We sincerely appreciate Referee 1 for the review of the paper "Sensitivity of hydrological model to the temporal and spatial resolutions of rainfall input". We have considered the reviewer's comments and will revise our manuscript according to the suggestions. The detailed answers to the comments are presented as below.

Major Comments

Introduction – The introduction is quite short and I don't think gives the reader a thorough overview of previous literature on this topic and where this research sits within the field. There have been lots of other studies that have focused on the impacts of spatial and temporal resolution of rainfall on hydrological model output and you need to clearly explain how your research builds on these previous studies. I found it difficult to identify from the introduction what the research gap was and how this study addressed that research gap.

Response: Thank you for the comments. We will rewrite the literature review for the manuscript. The revised version will contain an updated introduction, referring to the ongoing progress of the study for the sensitivity analysis of rainfall data to model performance both on temporal and spatial scales. We will describe in more details about the attempts for improving model performance and the monition of our study. We will also compare and discuss our idea with previous work on impacts of input variables in hydrological models.

Study area and hydrometeorological datasets – The rationale for your choice of catchments needs to be outlined. Why were these four catchments chosen? Do they have different climatological characteristics that make them interestingly different? A lot of the following analysis focuses on differences between these mesoscale catchments so it is important that the reader understands what these key differences are. Table 1 contained some interesting catchment characteristics but then these were not further explained.

Response: Thank you for the constructive comments. According to available flow records, four upstream catchments which are minimally impacted by human influences were considered in this study. These four catchments ranging in size from 417 km^2 to about 1300 km^2, along with a large difference in elevation and annual precipitation. Meanwhile, the map for raingauge locations also shows different observation density for them. We will further describe the study catchments in the revised paper.

Performance criteria – The choice of performance criteria needs to be better justified as this has a large impact on the sensitivity of your results.

Response: We agree that model performance depends strongly on the performance criteria used in calibration. In our previous study, we compared the lumped HBV model performance for difference objective functions in a number of catchments on daily scale. Three criteria: (1) the Nash-Sutcliffe (*NS*), (2) Kling-Gupta efficiency (*GK*) that accounts for the water balances and the correlation of observed and simulated discharge series(Gupta et al., 2009), (3) the combination of *NS* and the *NS* of logarithm of the discharge (*NS+LNS*), were used to calibrate the HBV for 15 catchments(Bárdossy et al., 2016). The model parameters calibrated for every catchment were used to simulate the remaining 14 catchments for testing the transferability of parameters. As shown in the figure below, results for different performance criteria differ considerably. The difference of model performance for the performance measures can be explained by different focuses: *NS* is mainly focusing on high flows as it represents the squared difference between the observed and discharge series, *GK* focuses on water balances and good timing, and *NS+LNS* criterion is strongly influenced by low flow events. Model behavior is dependent on how one evaluates the performance of the model. From the matrix we could find that the model performance for different criteria shows similar treads. In this study, we hope to investigate the sensitivity of model to the input variables and sequentially find effective way for increasing accuracy of flood prediction. We pay for attention to high flows, therefore *NS* was selected as objective function to evaluate model performance. We will add the discussion of the choice of performance measures in conclusion.

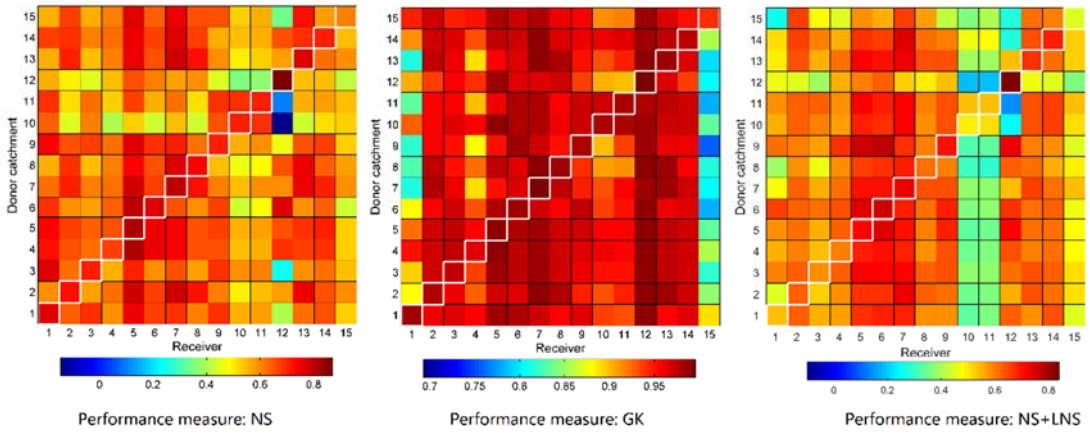

Figure. Color-coded matrices for the model performance of parameter transfer for 15 catchments using three difference performance criteria.

Minor Comments

Abstract P1 L6 'Two different flavors of HBV' – this doesn't make sense to me. It would be better to just say two different formulations or types.

Response: Thanks for the suggestion. We replaced "Two different flavors of HBV" with "two different types of HBV".

P3 L20 'illustrates the frame of these four datasets' – again, this sentence doesn't make sense to me and needs rewriting.

Response: Revision made. We replaced this sentence with "Figure 3 shows the flow chart of the data collection and process".

Figure 6 As you are focusing on higher flows, I would also find it useful to have another plot (or combined with Figure 6) that focuses on the flow duration curve for flows higher than the 10th percentile of flow.

Response: We appreciate the referee's suggestion, and will add the flow duration curve for flows higher than the 10th percentile of flow (Figure 6(b)) in the revised manuscript.

