Responses to referees' comments

To referee #2

Major comments:

##

1. The statistical forecasting methodology adopts the predictor with the highest correlation after model fitting (Figure 1 and lines 82-88). This means the selected climate indices are based on peeking during the evaluation period and selected to maximise the performance over the evaluation period. With 18 climate indices, the risk of artificially inflating the performance of the statistical system is high. I suggest that the selected climate index for each year ought to vary based on total exclusion of the data for that year through all steps up to and including the calculation of the correlation.

##

The description on the development of statistical models was confusing in the previous manuscript. Actually we first constructed statistical models by removing the data in forecast years and then produced precipitation forecasts by using the removed data. We improved the explanation on the development of the statistical models and modified Figure 1, as follows:

> "Next, in Step 2-1, 18 monthly precipitation forecasts were produced for each grid by using 18 climate indices and GPCP v2.3. The following leave-one-out method was used for producing the forecasts. First, the data of the forecast year was removed before constructing statistical models with monthly values of a climate index as the explanatory variable and monthly precipitation as the objective variable (see Section 2.2.2) through the smoothing spline method. Then, the forecast values were obtained using the constructed statistical models and removed data as inputs for the models."

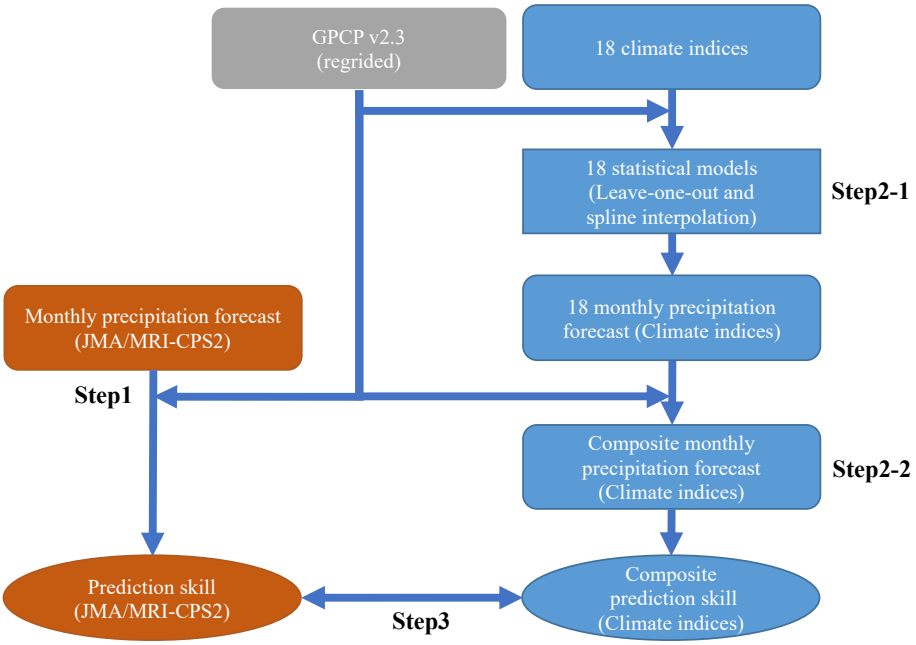

Figure 1: Research outline

##

2. Seasonal forecasting is highly uncertain, and the deterministic forecasts of the type produced here may be of limited value to end users. I suggest that the authors consider a means to produce uncertainty estimates or "ensembles" and include some more typical SCF analysis and verification results. I believe that is important to analyse the underlying skill in the ensemble mean as the authors have done, however, adopting it as the only verification approach side-steps a major advantage of the dynamical forecasts in having ensembles.

##

Thank you for your very important comment. We recognize that it is a very important issue to compare the accuracy of forecasts, taking into account the uncertainty of forecasts. However, this includes a difficult challenge as described below. Therefore, we would like to make this a future challenge and describe it in Section 4 (Discussion and conclusions).

First of all, even in statistical models, it is possible to show the uncertainty of forecasts by showing the confidence interval of forecasts or to make ensembles by using resampling methods such as the bootstrap method. Moreover, by using the resampling methods, statistical forecast systems can produce very large ensembles (e.g., 10,000) much more easily than dynamical ones. In order to compare the accuracy of two systems that handle uncertainty in a different way, we must first

understand the characteristics of the uncertainty in each system, and then consider how to compare them. For example, how do we compare the accuracy of a Dyn-SCF with an ensemble of approximately 10 and a St-SCF with an ensemble of about 10,000, taking uncertainty into account? We believe that intensive studies on this issue are needed. However, this is clearly beyond the scope of the present research. We would like to discuss the point in Section 4 (Discussion and conclusions) as follows:

> "How to compare the accuracy of Dyn-SCFs and St-SCFs considering the uncertainty in forecasts is an important future challenge. Forecasting inherently includes uncertainty. For this reason, ensemble forecasting is widely conducted in Dyn-SCFs systems. In St-SCFs systems, a huge ensemble of forecasts can be easily made by using resampling methods such as the bootstrap method or the confidence interval of the forecasts can be calculated. In order to compare the accuracy of two forecast systems that handle uncertainty in a different way, we must first understand the characteristics of the uncertainty in each system, and then consider how to compare them. For example, how do we compare the accuracy of a Dyn-SCF with an ensemble of approximately 10 and a St-SCF with an ensemble of about 10,000, taking uncertainty into account? We believe that intensive studies on this issue are needed and it is expected that such studies will be conducted in the future."

