive performance is a critical issue in applying GCM forecasts. Spatial plotting is a powerful tool to identify where forecasts perform well and where forecasts are not satisfactory. Here we build upon the spatial plotting of anomaly correlation between forecast ensemble mean and observations ~~and~~ to derive significant spatial patterns to illustrate the predictive performance. For the anomaly correlation derived from the ten sets of forecasts archived in the North America Multi-Model Ensemble (NMME) experiment, the global and local Moran's I are calculated to associate anomaly correlation at neighbouring grid cells to one another. The global Moran's I associates anomaly correlation at neighbouring grid cells to one another and indicates that at the global scale anomaly correlation at one grid cell relates significantly and positively to anomaly correlation at surrounding grid cells. The local Moran's I links anomaly correlation at one grid cell with its spatial lag and reveals clusters of grid cells with high, neutral, and low anomaly correlation.~~The global Moran's I indicates that at the global scale anomaly correlation at one grid cell relates significantly and positively to anomaly correlation at surrounding grid cells, while the local Moran's I reveals clusters of grid cells with high, neutral, and low anomaly correlation.~~ Overall, the forecasts produced by GCMs of similar settings and at the same climate center exhibit similar clustering of anomaly correlation. In the meantime, the forecasts in NMME show complementary performances. About 80% of grid cells across the globe fall into the cluster of high anomaly correlation under at least one of the ten sets of forecasts. While anomaly correlation exhibits substantial spatial variability, the clustering approach serves as a filter of noise to identify spatial patterns and yields insights into the predictive performance of GCM seasonal forecasts of global precipitation.

1 Introduction

Global climate models (GCMs) have been steadily improved over the past decades and are being employed by major climate centers around the world to generate operational long-range forecasts [Doblas-Reyes et al., 2013; Saha et al., 2014; Bauer et al., 2015; Hudson et al., 2017; Kushnir et al., 2019], providing physically-based forecasts in comparison to conventional statistical forecasts [Mason and Goddard, 2001; Wu et al., 2009; Schepen et al., 2012]. In particular, the fully-coupled GCMs assimilate world-wide observational information to predict the global hydrological cycle [Merryfield et al., 2013; Saha et al., 2014; Jia et al., 2015]. Equipped with physical and dynamical laws, GCMs can potentially make forecasts of longer lead time and higher skill than statistical models [Kirtman et al., 2014; Becker et al., 2014; Chen et al., 2017]. In terms of computation, global climate forecasting is as complex as the simulation of the human brain and of the evolution of the early Universe [Bauer et al., 2015]. Advances in super-computing facilitate the forecasting and make GCM forecasts readily available for hydrological, environmental, and agricultural modelling [Sheffield et al., 2014; Vecchi et al., 2014; Bellprat et al., 2019; Pappenberger et al., 2019; Zhao et al., 2019].

GCMs generate a vast amount of high-dimensional forecast data, including retrospective forecasts of past climate and real-time forecasts [Kirtman et al., 2014; Saha et al., 2014; Jia et al., 2015]. Due to the complexity of atmospheric processes and model physics, the predictive performance of GCM forecasts is not uniform but varies considerably across the globe [Yuan et al., 2013; Tian et al., 2017; Zhao et al., 2018]. Therefore, interpreting and understanding the predictive performance is a critical issue in the applications of GCM forecasts [Doblas-Reyes et al., 2013; Merryfield et al., 2013; Saha et al., 2014; Jia et al., 2015; Hudson et al., 2017]. There are various metrics to verify the attributes of forecasts [Murphy, 1993]. For example, bias in percentage indicates the extent to which the forecasts are persistently higher, or lower, than the corresponding observations; probability integral transform (PIT) evaluates the reliability of the spread of ensemble GCM forecasts in capturing the distribution of observations; and the continuous ranked probability score (CRPS) is a probability-weighted measure of the errors of ensemble members in relation to the observations [Murphy, 1993; Hersbach, 2000; Gneiting et al., 2007]. The anomaly correlation that indicates how well large (small) values of forecasts correspond to large (small) values of observations is one of the most popular metrics [e.g., Yuan et al., 2011; Saha et al., 2014; Crochemore et al., 2016; Hudson et al., 2017; Zhao et al., 2017a]. Compared to PIT that requires a diagnostic plot and CRPS that relies on numerical integration, anomaly correlation is conceptually simple, easy to implement, and also robust to missing and censored values [Yuan et al., 2011; Luo et al., 2014; Slater et al., 2017].

Spatial plotting with latitude and longitude has been extensively used to handle the dimensionality for the verification of GCM forecasts [Kirtman et al., 2014; Hudson et al., 2017; Slater et al., 2017]. The fact that forecasts are commonly generated by GCMs as grid-based data makes spatial plotting a particular tool of choice for verification [Merryfield et al., 2013; Saha et al., 2014; Jia et al., 2015]. As to anomaly correlation, spatial plotting overcomes tedious eyeball search by grid cell and is effective in locating where there is a good correspondence between forecasts and observations and where the correspondence is not satisfactory [~~Yuan et al., 2011;~~ Luo et al., 2013; Saha et al., 2014; ~~Sheffield et al., 2014;~~ Crochemore et al., 2016; Zhao et al., 2018, 2019b]. Similarly, spatial plotting applies to other verification metrics, such as bias and CRPS, and facilitates the examination of forecast attributes [Hersbach, 2000; Gneiting et al., 2007; Kirtman et al., 2014].

The extensive use of spatial plotting underlines the importance of testing the significance of spatial patterns. In spatial statistics, one of the fundamental issues is “are the spatial patterns displayed by the spatial plots significant in some sense and therefore worth interpreting?” [Cliff and Ord, 1981; Anselin, 1995; Getis, 2007]. However, the test of significance is commonly missing in the spatial plotting of GCM forecasts. In other words, verification metrics, such as anomaly correlation, are calculated for each grid cell and then shown as they are. To some extent, the interpretation of predictive performance depends on the color schemes, which are selected subjectively to represent the scale of verification metrics. There is the first law of geography – “everything is related to everything else, but near things are more related than distant things” [Tobler, 1970]. As to spatial plotting, the indication is that when verifying forecasts at one grid cell, attention also needs to be paid to forecasts at surrounding grid cells. For anomaly correlation, a grid cell with high correlation between forecasts and observations can be surrounded by grid cells with similarly high correlation, or by grid cells with low correlation. In the former case, the grid cell is located in a region where the GCM forecasts tend to perform well. But in the latter case, the high correlation can be a suspicious outlier. Moreover, previous studies observed ~~there exist~~ clusters of grid cells with negative anomaly correlation, ~~which means~~i.e., large (small) values of forecasts correspond to small (high) values of observations [Zhao et al., 2017b, 2018, 2019b]. In such a case, forecasts are cautiously wrong. Therefore, it is critical to characterize the different cases ~~for~~ in spatial plotting and test whether the spatial patterns are significant and worth further attention.

In this paper, we are motivated to introduce spatial statistics [e.g., Di Luzio et al., 2008; Lu and Wong, 2008; Woldemeskel et al., 2013] to investigate the spatial plotting of anomaly correlation at the global scale. As will be shown later in this paper, the technique of spatial clustering facilitates the identification of significant patterns of high, neutral, and low anomaly correlation and provides an objective approach to interpreting the predictive performance of GCM forecasts. For the purpose of inter-comparison, the examination of significant patterns in spatial plotting has been conducted for ten sets of GCM seasonal precipitation forecasts in the North American Multi-Model Ensemble (NMME) experiment [Kirtman et al., 2014; Ma et al., 2016; Zhang et al., 2017]. In the remainder of the paper, the dataset of GCM seasonal forecasts is illustrated in Section 2; the spatial clustering using global and local Moran’s I is detailed in Section 3; the results of anomaly correlation at the global scale and its clustering are shown in Section 4; the discussion and conclusions are respectively presented in Sections 5 and 6.

2 Data description

The NMME builds on existing GCMs in North America to provide quality-controlled forecast data to the community of climate research and applications. More than ten sets of GCM precipitation forecasts have been spatially regridded and temporally aggregated to form a consistent dataset [Kirtman et al., 2014]. Each set of forecasts overall has 5 dimensions. They are 1) start time s, when the forecasts are initialised; 2) lead time l, whose unit is month for the forecasts; 3) ensemble member n, which is meant to represent forecast uncertainty; 4) latitude y; and 5) longitude x. Taking the precipitation forecasts of the Climate Forecast System version 2 [CFSv2, Saha et al., 2014] in NMME as an example, s is the beginning of each month and its value represents the number of months since January 1960; l is 0, 1, …, 9, i.e., the forecasts are for month 0 head (current month), month 1 ahead, …, and month 9 ahead; n is numbered from 1 to 24, i.e., 24 ensemble members; y is from -90 to 90 while x is from 0 to 359, i.e., the spatial resolution is 1 degree by 1 degree (approximately 100 kilometres). In the meantime, NMME provides precipitation observations corresponding to the forecasts. Specifically, the Climate

Prediction Center (CPC)'s merged analysis of precipitation [CMAP; Xie and Arkin, 1997; Xie et al., 2007], which is monthly, has been regridded to 1-degree resolution to verify GCM forecasts [Kirtman et al., 2014; Chen et al., 2017; Zhao et al., 2018].

