We are very grateful to the Reviewer for the positive and careful review. The thoughtful comments have helped improve the manuscript. The reviewer's comments are italicized and our responses immediately follow.

Responses to the comments from Reviewer #1

This article by Wang and Yuan investigates seasonal predictability of water vapor flux and precipitation in the CFSv2 forecasts to see if water vapor flux could provide higher predictability of floods. I enjoyed reading the study and believe that this is an interesting study and one that is relevant to forecast users. I currently have two major concerns which I lay out below, and some further more minor comments. I hope the authors will find them of use. Major comments:

Response: We would like to thank the reviewer for the positive comments. Please see our responses below.

1. While the language is generally very good, there are grammatical issues, so the paper would benefit from being read by a native English speaker. Also, I think the authors need to be careful with their terminology. For example, on many occasions they say "moisture" when in fact they mean "moisture flux" or "water vapor flux". This needs to be corrected so that readers are not confused by the terms (moisture could be viewed as total column water vapor).

Response: Thanks for the comments. We have corrected inappropriate terminology and proofread the manuscript carefully. And the "moisture" has been replaced with "moisture flux" throughout the manuscript.

2. The authors only analyse monthly fields and I believe that this is not sufficient. This is useful, but I think some testing should be done on at least weekly or 2-weekly averaged fields. The reason is that it will be easier (and more beneficial) to see when the predictability drops off in the model at a finer time resolution. For example, Luo and Wood (2006) have some good plots that may help you to consider other time averages.

Response: Thanks for the comments. According to the reviewer's suggestion, we have incorporated the analysis for the potential predictability of precipitation and moisture flux for weekly averaged fields in the revised manuscript.

We have used the S2S daily data, and have clarified them in the Section "2.2 CFSv2 seasonal hindcast and real-time forecast data" as follows:
"In order to investigate the predictability at finer temporal resolution (e.g., weekly mean fields), the CFSv2 daily reforecasts were obtained from the Subseasonal to Seasonal (S2S) prediction project for the period of 1999-2010, with forecast lead times up to 45-days (Vitart et al. 2017). As for the June 1-7 weekly mean fields, the reforecasts started from May 18 were used as the first ensemble member, the reforecasts started from May 19 were used as the second, and so on. This resulted in 14 ensemble members, with forecast lead times from 1-day to 14-days. The above process was repeated for other weekly averaged fields during June and July. This is called as the first group of ensemble subseasonal forecasts, with lead times of 1-14days. The second group of ensemble reforecasts started from 17 May, 18 May …, and 30 May were formed similarly, with lead times of 2-15days, and so on."

The AC values for both weekly mean precipitation and moisture flux during June-July at different lead times were obtained and shown in Figure R1 (Figure 7 in the revised manuscript). Results show that the moisture flux has a higher predictability than precipitation for weekly averaged fields at different lead times, which is consistent with the results on seasonal averaged fields. We have added the related discussion in Section 4 as follows:

"In addition, we also investigated potential predictability of precipitation and moisture flux on weekly averaged fields in June-July at subseasonal scale. Results are similar to seasonal time scale, where the moisture flux has a higher predictability than precipitation at different lead times (Fig. 7)."

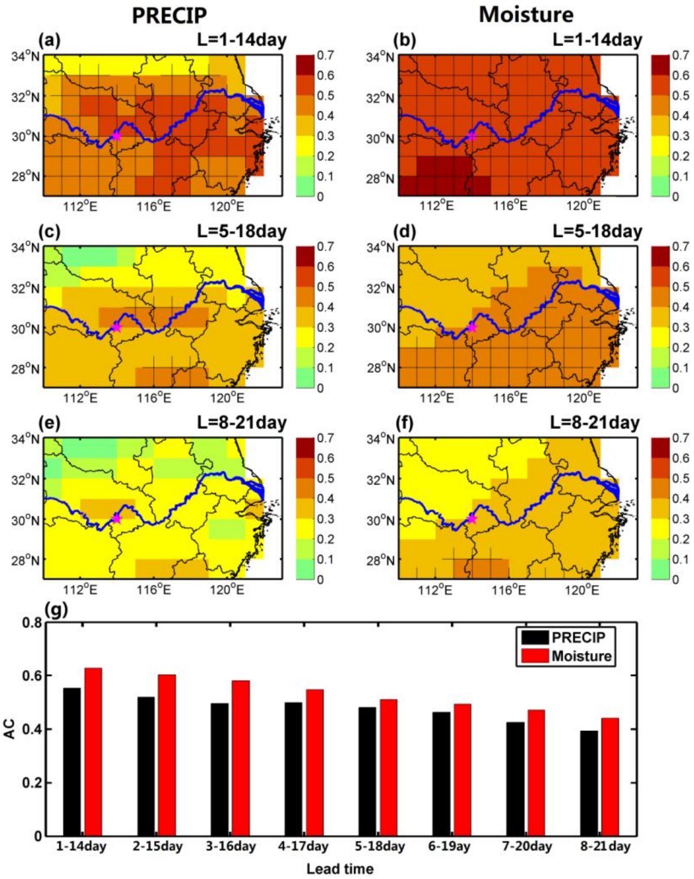

Figure R1. (a-f) Potential predictability (AC value) for weekly mean precipitation and atmospheric moisture flux at different lead times during June-July of 1999-2010 over the middle and lower reaches of Yangtze River for the 1-14, 5-18 and 8-21 days leads; the stippling indicates a 95% confidence level according to a two-tailed Student's t-test. (g) Potential predictability throughout study region at different lead times. Here, the daily CFSv2 reforecast were obtained from the S2S prediction project for the period of 1999-2010.