Minor comments:

##########

I suggest that the motivation for studying precipitation should be established earlier in the introduction, and a comment that statistical forecasts are suited to forecasting a small number of variables.

##########

We made new paragraph for the explanation on the motivation for focusing on precipitation, as follows:

> "Precipitation forecasting is very important for effective water management and disaster reduction. It has been shown that the accuracy of precipitation forecasts in Dyn-SCF systems is lower than that that of temperature forecasts, and areas with highly accurate precipitation forecasts is limited in the tropics (Doblas-Reyes et al., 2013). So far, it is

not well understood to what extent the accuracy of Dyn-SCF systems for precipitation forecasts adds value compared to St-SCF systems."

In addition, we added the explanation that statistical forecast systems can be alternative for dynamical ones in case of focusing a small number of climate variables, as follows:

"In case of forecasting a small number of specific climate variables, statistical SCF (St-SCF) systems are an alternative and simpler method for Dyn-SCF systems"

##

The process of selecting a different climate index for each cell could lead to very spatially inconsistent forecasts. I suggest this is addressed in the revised discussion.

##

In the revised manuscript, we added a figure showing the selected climate indices for each grid. This figure shows that the same index tends to be selected in each small area. This implies that St-SCFs using climate indices have a consistent spatial forecast within each small area. We added the following figure and the explanation above, as follows:

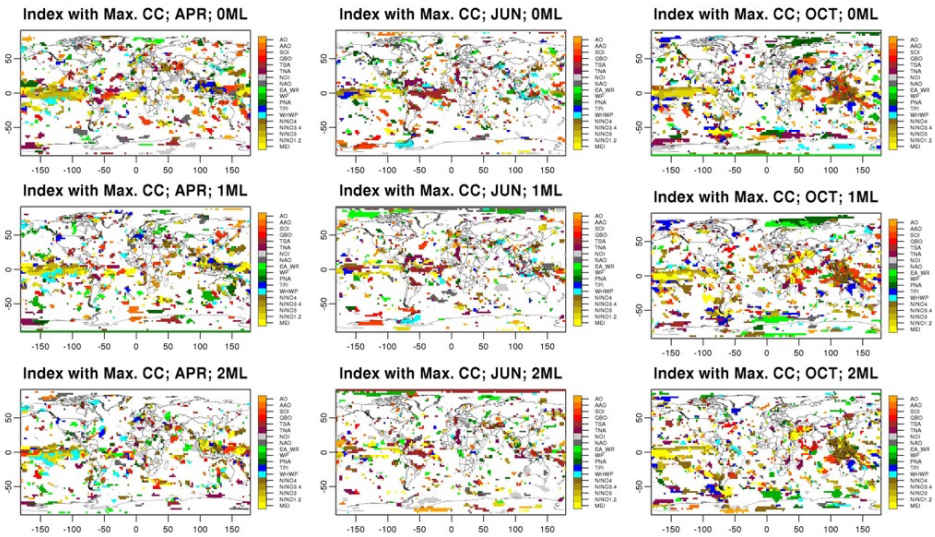

Figure 10: Climate indices selected for grids with significant CC. Left, center, and right columns denote Apr, Jun, and Oct, respectively. Top, middle, and bottom denote 0 to 2 months lead.

"Fig. 10 shows the climate indices selected for each grid. This figure shows that the selected indices depend on the regions and forecast month. The indices selected do not change significantly from zero to two months lead forecasts, and while the area with a significant CC has a patchy distribution, the same index tends to be selected in each small area. This implies that St-SCFs using climate indices have a consistent spatial forecast within each small area."

##########

It is one intention of this study to highlight that dynamical models don't always capture the slow dynamics and teleconnections. However, the paper is rather brief and doesn't attempt to explain why particular climate indices are important, or even the frequency at which they are selected. I suggest in the revision some further attention is given to highlighting the most important indices.

##########

As mentioned above, we added a figure showing the selected climate indices for each grid. This figure shows that the selected indices depend on the regions and forecast month. Therefore, we can see that there is a particular index that is important for each region and lead month (Figure 10 and revised description shown).

##########

Mean square error and skill scores are referred to in the introduction, however they are not used in the paper. I suggest making the introduction and methods more consistent.

##########

We removed the sentence in the revised manuscript.

##

It sounds to me like the climate indices wouldn't be available in time to make real time forecasts due the week delay (e.g. section 4.3). I suggest addressing this issue in the revision.

##

Our intention in the sentences is that the forecast accuracy of the St-SCF systems could be improved if a larger number of climate indices than the present study. We modified the sentences, as follows:

> "In future studies, using a larger number of climate indices will improve the forecast accuracy."