Table 1: Basic information on the ten sets of GCM forecasts from the NMME experiment

Climate Centre	GCM	Number of ensemble members	Lead time (month)
Canadian Meteorological Center (CMC)	Canadian coupled model version 3 (CanCM3)	10	0–11
	Canadian coupled model version 4 (CanCM4)	10	0–11
Center for Ocean-Land-Atmosphere Studies, Rosenstiel School of Marine and Atmospheric Science (COLA-RSMAS)	Community climate system model version 3 (CCSM3)	6	0–11
	Community climate system model version 4 (CCSM4)	10	0–11
Geophysical Fluid Dynamics Laboratory (GFDL)	Climate model version 2.1 (CM2p1)	10	0–11
	Climate model version 2.1 (CM2p1-aer04)	10	0–11
	Climate model version 2.5 with forecast-oriented low ocean resolution (CM2p5-FLOR-A06)	12	0–11
	Climate model version 2.5 with forecast-oriented low ocean resolution (CM2p5-FLOR-B01)	12	0–11
National Center for Atmospheric Research (NCAR)	Community earth system model version 1 (CESM1)	10	0–11
National Centers for Environmental Prediction (NCEP)	Climate forecast system version 2 (CFSv2)	24	0–9

Ten sets of precipitation forecasts, as well as CMAP observations, in the NMME are downloaded from the International Research Institute at the Columbia University (https://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/). Their retrospective forecasts are complete in the period from 1982 to 2010~~, and their real-time forecasts are updated by month since January 2010~~ [Merryfield et al., 2013; Saha et al., 2014; Jia et al., 2015]. In the meantime, their real-time forecasts are updated periodically in a slightly different setting; for example, CFSv2 forecasts are generated since January 2011 using initial conditions of the last 30 days, with 4 runs from each day (https://www.cpc.ncep.noaa.gov/products/CFSv2/CFSv2_body.html). Basic information on the forecasts are provided in Table 1. In the analysis, the attention is paid to the retrospective forecasts~~:~~

$$F_{GCM} = \left[f_{s,l,n,y,x} \right]_{GCM} \quad (1)$$

In Eq. (1), f represents forecast values that are specified by the 5 dimensions; F, which is the set of forecasts, is marked by the GCM that generates the forecasts. It is noted that in NMME, F_{GCM} are raw forecasts generated by GCMs and are not bias-corrected or downscaled.

The observed precipitation corresponding to the forecasts is denoted as

$$O = \left[o_{t,y,x} \right] \quad (t = s + l) \tag{2}$$

As shown in Eq. (2), the observation in total has three dimensions: time t, whose value is the addition of lead time l to start time s in the alignment of observations with forecasts; latitude y; and longitude x. It is pointed out that while F differs by GCM, O is the same across the ten sets of forecasts.

The start time s in Eqs. (1) and (2) comprises year k, i.e., 1982, 1983, …, 2010, and month m, i.e., January, February, …, and December. The predictive performance of GCM forecasts exhibits seasonality [Yuan et al., 2011; Zhao et al., 2017a, 2017b]. Accordingly, in the analysis, forecasts are selected by fixing m while varying k, e.g., pooling forecasts initialised in January 1982, January 1983, …, January 2010. The anomaly correlation is calculated by relating forecasts to the corresponding observations

$$r = \frac{\sum_k \left(rf_k - \overline{rf} \right)\left(ro_k - \overline{ro} \right)}{\sqrt{\sum_k \left(rf_k - \overline{rf} \right)^2}\sqrt{\sum_k \left(ro_k - \overline{ro} \right)^2}} \tag{3}$$

The above formulation focuses on k and omits other dimensions, including m, l, y and x, for the sake of simplicity. In Eq. (3), rf_k (ro_k) is the rank of year k's forecast ensemble mean (observation) in the 29 years' ensemble mean (observations); and $\overline{rf}$ ($\overline{ro}$) is the mean value of rf_k (ro_k). In general, the anomaly correlation characterises how well large (small) values of ensemble mean correspond to large (small) values of observations. Good (poor) correspondence makes r tend towards 1 (–1). With Eq. (3), ~~T~~the set of anomaly correlation between F_{GCM} and O is ~~obtained~~evaluated~~:~~

$$R_{GCM} = \left[r_{m,l,y,x} \right]_{GCM} \tag{~~3~~4}$$

In which r and R are respectively the correlation coefficients and the set of correlation. R, which differs by GCM, has four dimensions: 1) ~~a new~~month~~ dimension~~ m~~, which is the month when the forecasts are generated, is created to~~ , which ~~substitute~~ substitutes start time s~~.~~ in Eq. (1)~~It is owing to that the calculation of anomaly correlation pools forecasts at multiple start times, e.g., January 1982, January 1983, …, January 2010, since the performance of forecasts exhibits seasonality [Luo et al., 2014; Ma et al., 2016; Slater et al., 2017]~~; 2) lead time l; 3) latitude y; and 4) longitude x. Comparing Eq. (~~3~~4) to Eq. (1), the dimension n of ensemble member is eliminated since the forecast- ensemble mean ~~mean~~ is taken in the calculation of anomaly correlation (Eq. 3).

域代码已更改

域代码已更改

For selected GCM forecasts in month m and at lead time l, the anomaly correlation between ensemble mean and observation forms a two-dimensional array by latitude and longitude. Here, spatial plotting applies to the presentation of anomaly correlation at the global scale, and also at the regional scale. Following Eq. (~~3~~4), the set of anomaly correlation is denoted as

$$R_{GCM,m,l} = \left[r_{y,x} \right] \qquad \text{~~(4)~~(5)}$$

In Eq. (~~4~~5), y and x specify the location of grid cells. Denoting grid cell as i, the subscripts of latitude y and longitude x are merged into i for the purpose of simplicity

$$R_{GCM,m,l} = \left[r_i \right] \qquad \text{~~(5)~~(6)}$$

In which r_i represents the anomaly correlation at grid cell i, of which the latitude is y_i and the longitude is x_i.

3 Methods

The spatial plotting employs certain pre-selected colour schemes to represent the value of anomaly correlation and show the grid cell-wise anomaly correlation as it is [e.g., Yuan et al., 2011; Kirtman et al., 2014; Ma et al., 2016]. Spatial patterns that represent clusters of grid cells with high anomaly correlation ~~are~~ have been observed and highlighted in peer studies [e.g., Saha et al., 2014; Jia et al., 2015; Slater et al., 2017]. The spatial clustering associates anomaly correlation at neighbouring grid cells to one another and tests the significance of the patterns by random permutation. Following the standard formulations of spatial statistics [Anselin, 1995, 2006; Rey and Anselin, 2010], the global Moran's I is calculated to examine the association among anomaly correlation at the global scale

$$I = \frac{\dfrac{1}{\sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} w_{i,j}} \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} w_{i,j} \left(r_i - \bar{r} \right) \left(r_j - \bar{r} \right)}{\dfrac{1}{N} \sum_{i=1}^{N} \left(r_i - \bar{r} \right)^2} \qquad \text{~~(6)~~(7)}$$

In which N is the number of grid cells indexed by i and j across the globe; $\bar{r}$ is the mean value of anomaly correlation; and $w_{i,j}$ is the spatial weighting coefficient that usually decays with the distance between i and j [Miller, 2004; Hao et al., 2016; Schmal et al., 2017]. At the right-hand side of Eq. (~~6~~7), the denominator is the variance of r_i across all the grid cells; and the numerator is the spatially-weighted and -averaged covariance between r_i and r_j. Generally, the value of the global Moran's I ranges from -1 to 1. The similarity (dissimilarity) of r_i to the surrounding r_j makes I tend toward 1 (-1), while the random distribution of anomaly correlation makes I close to 0.

The spatial weight $w_{i,j}$ plays an important part in the calculation of I [Rey and Anselin, 2010]. Following the inverse distance weighting (IDW) interpolation in geosciences [Di Luzio et al., 2008; Lu and Wong, 2008; Woldemeskel et al., 2013], $w_{i,j}$ is formulated as follows

$$w_{i,j} = \frac{1}{d(i,j)^2} \qquad \text{~~(7)~~(8)}$$

In ~~Eq. (6),~~which $d(i, j)$ ~~$d(i,j)$~~ is the Euclidean distance between grid cells i and j, i.e., $d(i,j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$. In addition, the cut-off threshold for $d(i, j)$ ~~$d(i,j)$~~ is set as 10 degrees (approximately 1,000 kilometres) to reduce the computational burden. That is, $w_{i,j}$ is set as 0 if $d(i, j)$ ~~$d(i,j)$~~ exceeds 10.

Adding to the global Moran's I, the local Moran's I is obtained to test whether r_i at a certain grid cell i significantly relates to surrounding r_j at the local scale [Anselin, 2006; Hao et al., 2016; Yuan et al., 2018]

$$I_i = \frac{\left(r_i - \bar{r}\right) \dfrac{\sum_{j=1, j\neq i}^{N} w_{i,j}\left(r_j - \bar{r}\right)}{\sum_{j=1, j\neq i}^{N} w_{i,j}}}{\dfrac{1}{N}\sum_{i=1}^{N}\left(r_i - \bar{r}\right)^2} \qquad \text{~~(8)~~(9)}$$

As shown in the above formulation, I_i is positive when r_i and the surrounding r_j are similarly high, or similarly low. On the other hand, I_i is negative when a high (low) value of r_i correspond to low (high) values of neighbouring r_j. Also, I_i can be close to zero when r_i or the surrounding r_j is close to the mean value. The significance of I_i is tested by random permutations [Rey and Anselin, 2010]. For each permutation, the values of r_j are randomly rearranged, and then the local Moran's I is re-calculated. The permutations obtained a reference distribution for I_i under the null hypothesis of randomly distributed anomaly correlation [Anselin, 1995, 2006; Rey and Anselin, 2010]. Given a significance level α, the quantiles $I_{\alpha/2}$ and $I_{1-\alpha/2}$ are retrieved from the reference distribution. Therefore, the two-tailed test of I_i along with the anomaly correlation r_i facilitates spatial clustering and derives five cases:

$$case_i = \begin{cases} HH & (I_i > I_{1-\alpha/2}) \cup (r_i > \bar{r}) \\ HL & (I_i < I_{\alpha/2}) \cup (r_i > \bar{r}) \\ NS & (I_{\alpha/2} \le I_i \le I_{1-\alpha/2}) \\ LH & (I_i < I_{\alpha/2}) \cup (r_i < \bar{r}) \\ LL & (I_i > I_{1-\alpha/2}) \cup (r_i < \bar{r}) \end{cases}$$ ~~(9)~~(10)

As illustrated in Eq. (9), the first case HH, which is short for high-high, indicates that a high value of r_i is surrounded by high values of r_j; the second case is HL – ~~, i.e.,~~ high-low – a high value of r_i surrounded by low values of r_j; the third case is NS – ~~, i.e.,~~ not significant – the local association of r_i with surrounding r_j is not significant; the fourth case is LH – ~~, i.e.,~~ low-high – a low value of r_i surrounded by high values of r_j; and the fifth case is LL – ~~, i.e.,~~ low-low – a low value of r_i surrounded by low values of r_j. In this way, the significance of patterns, which generally represent clusters of grid cells with high (low) anomaly correlation, is examined for spatial plotting of anomaly correlation. α is set to be 0.05 in this paper.