Other comments:

Line 19: "atmospheric moisture" – see major comment 1.

Response: Thanks for the suggestion. We have corrected them throughout the manuscript.

Line 33: "ability to be predicted" is a strange phrase. Perhaps consider re-phrasing.

Response: We have changed it to "the ability of the model to predict itself".

Line 40: Precipitation is connected to mesoscale (or more local scale) circulation and orography.
Response: Thanks for the comments. We have revised it as follows:
"The atmospheric moisture flux is supposed to be better predicted by large-scale climate models than precipitation that is connected to mesoscale (or more local scale) circulation and orography".

Line 61: I think Lavers et al (2016a) investigated prediction skill, not predictability.
Response: Revised as suggested.

Line 75: What pressure levels were used? Please add some details here.
Response: We have specified as "Monthly mean atmospheric fields including geopotential height, u-wind, v-wind, and specific humidity at 300, 400, 500, 700, 850, 925 and 1000 hPa were derived from the ERA-Interim reanalysis".

Lines 103-112: Why does the AC go from 0.33 to 0.44 (compare Figure 1b and 1d)? You would expect the predictability to drop with lead time.
Response: In general, the predictability drops over lead times, but not necessarily for any cases. We plotted the results for all 24 ensemble members in Figure R2, and found that the AC for 0.5-month lead is not necessarily higher than 1.5-month lead. However, the average results for the 24 AC (Fig. R2c) shows that AC decreases over leads on average.

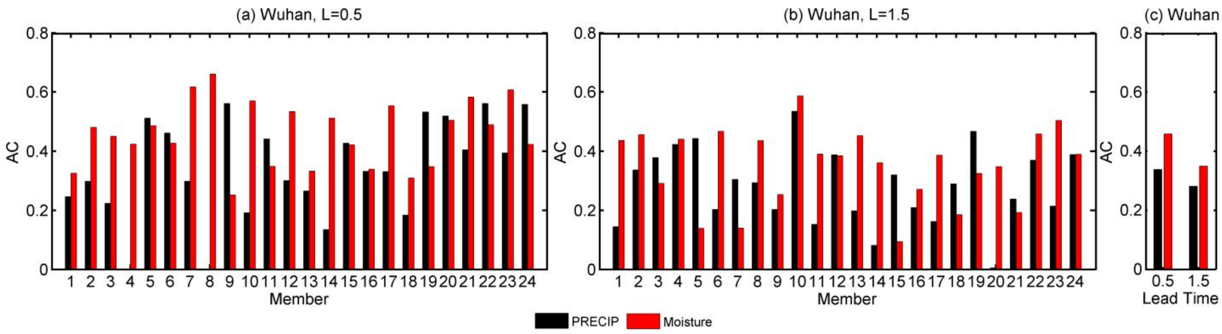

Figure R2. Potential predictability (AC value) when different ensemble member was taken as the truth and the mean of the members was the prediction at Wuhan city for the (a) 0.5- and (b) 1.5-month leads. (c) the final estimate of the potential predictability in Wuhan city.

Line 116: I would advise not to use the "upper limit of forecasting skill". This is not always true, as explained in Kumar et al (2014).
Response: We have removed this statement.

Line 119: "Seasonal predictability" is probably seasonal predictive skill.
Response: As mention in L124-130, "a forecast for flooding event can be counted at a given grid or region when taking ensemble member 1 as observation and the average of members 2–24 as the prediction". So, the hit rate used here is to assess the seasonal predictability for extreme hydrologic events.

Line 124: "b" is not in the equation. Please remove.

Response: Revised as suggested.

Line 131: It is hard to see the Yangtze River. Can this figure be edited so that the river is more clear?
Response: Revised as suggested. Please see Figure 2 below.

Line 133: "pummelled" is not really scientific. Can this be rephrased?
Response: We have changed to "hit".

Line 183: Can you consider plotting the differences between precipitation and moisture flux in Figure 4g? This would more clearly show any significant differences between the variables.
Response: Thanks for the comments. We have examined the statistical significance, and revised the manuscript as follows:
"The difference between precipitation and moisture flux is statistically significant ($p<0.05$) with a two-tailed Student's t-test"

Figure 2: In panel c can you use the total column-integrated moisture flux instead of just the 850 hPa level? This would match the rest of the paper.
Response: Revised as suggested.

