The authors analyzed the effect of precipitation post-processor (CSGD) for four different four hydro-meteorological forecasting systems. Those forecasting systems were differed by the degree of uncertainty considered: system A) forcing, system B) forcing and initial conditions, system C) forcing and model structure, and system D) forcing, initial conditions, and model structure. The results showed that precipitation post-processor worked better for less complex hydrological systems (system A/B/C). Quantifying all sources of uncertainty (forcing, initial conditions, and model structure) did not always lead to the best results in streamflow forecasting. The authors also compared the contribution of the post-processor for different catchment sizes.

The subject clearly fits into the scope of the journal and provides a useful guidance for the choice of forecasting systems. The authors sufficiently draw upon the existing body of literature and the research is interesting. However, I have some concerns with the research method. For these reasons, I recommend a decision of Minor Revision for this manuscript. Please find my comments in the following. I hope you find them useful.

1. The research is a comparison between different forecasting systems. Quantitative results are needed to justify the conclusions. However, the authors tend to use sentences like 'Line 467-468: We also observe that post-processing precipitation forecasts have a **much higher** impact on the quality of precipitation forecasts (Fig. 9) than on the quality of streamflow forecasts, as evaluated by the BIAS score', or 'Line 491-491: It is interesting to note that, for systems B, C, and D, streamflow forecasts based on raw precipitation forecasts are always **much better** than streamflow forecasts based on post-processed precipitation forecasts in system'. They don't give the readers an objective description for the research. The qualitative results cannot help the scientific choice. Please provide more numerical results, especially in conclusion.

2. Concerns about the multimodel approaches. From Figure 8, the multimodel approach seems to bring additional uncertainty to streamflow forecasts, as system C always has the worst results in term of BIAS and IQR. From Figure 5, the models are with an average KGEm of 0.64 in validation without EnKF. This is a quite low value. When the hydrological uncertainty is dominant, it is difficult to analyze the effect from precipitation post-processor. So, the boxplot for system C with or without post-processor for lead time 1 day in Figure 11 is similar. It is not indicated that the precipitation post-processor brings in no improvements. The improvements probably are too minor to offset the hydrological errors.

3. The author's language usage was difficult to read at times. Too many adverbial clauses and attributive clauses make the sentence too long to understand.

4. In Section 2.1.1, the authors reduced the ECMWF database to $0.1°$ and then spatially averaged forecasts to the catchment scale. I am confused by the resolution reduction as the catchment areal forecasts were used. It is more simple to use the archived $0.25°$ to calculate the catchment average. The resolution reduction might bring in additional

uncertainty to precipitation forecasts.

5. Line 170-171: The authors mentioned they used an adapted CSGD to post-process ECMWF precipitation. Whether the only difference is that the original CSGD used neighboring grid points while they used all grids in a catchment?

6. Line 221-222: the same….as….

7. Line 506-509: NWP products often fail to capture precipitation forecasts in small domain, yet behave better in large catchments. Lumped hydrological models are more likely to better model streamflow at small basins, where the hydrological process is simpler and easier to be simulated by those simple lumped models.

8. Since the authors elaborately analyze the performance for different catchment size in Section 3.4, they should provide detailed validation results of the hydrological models in different catchments in Section 3.1. It would help the readers to understand the forecast skill over catchment sizes.