4 Results

The spatial clustering is performed for the anomaly correlation ~~of~~ across the ten sets of ~~GCM seasonal precipitation~~ forecasts in NMME. In the analysis, the attention is mainly ~~concentrated on~~paid to ~~the forecasts of total precipitation in~~ June, July, and August (JJA), which are generally boreal summer and Austral winter. That is, ~~The~~ the start time of the forecasts is June, and the forecasts at the lead times of 0, 1, and 2 months are aggregated to form the seasonal forecasts. In the meantime, forecasts initialized in September of total precipitation in September, October, and November (SON), forecasts initialized in December of total precipitation in December, (the next) January, and (the next February) (DJF), and forecasts initialized in March of total precipitation in March, April, and May (MAM) are also investigated, with the results presented in the supplementary material.

4.1 Anomaly correlation ~~of GCM forecasts~~in JJA

The anomaly correlation between ensemble mean and observation is evaluated for the ten sets of ~~GCM~~ seasonal precipitation forecasts ~~in NMME~~. In Figure 1, the spatial plots employ a diverging red-blue colour scheme to represent the value of anomaly correlation. Red pixels indicate positive correlation, while blue pixels negative correlation. For each set of forecasts, many instances of red pixels can be observed. That is, forecasts exhibit promising performance with ensemble mean positively correlated with observation in many instances. Meanwhile, there also exist instances of blue pixels. In those instances, forecasts are generally not right because large (small) values of ensemble mean coincide with small (large) values of observation. While an inter-comparison of the ten sets of GCM forecasts in terms of anomaly correlation is presented in Figure 1, the anomaly correlation exhibits considerable spatial variability that hinders the analysis across the ~~the multiple~~

different sets of forecasts. As a result, it is none too easy to identify regions where the forecasts persistently exhibit promising predictive performance.

The first row of Figure 1 is for the forecasts generated by two Canadian GCMs. Although CanCM3 and CanCM4 share the ocean components and have slightly different atmospheric components [Merryfield et al., 2013], their anomaly correlation shows differences. For example, in Asia and Africa, the clusters of red pixels do not seem to overlap but differ instead; and in Australia, the anomaly correlation is high in Southeast and part of Western Australia for CanCM3 while it is high in East Australia for CanCM4. These results are in accordance with a previous finding that CanCM3 and CanCM4 tend to complement each other [Merryfield et al., 2013]. The second row of Figure 1 shows the performance of two sets of forecasts by COLA-RSMAS GCMs. Complementary performance is no longer seen. Instead, CCSM4 forecasts show higher anomaly correlation and largely outperform CCSM3 forecasts in North and South America, Africa, and Australia. The outperformance can be attributed to the developments in ocean, atmospheric, and land components and the new coupling infrastructure of CCSM4 [Gent et al., 2011].

The third and fourth rows of Figure 1 are for the forecasts produced by four GFDL GCMs. In the third row, CM2p1 and CM2p1-aer04 forecasts seem to exhibit similar anomaly correlation, which tends to be high in Northeast South America, Western Africa, and Southeast Australia. In the fourth row, CM2p5-FLOR-A06 and CM2p5-FLOR-B01 forecasts show similarly high anomaly correlation in Northeast and Southeast South America, Northeast Australia and part of West Australia. On the other hand, the anomaly correlation differs from the CM2p1/CM2p1-aer04 forecasts to the CM2p5-FLOR-A06/CM2p5-FLOR-B01 forecasts. Jia et al. [2015] illustrated that CM2p5-FLOR GCMs have higher-resolution atmospheric and land components but coarser-resolution ocean components than CM2p1 GCMs. It is likely the ~~modifications~~ changes ~~to~~ in the setting of GCMs that lead to the difference in predictive performance. The fifth row of Figure 1 is for NCAR-CESM1 and NCEP-CFSv2 forecasts. Compared to CESM1 forecasts, CFSv2 forecasts tend to exhibit similar anomaly correlation in South America and show higher anomaly correlation in Asia, Africa and Australia.

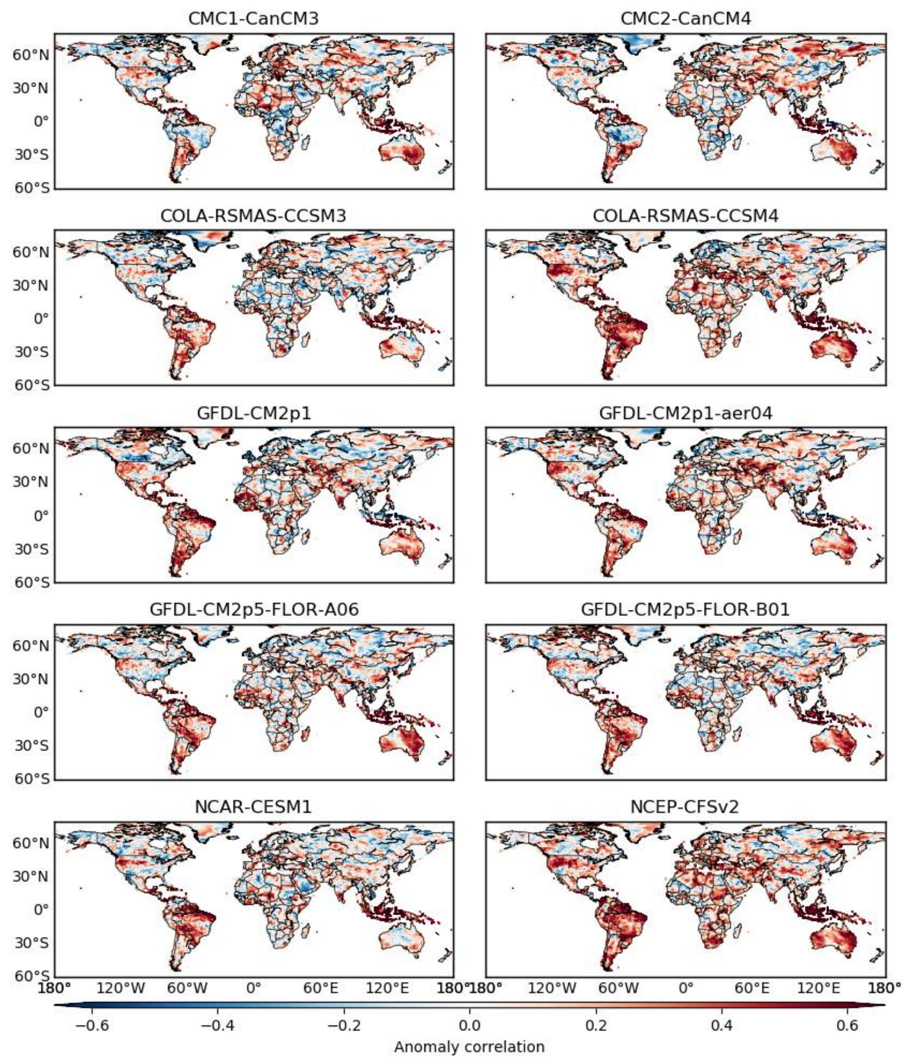
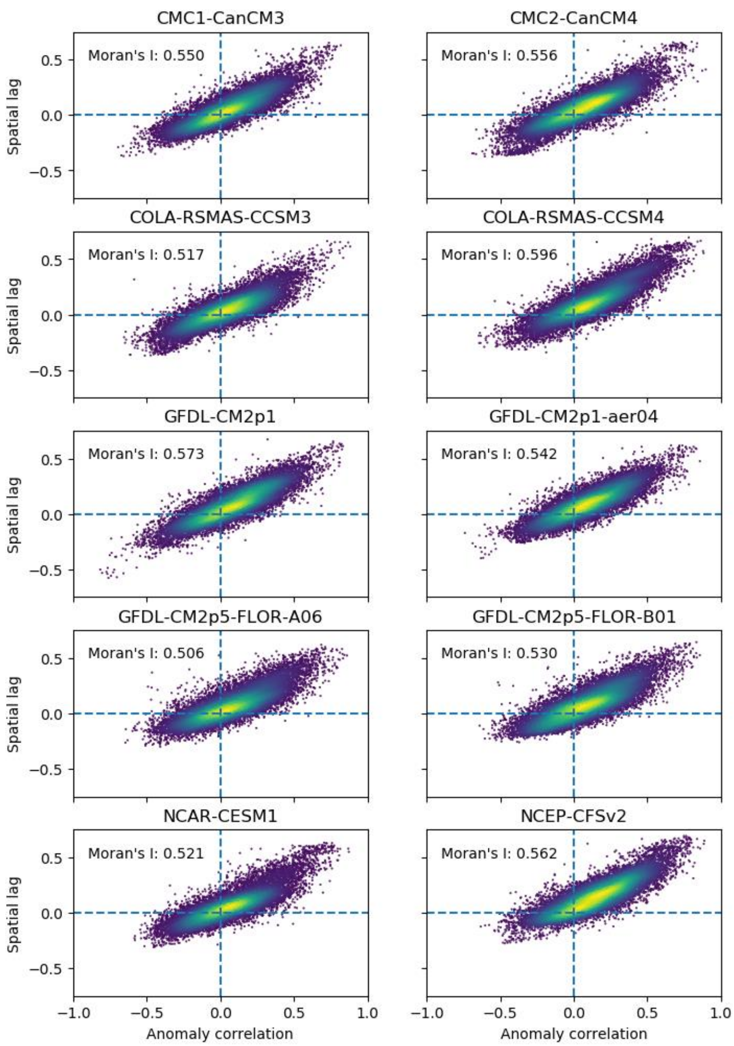


Figure ~~2~~1: Anomaly correlation between ensemble mean and observation for ten sets of GCM forecasts of seasonal precipitation. The forecasts are initialized in June and are for the total precipitation in June, July, and August

4.2 Anomaly correlation ~~of GCM forecasts~~and its spatial lag in JJA

In spatial analysis, one critical issue is how an attribute at one location relates to the attribute at neighbouring locations [Cliff and Ord, 1981; Anselin, 1995; Getis, 2007]. For anomaly correlation, the subplots of Figure 1 imply the existence of some relationships as there are clusters of red pixels and of blue pixels. As for the clusters, Figure 2 presents a statistical ~~thorough~~ test of the relationships using the global Moran's I. Specifically, for all the grid cells across the globe, the anomaly correlation at each grid cell is plotted against the spatially-weighted and -averaged anomaly correlation, i.e., spatial lag [Miller, 2004; Hao et al., 2016; Schmal et al., 2017], at the surrounding grid cells.