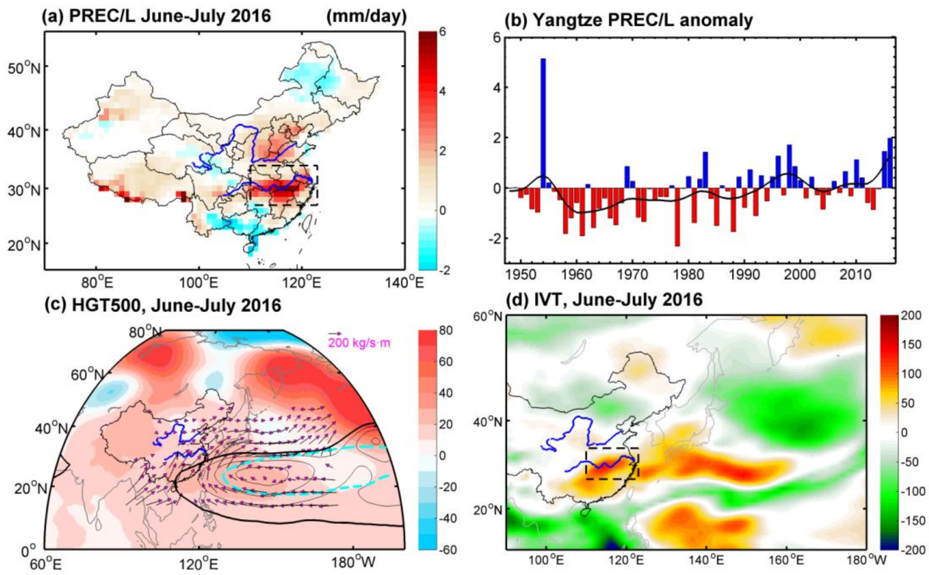

Figure 2. The 2016 extreme summer flood. (a) Mean precipitation anomaly (shading, mm/day) during the June-July of 2016. (b) Time series of the June-July mean precipitation anomaly averaged over the middle and lower reaches of Yangtze River basin (110-123°E, 27-34°N) in (a). (c) Anomaly of 500 hPa geopotential height (shading, gpm) superimposed by absolute integrated horizontal moisture transport between 1000 to 300 hPa layers (vectors, kg•m-1s-1). The thick contour lines are 5880 gpm, implying the location of the West Pacific Subtropical High, where the black denotes the June-July 2016 and the cyan is the climatology during 1982-2010. (d) Anomaly of integrated horizontal moisture transport amount (shading, kg•m-1s-1).

Figure 3: What initialisation times are used in this figure (e.g. 1st May 2016)? Please consider adding to the caption.

Response: We have revised as "where the 0.5-month lead forecasts were initialized from mid-May to early June in 2016, 1.5-month lead forecasts were initialized from mid-Apr to early May in 2016, and so on."

Figure 6: Panels c-d. Perhaps a few extra contours should be added to more clearly show the 500 hPa geopotential height?

Response: Revised as suggested.

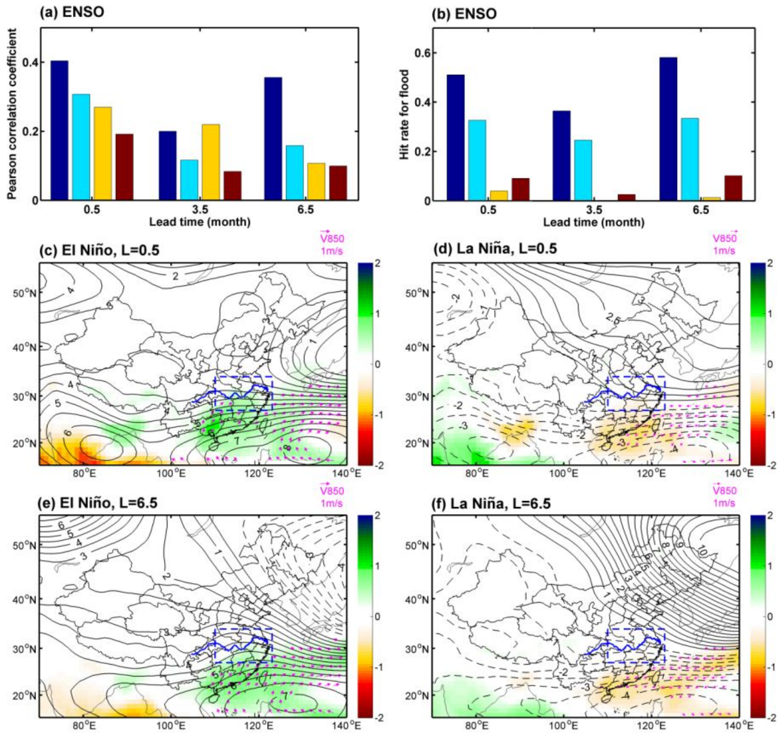

Figure 6. Potential predictability at different lead times in terms of (a) anomaly correlation (AC) for precipitation and moisture, and (b) hit rate (HR) for flood events (>90th percentiles) across the Yangtze River region conditioned on ENSO phases. (c-d) Composites of predicted anomalies of 500 hPa geopotential height (contour, gpm) superimposed by 850 hPa wind (vectors, m/s) and moisture flux (shading, g/cm•hPa•s) at the 0.5-month lead during different ENSO phases. (e-f) The same as (c-d), but for 6.5-month lead time.