Figure 2 uses a viridis heatmap to indicate the density of scatter points. It can be observed that the points frequently fall in the first quadrant under all the ten sets of forecasts. In accordance with clusters of red pixels in Figure 1, this result suggests that many grid cells are with positive anomaly correlation and that they tend to be surrounded by grid cells with positive anomaly correlation. Meanwhile, some points are in the third quadrant. It is due to that some grid cells are of negative anomaly correlation and are surrounded by grid cells with negative anomaly correlation. This outcome corresponds to the existence of clusters of blue grid cells in Figure 1. Also, there are points in the second and fourth quadrants. Overall, anomaly correlation at one grid cell positively relates to anomaly correlation at the neighbouring grid cells. The global Moran's I is above 0.500, with the p-value far smaller than 0.01, for all the ten sets of NMME seasonal forecasts. ~~Adding to empirical findings from spatial plots~~Therefore, it is statistically verified that ~~the spatial patters implied by clusters of red (blue) pixels in Figure 1 are significant. That is,~~ at the global scale, a grid cell with high (neutral, or low) anomaly correlation tends to be surrounded by grid cells with high (neutral, or low) anomaly correlation.



240 **Figure 2: Scatter plots of anomaly correlation at one grid cell against the corresponding spatial lag, i.e., spatially-weighted and -averaged anomaly correlation at surrounding grid cells. The density of points is estimated by kernel density function and shown by the viridis heatmap, with yellow (blue) colour indicating high (low) density**

4.3 ~~Classification of grid cells~~Spatial clustering ~~by spatial analysis~~in JJA

Furthermore, the local Moran's I classifies the grid cells across the globe into 5 cases under each of the ten sets of forecasts. In Figure 3, the five cases are marked by different colours. Specifically, the case HH is in orange, the case HL in red, the case NS in grey, the case LH in green, and the case LL in blue. A prominent finding is that the three cases of HH, NS, and LL have more instances than the other two cases of HL and LH. This result agrees to the spatial clustering of anomaly correlation in Figure 1 and to the distribution of scatter points in Figure 2. Comparing Figure 3 to Figure 1, it can be observed that orange regions generally correspond to clusters of red pixels,~~ i.e., ~~ which represent positive anomaly correlation, and that blue regions coincide with clusters of blue pixels, ~~i.e., ~~which show negative anomaly correlation. In the meantime, in-between orange and blue regions are grey regions. The implication is that regions with high and low anomaly correlation tend to be separated by regions with neutral anomaly correlation. While~~ in Figure 1~~ the spatial variability of anomaly correlation in Figure 1 is complicates the analysis of predictive performance, the classification in Figure 3 facilitates ~~a more ~~effective inter-comparison of the ten sets of GCM forecasts.

The orange regions that correspond to clusters of grid cells with high anomaly correlation are of particular interest. Three findings are made from the spatial extent of orange regions. First of all, they tend to be similar under forecasts generated by the same climate center. For example, orange regions exist in a large part of South America for ~~all the~~the ten sets of forecasts. Meanwhile, ~~orange regions~~they seem missing in the Amazon Basin under the CMC1-CanCM3 and CMC2-CanCM4 forecasts while they tend to cover Amazon under the COLA-RSMAS-CCSM3 and COLA-RSMAS-CCSM4 forecasts. The similarity versus difference can be owing to that GCMs developed at the same climate center tend to share certain components [Gent et al., 2011; Merryfield et al., 2013; Jia et al., 2015]. Secondly, orange regions seem to be affected by the setting of GCMs. There are four sets of forecasts by GFDL. In Western United States, orange regions are extensive under the GFDL-CM2p1 and GFDL-CM2p1-aer04 forecasts but tend to be limited under the GFDL-CM2p5-FLOR-A06 and GFDL-CM2p5-FLOR-B01 forecasts. This drastic difference can be due to the setting of FLOR, i.e., forecast-oriented low ocean resolution [Vecchi et al. 2014; Jia et al., 2015]. Thirdly, there are substantial regional variations possibly due to the predictability of seasonal precipitation [Doblas-Reyes et al., 2013; Becker et al., 2014; Zhang et al., 2017]. For example, orange regions cover large part of Australia, in particular Southwest and Southeast Australia. However, they are not as extensive in Europe, Asia and Africa. It is possibly owing to that the climate in Australia is strongly affected by ENSO ~~[Schepen et al., 2012; Hudson et al., 2017]~~ and that the 10 GCMs in NMME tend to capture the effect of ENSO on June-July-August precipitation [Schepen et al., 2012; Wang et al., 2012; Hudson et al., 2017].

The blue regions correspond to clusters of grid cells with low anomaly correlation. They are generally indicative of locations where forecasts are not satisfactory. Under the ten sets of forecasts, blue regions can be observed in large parts of Europe, Asia, Africa, Canada, and Eastern United States. While orange regions show some relationships with the source and setting of GCMs, blue regions are more varying. In addition, they tend to be mixed with grey regions, which are indicative of neutral anomaly correlation, and also with red and green regions. ~~Overall~~Generally, this outcome ~~reflects~~ implies the difficulty of climate forecasting at the global scale as there are complex land-ocean-atmosphere processes [Bauer et al., 2015; Kapnick et al., 2018; Kushnir et al., 2019]. It is noted that some red regions that represent the case HL are observed to be located inside blue regions. The implication is that some grid cells may happen to exhibit high anomaly correlation but their surrounding grid cells are of low anomaly correlation. From the perspective of spatial statistics, the high correlation is not trustworthy and can be ~~regarded as~~ outlier.

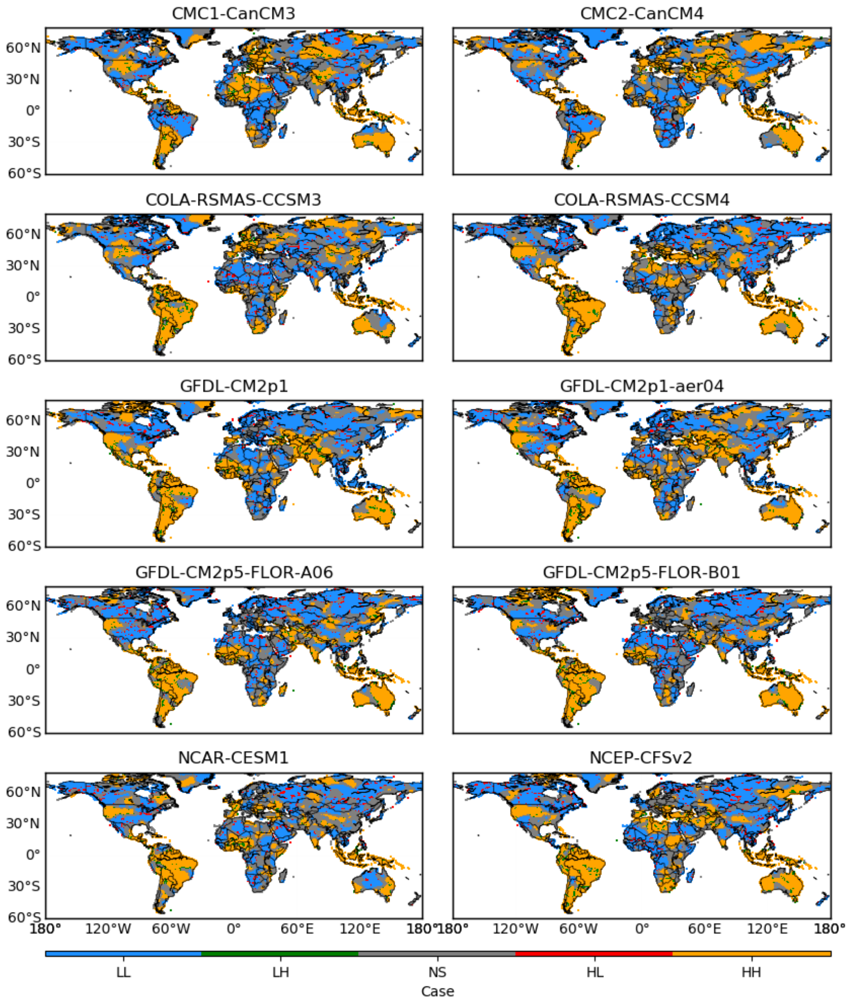


Figure 3: Classification of grid cells across the globe into five cases ~~using~~ based on spatial clustering of anomaly correlation. The case HH is marked by orange, the case HL by red, the case NS by grey, the case LH by green, and the case LL by blue. H and L are respectively short for high and low; the case HH (HL, LH, and LL) indicates that a grid cell with high (high, low, and low) anomaly correlation is surrounded by grid cells with high (low, high, and low) anomaly correlation. NS is short for not significant; the case NS means that the anomaly correlation at a grid cell and surrounding grid cells is neutral

4.4 Frequency of the case HH ~~across the forecasts~~in JJA

While the orange regions of the case HH are indicative of promising predictive performance, grid cells classified ~~under~~ as this case differ across the ten sets of forecasts. To handle the differences, the frequency that a grid cell falls into orange regions is counted for Figure 3. ~~Figure 4 illustrates the frequency using a sequential colour scheme.~~ For one grid cell, the frequency ranges from 0 to 10. That is, across the 10 sets of forecasts, one grid cell is with high anomaly correlation and is surrounded by grid cells with high anomaly correlation at the minimum for 0 times and at the maximum for 10 times. Figures 4 and 5 illustrate the spatial and statistical distributions of the frequency, respectively.

Substantial regional variation ~~is~~ can be observed for the frequency of the case HH from Figure 4. In North America, the frequency is evidently higher in Western United States than in Eastern United States, Canada, and Mexico. Globally, the frequency is higher in South America than in Europe, Asia, and Africa. Also, the frequency is high in Australia and Southeast Asia. ~~These results, which is a summary of orange regions in Figure 3, correspond to the spatial distribution of anomaly correlation in Figure 1.~~ Mason and Goddard [2001] elaborated on the relationship between ENSO and global seasonal precipitation anomalies~~.~~ ; ~~For~~ for the total precipitation in ~~June, July, and August~~JJA, El Niño was shown to coincide with above-normal precipitation in parts of South and North America and below-normal precipitation in parts of Australia and Southeast Asia~~.~~ ; ~~By~~ by contrast, the impact of El Niño is not prominent for large parts of Europe, Asia, and Africa. ~~From this perspective, i~~It is thus speculated that the results in Figure 4 to some extent reflect the impact of ENSO at the global scale. ~~Adding to previous findings that GCMs in NMME generate skilful forecasts of ENSO [e.g., Kirtman et al., 2014; Saha et al., 2014; Zhang et al., 2017], Figure 4 suggests that the GCMs also tend to capture ENSO-related precipitation.~~

~~Besides ENSO, there are other drivers of global climate. For example, North Atlantic Oscillation (NAO) and Arctic Oscillation (AO) extensively affect the climate in Europe, Asia, and North America [Hurrell et al., 2001; Ambaum et al., 2002]. Several sea surface temperature indices of the Atlantic and Indian Oceans and ENSO jointly impact the climate in Africa [Rowell, 2013]. According to Figure 4, there are still substantial rooms for improvements of seasonal precipitation forecasts in June, July, and August for large parts of Europe, Asia, and Africa. It is possibly because GCM formulations of other climate indices are not as effective as the formulations of ENSO. In the meantime, the difficulty of global climate forecasting due to spatially-temporally varying teleconnections between regional precipitation and global climate drivers is noted [Merryfield et al., 2013; Saha et al., 2014; Jia et al., 2015; Hudson et al., 2017; Kushnir et al., 2019].~~

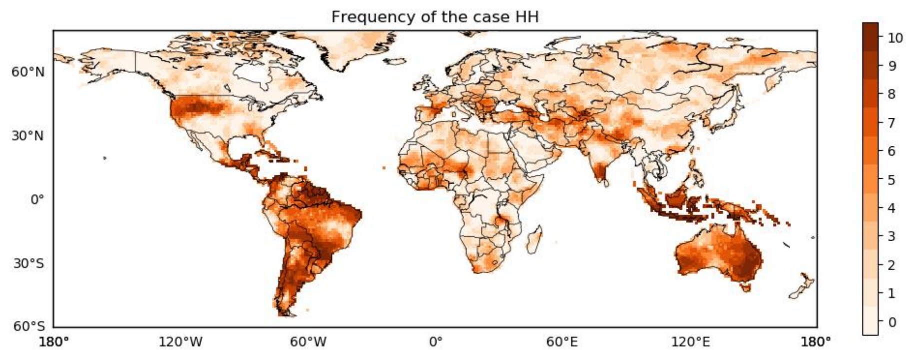


Figure 4: The number of the case HH across the ten sets of NMME seasonal forecasts at the global cells

~~4.5 Distribution of the frequency of the case HH~~

The percentage and cumulative percentage of the frequency of the case HH are shown by bar and line plots in Figure 5, respectively. The frequency of 0 corresponds to a percentage of nearly 20%. This outcome means that about 20% of the grid cells across the globe do not fall into the case HH ~~under~~ in any of the ten sets of forecasts. Another interpretation of this result is that about 80% of the grid cells fall into the case HH ~~under~~ in at least one of the ten sets of forecasts. This result is in contrast to Figure 3 suggesting that orange regions are limited under each of the ten set of forecasts. It highlights the spatial complementarity among the multiple sets of GCM forecasts [Doblas-Reyes et al., 2013; Merryfield et al., 2013; Jia et al., 2015]. In the meantime, the percentages corresponding to the frequencies of 5, 6, …, 10 are all below 5% and the cumulative percentage reaches 80% at the frequency of 4. This result reflects that the performances of the different sets of forecasts are not the same. Instead, they tend to differ from one another. In other words, for certain regions, some sets of GCM forecasts may not be satisfactory while some other sets of GCM forecasts may be promising. ~~Generally~~Overall, ~~the results in~~ Figure 5 suggests ~~the complementarity among the~~ that GCM forecasts in NMME can complement each other ~~and call for the use of model averaging methods to assign spatially-varying weighting coefficients and take advantage of the multiple sets of forecasts~~ [Wang et al., 2012; Becker et al., 2014; Kirtman et al., 2014].

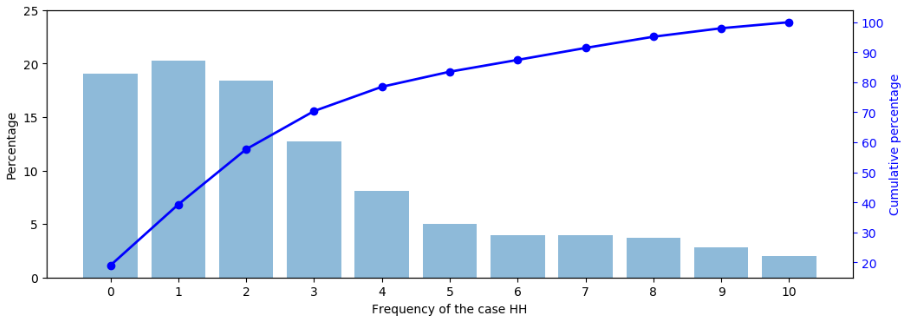


Figure 5: Percentage (bar plot) and cumulative percentage (line plot) of the number of the case HH

4.5 Frequency of the case HH in SON, DJF, and MAM

Besides JJA, spatial clustering has been performed for the anomaly correlation of GCM seasonal forecasts of total precipitation in SON, DJF, and MAM. Similarly, it is observed that the anomaly correlation varies across the globe (Figures S1, S4, and S7 in the supplementary material), correlates with its spatial lag (Figures S2, S5, and S8), and exhibits significant spatial patterns (Figures S3, S6, and S9). Similar to Figures 4 and 5, the frequency of the case HH is counted for the other three seasons and shown in Figures 6 and 7.

ENSO is one of the most important drivers of global climate [Mason and Goddard, 2001; Saha et al., 2014; Bauer et al., 2015], and the Climate Prediction Center (CPC) of NOAA has summarised the correlation between ENSO and global precipitation in different seasons (https://www.cpc.ncep.noaa.gov/products/precip/CWlink/ENSO/regressions/geplr.shtml). Therefore, this paper associates the results in Figure 6 with the global effects of ENSO. In SON, the CPC shows that ENSO correlates negatively with precipitation in Eastern Australia and Southeast Asia, and positively with precipitation in part of Middle East and East Africa. From the upper part of Figure 6, it is observed that the frequency of the case HH is high in these regions. In DJF, ENSO is shown to correlate positively with precipitation in Southern North America and negatively with precipitation in Northern South America. In these two regions, the frequency of the case HH is high (middle part of Figure 6). In MAM, ENSO is illustrated to correlate negatively with precipitation in part of Southeast Asia, Eastern Brazil, and Eastern Australia. Therein, the frequency of the case HH seem to be high (lower part of Figure 6). As previous studies found that GCMs in NMME generate skilful forecasts of ENSO [e.g., Kirtman et al., 2014; Saha et al., 2014; Zhang et al., 2017], Figure 6 suggests that the anomaly correlation of GCM forecasts in NMME can also be related to ENSO. In Figure 7, the percentage and cumulative percentage of the frequency of the case HH are illustrated for SON, DJF, and MAM. Similar to Figure 5, the results highlight the complementarity among the ten sets of forecasts.

Besides ENSO, there are other drivers of global climate. For example, North Atlantic Oscillation (NAO) and Arctic Oscillation (AO) extensively affect the climate in Europe, Asia, and North America [Hurrell et al., 2001; Ambaum et al., 2002]. Several sea surface temperature indices of the Atlantic and Indian Oceans and ENSO jointly impact the climate in Africa [Rowell, 2013]. According to Figures 4, 5, 6, and 7, there ~~are~~is still substantial room~~s~~ for improvement~~s~~ of seasonal precipitation forecasts ~~in June, July, and August~~ for large parts of Europe, Asia, and Africa. ~~It~~The unsatisfactory precipitation forecasts in these regions can ~~is~~ possibly be ~~because~~due to that GCM formulations of other climate ~~indices~~drivers are not as effective as the formulations of ENSO. In the meantime, the difficulty of global climate forecasting due to spatially-temporally varying teleconnections between regional precipitation and global climate drivers is noted [Merryfield et al., 2013; Saha et al., 2014; Jia et al., 2015; Hudson et al., 2017; Kushnir et al., 2019].

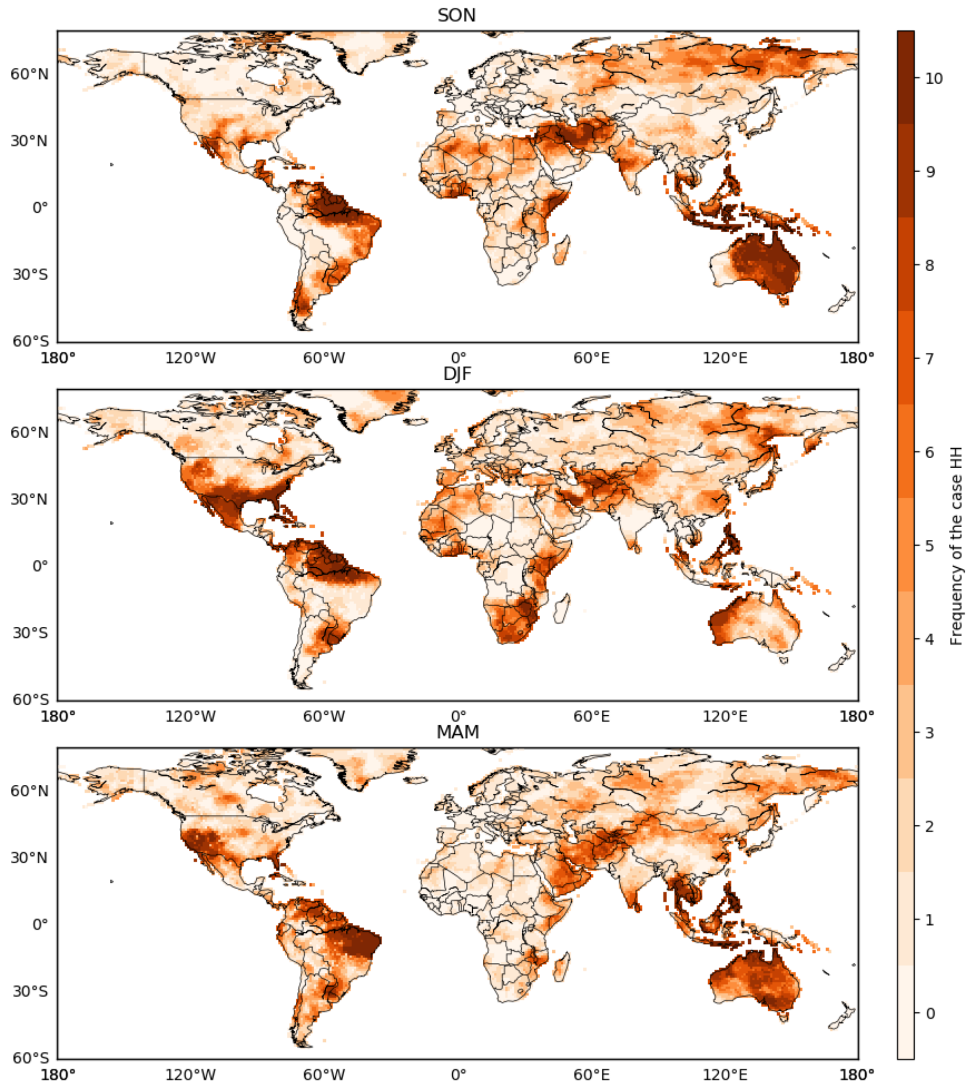


Figure 6: As for Figure 4, but for SON, DJF, and MAM

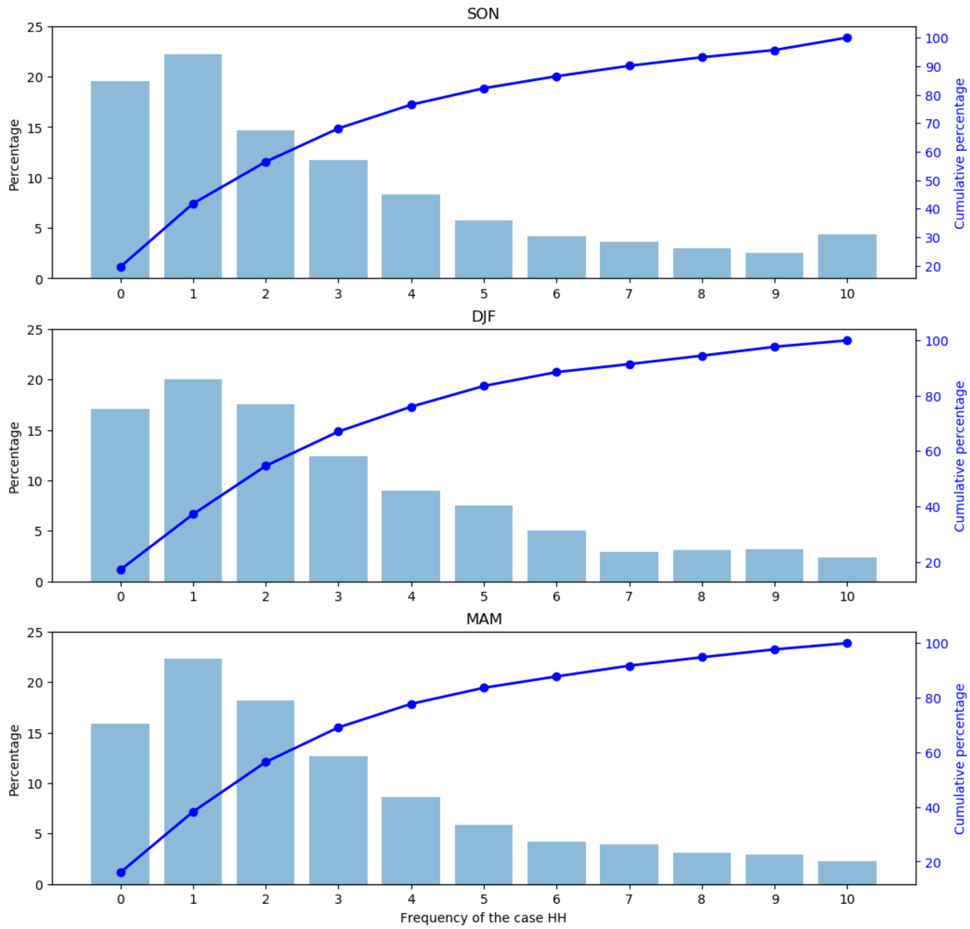


Figure 7: As for Figure 4, but for SON, DJF, and MAM

5 Discussion

~~We have identified~~This paper proposes to use spatial clustering to identify significant spatial patterns [Anselin, 1995; Miller, 2004; Schmal et al., 2017] from spatial plots of anomaly correlation, which have been widely used to illustrate the predictive performance of GCM forecasts. The test of significance is ~~conducted by~~based on global and local Moran's I~~ for ten sets of forecasts in NMME~~. The global Moran's I indicates that at the global scale anomaly correlation at one grid cell significantly relates to anomaly correlation at neighbouring grid cells, ~~while~~ and the local Moran's I reveals clusters of grid cells with high anomaly correlation. Across the ten sets of GCM forecasts in NMME, the clusters are observed in different regions across globe, which suggests that the skill of forecasts differs from region to region; in the meantime, the clusters vary by season, which indicates the seasonality in the skill of GCM forecasts ~~.~~[Doblas-Reyes et al., 2013; Becker et al., 2014; Yuan et al., 2015; Hudson et al., 2017; Kushnir et al., 2019]. To test the whether the spatial patterns are robust, the observations of precipitation are also sourced from the Global Precipitation Climatology Centre (GPCC) [Becker et al., 2011; Schamm et al., 2014]. The anomaly correlation is re-calculated, and the spatial clustering is re-conducted. The results of GPCC precipitation, which are shown in Figures S10 to 25 in the supplementary material, are overall similar to the results of CMAP precipitation. In particular, as to the two sets of precipitation observations, the spatial distributions of the case HH resemble in JJA (Figures 4 and S10) and also in SON, DJF, and MAM (Figures 6 and S12). This outcome highlights the existence of significant spatial patterns and confirms that ~~tend to be persistently observed in the Western United States, South America, Australia, and Southeast Asia, possibly due to the fact that the GCMs tend to capture the ENSO-related precipitation. Besides the seasonal forecasts for June, July, and August, the spatial clustering has also been applied to forecasts for September, October, and November, forecasts for December, January, and, and forecasts for March, April, and May. The results, which are illustrated in the supplementary material, are in accordance with Figures 1 to 5 and indicate the existence of significant patterns. Meanwhile, the spatial extents of clusters vary by season. The variability can be due to that the predictability of precipitation varies by season and that the effectiveness of GCMs also differs by season [Doblas-Reyes et al., 2013; Becker et al., 2014; Yuan et al., 2015; Hudson et al., 2017; Kushnir et al., 2019]. Nevertheless,~~ the spatial clustering ~~yields insights into the predictive performance and~~ can serve as an effective ~~approach~~ tool to yield insights into the predictive performance~~the inter-comparison~~ of GCM forecasts.

The spatial clustering ties anomaly correlation at neighbouring grid cells to one another and converts the continuous anomaly correlation into five categorical cases. Similar to the technique of moving average in time-series analysis, the categorical cases serve as a filter to reduce noise for the identification of spatial patterns. They handle the spatial variability of anomaly correlation and facilitate analysis across different sets of forecasts. The forecasts produced by the same climate center tend to exhibit similar predictive performance and the setting of GCMs leads to different predictive performance. These findings are obtained through ten sets of forecasts in NMME. Given the fact that the global and local Moran's I are flexible and easy to compute, they are ready to be extended in future analysis to other datasets of forecasts, such as forecasts generated by ~~Systems 4 and 5 at the European Centre for Medium-Range Weather Forecasts (ECMWF) and by other~~ GCMs in Europe and Asia or by regional climate models (RCMs) [Alfieri et al., 2013; Bellprat et al., 2019; Kushnir et al., 2019]. Also, the forecasts can be verified using global and regional datasets of precipitation [Funk et al., 2015; Zhao et al., 2017a, 2017b]. A more extensive in~~ter-comparison~~vestigation would contribute to better understanding of the predictive performance and illustrate the advantages of different sets of~~f GCM~~ forecasts. Of particular interest is to explore which ~~GCMs~~ forecasts achieve promising predictive performance in large parts of Europe, Asia, and Africa. In the meantime, it is meaningful to account for the dynamics of global climate and investigate the model physics that leads to the improved performance.

The spatial clustering is a popular approach to geographical, ecological, and environmental modelling [e.g., Anselin, 1995, 2006; Miller, 2004; Hao et al., 2016; Schmal et al., 2017]. Meanwhile, its use appears to be uncommon in the forecasting area. A possible cause is that the objective of forecasting is usually location-specific. In other words, forecasts are produced for a certain site/watershed and then verified using the corresponding observations, of which the process does not involve other sites/watersheds. In this paper, the analysis of GCM forecasts in NMME reveals that forecasts at neighbouring locations positively relate to one another. The indication is that the skill at one location can to some extent be inferred from adjacent locations. This result facilitates a new perspective for the verification of GCM forecasts. If a grid cell with high anomaly correlation is surrounded by grid cells with high anomaly correlation, the promising predictive performance at that grid cell can be confirmed. ~~However~~On the other hand, if the surrounding grid cells are with low, or even negative, anomaly correlation, then the high anomaly correlation is identified to be a suspicious outlier. Under that circumstance, further examination of the predictive performance is ~~highly~~ demanded to avoid undue optimism.

6 Conclusions

Fully-coupled GCMs perform physically-based forecasting of the global climate and generate a vast amount of spatial-temporal forecast data. The predictive performance is of both societal and scientific importance in the applications of these GCM forecasts. Focusing on the anomaly correlation between forecast ensemble mean and observation, we have conducted in-depth spatial analysis for ten sets of GCM forecasts in NMME and identified significant patterns from the spatial plotting of anomaly correlation. In the analysis of spatial clustering, grid cells across the globe are classified into five categories – HH, HL, NS, LH, and LL – depending on the anomaly correlation at that grid cell and the surrounding grid cells. The regions of grid cells with high, neutral, and low anomaly correlation are effectively identified. Further, effective inter-comparison across multiple sets of GCM forecasts is facilitated. While the analysis is concentrated on the spatial plotting of anomaly correlation, the framework readily applies to other metrics of GCM forecasts, such as bias, reliability, and skill. Moreover, the framework can be extended to GCM forecasts of other climate variables, for example ~~precipitation~~ temperature and wind speed, serving as a tool to explore GCM forecasts and interpret the predictive performance.

Acknowledgements

The work is supported by the Natural Science Foundation of China, the Guangdong Provincial Department of Science and Technology, and the Ministry of Science and Technology of China.

Data availability

Both the observations and the forecasts can be downloaded from the International Research Institute for Climate and Society, Earth Institute, Columbia University (https://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/)

Competing interests

The authors declare that they have no conflict of interest.

Author contributions

TZ, WZ, and YZ designed the experiment and performed the data analysis. TZ, ZL, and XC collected the data. TZ prepared the manuscript with contributions from all co-authors.

References

Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., and Pappenberger, F.: GloFAS - global ensemble streamflow forecasting and flood early warning, Hydrology and Earth System Sciences, 17, 1161-1175, https://doi.org/10.5194/hess-17-1161-2013, 2013.

Ambaum, M. H. P., Hoskins, B. J., and Stephenson, D. B.: Arctic oscillation or North Atlantic oscillation? (vol 14, pg 3495, 2001), Journal of Climate, 15, 553-553, https://doi.org/10.1175/1520-0442(2002)015<0553:c>2.0.co;2, 2002.

Anselin, L.: LOCAL INDICATORS OF SPATIAL ASSOCIATION - LISA, Geographical Analysis, 27, 93-115, https://doi.org/10.1111/j.1538-4632.1995.tb00338.x, 1995.

Anselin, L., Syabri, I., and Kho, Y.: GeoDa: An introduction to spatial data analysis, Geographical Analysis, 38, 5-22, https://doi.org/10.1111/j.0016-7363.2005.00671.x, 2006.

Bauer, P., Thorpe, A., and Brunet, G.: The quiet revolution of numerical weather prediction, Nature, 525, 47-55, https://doi.org/10.1038/nature14956, 2015.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., and Ziese, M.: GPCC full data reanalysis version 6.0 at 1.0: monthly land-surface precipitation from rain-gauges built on GTS-based and historic data, Global Precipitation Climatology Centre (GPCC): Berlin, Germany, https://doi.org/10.5065/D6000072, 2011.

Becker, E., van den Dool, H., and Zhang, Q.: Predictability and Forecast Skill in NMME, Journal of Climate, 27, 5891-5906, https://doi.org/10.1175/jcli-d-13-00597.1, 2014.

Bellprat, O., Guemas, V., Doblas-Reyes, F., and Donat, M. G.: Towards reliable extreme weather and climate event attribution, Nature Communications, 10, https://doi.org/10.1038/s41467-019-09729-2, 2019.

Chen, L.-C., Van den Dool, H., Becker, E., and Zhang, Q.: ENSO Precipitation and Temperature Forecasts in the North American Multimodel Ensemble: Composite Analysis and Validation, Journal of Climate, 30, 1103-1125, https://doi.org/10.1175/jcli-d-15-0903.1, 2017.

Cliff, A. D., and Ord, J. K.: Spatial processes: models & applications, Taylor & Francis, https://doi.org/10.1086/412797, 1981.

Crochemore, L., Ramos, M. H., and Pappenberger, F.: Bias correcting precipitation forecasts to improve the skill of seasonal streamflow forecasts, Hydrology and Earth System Sciences, 20, 3601-3618, https://doi.org/10.5194/hess-20-3601-2016, 2016.

Di Luzio, M., Johnson, G. L., Daly, C., Eischeid, J. K., and Arnold, J. G.: Constructing retrospective gridded daily precipitation and temperature datasets for the conterminous United States, Journal of Applied Meteorology and Climatology, 47, 475-497, https://doi.org/10.1175/2007jamc1356.1, 2008.

Doblas-Reyes, F. J., Garcia-Serrano, J., Lienert, F., Biescas, A. P., and Rodrigues, L. R. L.: Seasonal climate predictability and forecasting: status and prospects, Wiley Interdisciplinary Reviews-Climate Change, 4, 245-268, https://doi.org/10.1002/wcc.217, 2013.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., and Hoell, A.: The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes, Scientific data, 2, 150066, https://doi.org/10.1038/sdata.2015.66, 2015.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang, Z. L., and Zhang, M. H.: The Community Climate System Model Version 4, Journal of Climate, 24, 4973-4991, https://doi.org/10.1175/2011jcli4083.1, 2011.

Getis, A.: Reflections on spatial autocorrelation, Regional Science and Urban Economics, 37, 491-496, https://doi.org/10.1016/j.regsciurbeco.2007.04.005, 2007.

Gneiting, T., Balabdaoui, F., and Raftery, A. E.: Probabilistic forecasts, calibration and sharpness, Journal of the Royal Statistical Society Series B-Statistical Methodology, 69, 243-268, https://doi.org/10.1111/j.1467-9868.2007.00587.x, 2007.

Hao, Y., Liu, Y. M., Weng, J. H., and Gao, Y. X.: Does the Environmental Kuznets Curve for coal consumption in China exist? New evidence from spatial econometric analysis, Energy, 114, 1214-1223, https://doi.org/10.1016/j.energy.2016.08.075, 2016.

Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems, Weather and Forecasting, 15, 559-570, https://doi.org/10.1175/1520-0434(2000)015<0559:dotcrp>2.0.co;2, 2000.

Hudson, D., Alves, O., Hendon, H. H., Lim, E.-P., Liu, G., Luo, J.-J., MacLachlan, C., Marshall, A. G., Shi, L., Wang, G., Wedd, R., Young, G., Zhao, M., and Zhou, X.: ACCESS-S1: The new Bureau of Meteorology multi-week to seasonal prediction system, Journal of Southern Hemisphere Earth Systems Science, 67, 132-159, 10.22499/3.6703.001, 2017.

Hurrell, J. W., Kushnir, Y., and Visbeck, M.: Climate - The North Atlantic oscillation, Science, 291, 603-605, https://doi.org/10.1126/science.1058761, 2001.

Jia, L. W., Yang, X. S., Vecchi, G. A., Gudgel, R. G., Delworth, T. L., Rosati, A., Stern, W. F., Wittenberg, A. T., Krishnamurthy, L., Zhang, S. Q., Msadek, R., Kapnick, S., Underwood, S., Zeng, F. R., Anderson, W. G., Balaji, V., and Dixon, K.: Improved Seasonal Prediction of Temperature and Precipitation over Land in a High-Resolution GFDL Climate Model, Journal of Climate, 28, 2044-2062, https://doi.org/10.1175/jcli-d-14-00112.1, 2015.

Kirtman, B. P., Min, D., Infanti, J. M., Kinter, J. L., Paolino, D. A., Zhang, Q., van den Dool, H., Saha, S., Mendez, M. P., Becker, E., Peng, P. T., Tripp, P., Huang, J., DeWitt, D. G., Tippett, M. K., Barnston, A. G., Li, S. H., Rosati, A., Schubert, S. D., Rienecker, M., Suarez, M., Li, Z. E., Marshak, J., Lim, Y. K., Tribbia, J., Pegion, K., Merryfield, W. J., Denis, B., and Wood, E. F.: THE NORTH AMERICAN MULTIMODEL ENSEMBLE Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction, Bulletin of the American Meteorological Society, 95, 585-601, https://doi.org/10.1175/bams-d-12-00050.1, 2014.

Kushnir, Y., Scaife, A. A., Arritt, R., Balsamo, G., Boer, G., Doblas-Reyes, F., Hawkins, E., Kimoto, M., Kolli, R. K., Kumar, A., Matei, D., Matthes, K., Muller, W. A., O'Kane, T., Perlwitz, J., Power, S., Raphael, M., Shimpo, A., Smith, D., Tuma, M., and Wu, B.: Towards operational predictions of the near-term climate, Nature Climate Change, 9, 94-101, https://doi.org/10.1038/s41558-018-0359-7, 2019.

Lu, G. Y., and Wong, D. W.: An adaptive inverse-distance weighting spatial interpolation technique, Computers & Geosciences, 34, 1044-1055, https://doi.org/10.1016/j.cageo.2007.07.010, 2008.

Luo, L., Tang, W., Lin, Z., and Wood, E. F.: Evaluation of summer temperature and precipitation predictions from NCEP CFSv2 retrospective forecast over China, Climate Dynamics, 41, 2213-2230, https://doi.org/10.1007/s00382-013-1927-1, 2013.

Ma, F., Ye, A., Deng, X., Zhou, Z., Liu, X., Duan, Q., Xu, J., Miao, C., Di, Z., and Gong, W.: Evaluating the skill of NMME seasonal precipitation ensemble predictions for 17 hydroclimatic regions in continental China, International Journal of Climatology, 36, 132-144, https://doi.org/10.1002/joc.4333, 2016.

Mason, S. J., and Goddard, L.: Probabilistic precipitation anomalies associated with ENSO, Bulletin of the American Meteorological Society, 82, 619-638, https://doi.org/10.1175/1520-0477(2001)082<0619:ppaawe>2.3.co;2, 2001.

Merryfield, W. J., Lee, W. S., Boer, G. J., Kharin, V. V., Scinocca, J. F., Flato, G. M., Ajayamohan, R. S., Fyfe, J. C., Tang, Y. M., and Polavarapu, S.: The Canadian Seasonal to Interannual Prediction System. Part I: Models and Initialization, Monthly Weather Review, 141, 2910-2945, https://doi.org/10.1175/mwr-d-12-00216.1, 2013.

Miller, H. J.: Tobler's First Law and spatial analysis, Annals of the Association of American Geographers, 94, 284-289, https://doi.org/10.1111/j.1467-8306.2004.09402005.x, 2004.

Murphy, A. H.: What is a good forecast? An essay on the nature of goodness in weather forecasting, Weather and forecasting, 8, 281-293, https://doi.org/10.1175/1520-0434(1993)008<0281:WIAGFA>2.0.CO;2, 1993.

Pappenberger, F., Cloke, H. L., and Baugh, C. A.: Cartograms for Use in Forecasting Weather-Driven Natural Hazards, Cartographic Journal, 56, 134-145, https://doi.org/10.1080/00087041.2018.1534358, 2019.

Rey, S. J., and Anselin, L.: PySAL: A Python library of spatial analytical methods, in: Handbook of applied spatial analysis, Springer, 175-193, https://doi.org/10.1007/978-3-642-03647-7_11, 2010.

Rowell, D. P.: Simulating SST Teleconnections to Africa: What is the State of the Art?, Journal of Climate, 26, 5397-5418, https://doi.org/10.1175/jcli-d-12-00761.1, 2013.

Saha, S., Moorthi, S., Wu, X. R., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y. T., Chuang, H. Y., Iredell, M., Ek, M., Meng, J., Yang, R. Q., Mendez, M. P., Van Den Dool, H., Zhang, Q., Wang, W. Q., Chen, M. Y., and Becker, E.: The NCEP Climate Forecast System Version 2, Journal of Climate, 27, 2185-2208, https://doi.org/10.1175/jcli-d-12-00823.1, 2014.

Schamm, K., Ziese, M., Becker, A., Finger, P., Meyer-Christoffer, A., Schneider, U., Schroeder, M., and Stender, P.: Global gridded precipitation over land: a description of the new GPCC First Guess Daily product, Earth System Science Data, 6, 49-60, https://doi.org/10.5194/essd-6-49-2014, 2014.

Schepen, A., Wang, Q. J., and Robertson, D.: Evidence for Using Lagged Climate Indices to Forecast Australian Seasonal Rainfall, Journal of Climate, 25, 1230-1246, https://doi.org/10.1175/jcli-d-11-00156.1, 2012.

Schmal, C., Myung, J., Herzel, H., and Bordyugov, G.: Moran's I quantifies spatio-temporal pattern formation in neural imaging data, Bioinformatics, 33, 3072-3079, https://doi.org/10.1093/bioinformatics/btx351, 2017.

Sheffield, J., Wood, E. F., Chaney, N., Guan, K. Y., Sadri, S., Yuan, X., Olang, L., Abou, A., Ali, A., Demuth, S., and Ogallo, L.: A DROUGHT MONITORING AND FORECASTING SYSTEM FOR SUB-SAHARA AFRICAN WATER RESOURCES AND FOOD SECURITY, Bulletin of the American Meteorological Society, 95, 861-+, https://doi.org/10.1175/bams-d-12-00124.1, 2014.

Slater, L. J., Villarini, G., and Bradley, A. A.: Weighting of NMME temperature and precipitation forecasts across Europe, Journal of Hydrology, 552, 646-659, https://doi.org/10.1016/j.jhydrol.2017.07.029, 2017.

Tian, D., Wood, E. F., and Yuan, X.: CFSv2-based sub-seasonal precipitation and temperature forecast skill over the contiguous United States, Hydrology and Earth System Sciences, 21, 1477-1490, https://doi.org/10.5194/hess-21-1477-2017, 2017.

Tobler, W. R.: A computer movie simulating urban growth in the Detroit region, Economic geography, 46, 234-240, 1970.

Vecchi, G. A., Delworth, T., Gudgel, R., Kapnick, S., Rosati, A., Wittenberg, A. T., Zeng, F., Anderson, W., Balaji, V., Dixon, K., Jia, L., Kim, H. S., Krishnamurthy, L., Msadek, R., Stern, W. F., Underwood, S. D., Villarini, G., Yang, X., and Zhang, S.: On the Seasonal Forecasting of Regional Tropical Cyclone Activity, Journal of Climate, 27, 7994-8016, https://doi.org/10.1175/jcli-d-14-00158.1, 2014.

Wang, Q. J., Schepen, A., and Robertson, D. E.: Merging Seasonal Rainfall Forecasts from Multiple Statistical Models through Bayesian Model Averaging, Journal of Climate, 25, 5524-5537, https://doi.org/10.1175/jcli-d-11-00386.1, 2012.

Woldemeskel, F. M., Sivakumar, B., and Sharma, A.: Merging gauge and satellite rainfall with specification of associated uncertainty across Australia, Journal of Hydrology, 499, 167-176, https://doi.org/10.1016/j.jhydrol.2013.06.039, 2013.

Wu, Z., Wang, B., Li, J., and Jin, F.-F.: An empirical seasonal prediction model of the east Asian summer monsoon using ENSO and NAO, Journal of Geophysical Research-Atmospheres, 114, https://doi.org/10.1029/2009jd011733, 2009.

Xie, P., and Arkin, P. A.: Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates, and Numerical Model Outputs, Bulletin of the American Meteorological Society, 78, 2539-2558, https://doi.org/10.1175/1520-0477(1997)078<2539:gpayma>2.0.co;2, 1997.

Xie, P., Arkin, P. A., and Janowiak, J. E.: CMAP: The CPC merged analysis of precipitation, in: Measuring precipitation from space, Springer, 319-328, https://doi.org/10.1007/978-1-4020-5835-6, 2007.

Yuan, X., Wood, E. F., Luo, L., and Pan, M.: A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction, Geophysical research letters, 38, https://doi.org/10.1029/2011GL047792, 2011.

Yuan, X., Wood, E. F., Roundy, J. K., and Pan, M.: CFSv2-based seasonal hydroclimatic forecasts over the conterminous United States, Journal of Climate, 26, 4828-4847, https://doi.org/10.1175/JCLI-D-12-00683.1, 2013.

Yuan, X., Wood, E. F., and Ma, Z.: A review on climate-model-based seasonal hydrologic forecasting: physical understanding and system development, Wiley Interdisciplinary Reviews: Water, 2, 523-536, https://doi.org/10.1002/wat2.1088, 2015.

Yuan, Y. M., Cave, M., and Zhang, C. S.: Using Local Moran's I to identify contamination hotspots of rare earth elements in urban soils of London, Applied Geochemistry, 88, 167-178, https://doi.org/10.1016/j.apgeochem.2017.07.011, 2018.

Zhang, W., Villarini, G., Slater, L., Vecchi, G. A., and Bradley, A. A.: Improved ENSO Forecasting Using Bayesian Updating and the North American Multimodel Ensemble (NMME), Journal of Climate, 30, 9007-9025, https://doi.org/10.1175/jcli-d-17-0073.1, 2017.

Zhao, T. T. G., Bennett, J. C., Wang, Q. J., Schepen, A., Wood, A. W., Robertson, D. E., and Ramos, M. H.: How Suitable is Quantile Mapping For Postprocessing GCM Precipitation Forecasts?, Journal of Climate, 30, 3185-3196, https://doi.org/10.1175/jcli-d-16-0652.1, 2017a.

Zhao, T. T. G., Liu, P., Zhang, Y. Y., and Ruan, C. Q.: Relating anomaly correlation to lead time: Clustering analysis of CFSv2 forecasts of summer precipitation in China, Journal of Geophysical Research-Atmospheres, 122, 9094-9106, 10.1002/2017jd027018, 2017b.

Zhao, T. T. G., Chen, X. H., Liu, P., Zhang, Y. Y., Liu, B. J., and Lin, K. R.: Relating Anomaly Correlation to Lead Time: Principal Component Analysis of NMME Forecasts of Summer Precipitation in China, Journal of Geophysical Research-Atmospheres, 123, 6039-6052, https://doi.org/10.1029/2018jd028267, 2018.

Zhao, T. T. G., Wang, Q. J., Schepen, A., and Griffiths, M.: Ensemble forecasting of monthly and seasonal reference crop evapotranspiration based on global climate model outputs, Agricultural and Forest Meteorology, 264, 114-124, https://doi.org/10.1016/j.agrformet.2018.10.001, 2019a.

Zhao, T. T. G., Zhang, Y. Y., and Chen, X. H.: Predictive performance of NMME seasonal forecasts of global precipitation: A spatial-temporal perspective, Journal of Hydrology, 570, 17-25, 10.1016/j.jhydrol.2018.12.036, 2019b.