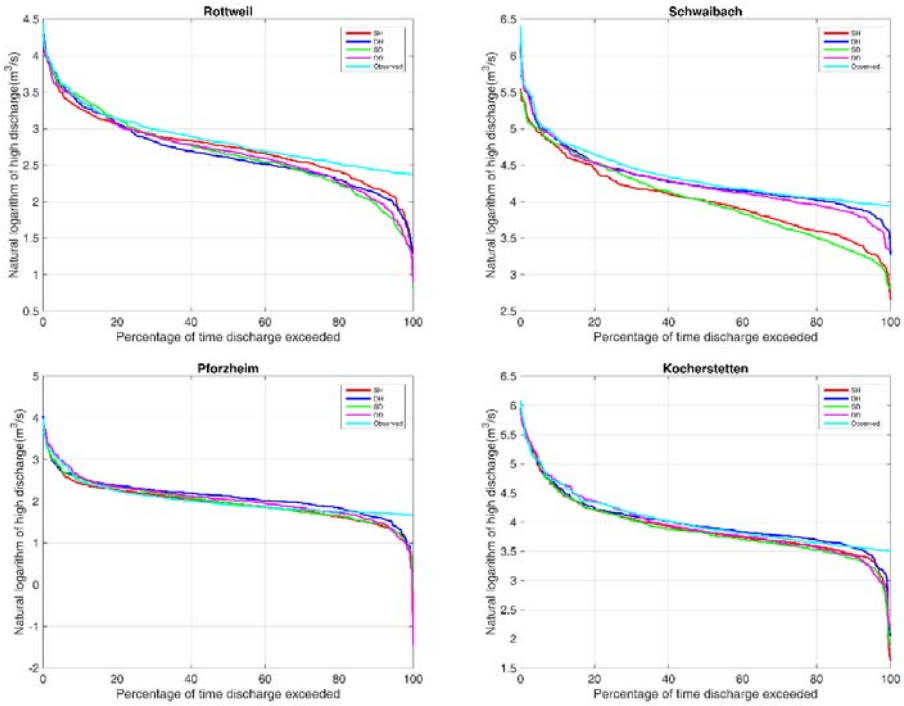

Figure 6(b). Comparison of the flow duration curve for flows higher than the 10th percentile of flow.

Figures 7 -10 need some improvement. The colour scheme needs to be changed in these plots so it is easier for the reader to distinguish between the different catchments. Currently it is difficult to pick out differences between catchments.

Response: Thanks for the suggestion. We have changed the colour of the plots for Figure 7-10 and Figure 12 to make a clear distinction between catchments.

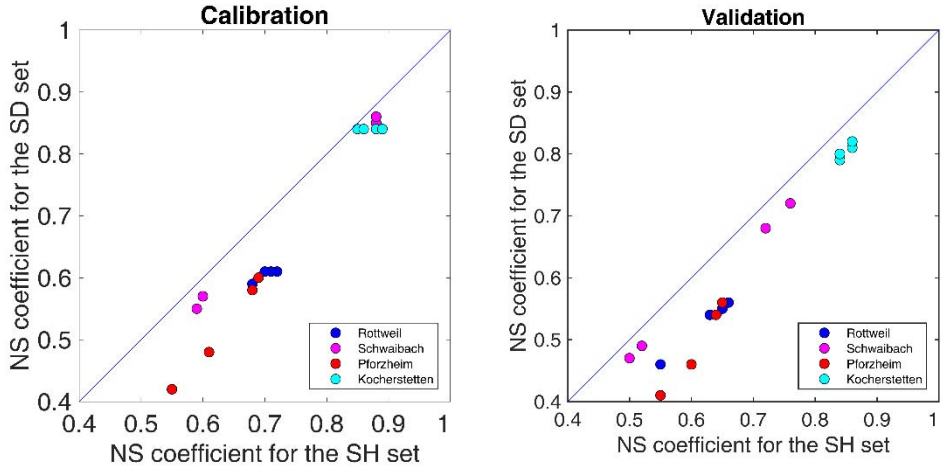

Figure 7. Comparison of NS model performance for using hourly and daily variables as model input for the SH and SD sets.

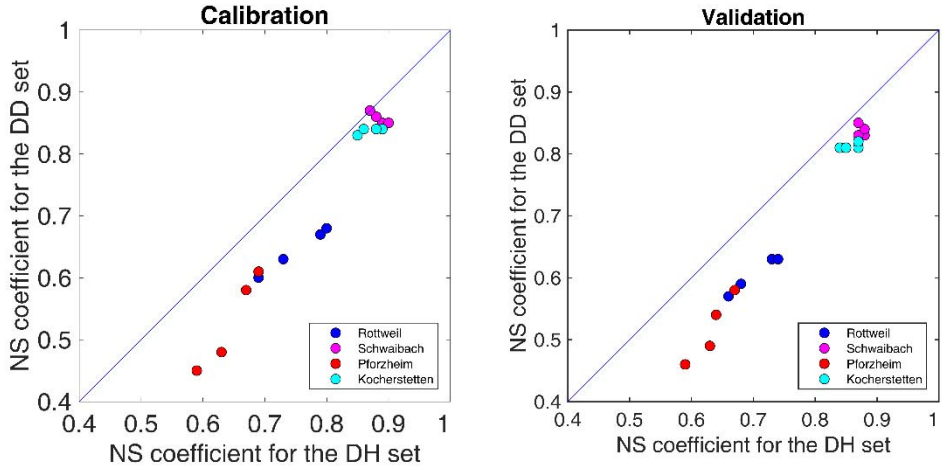

Figure 8. Comparison of NS model performance for using hourly and daily variables as model input for the DH and DD sets.

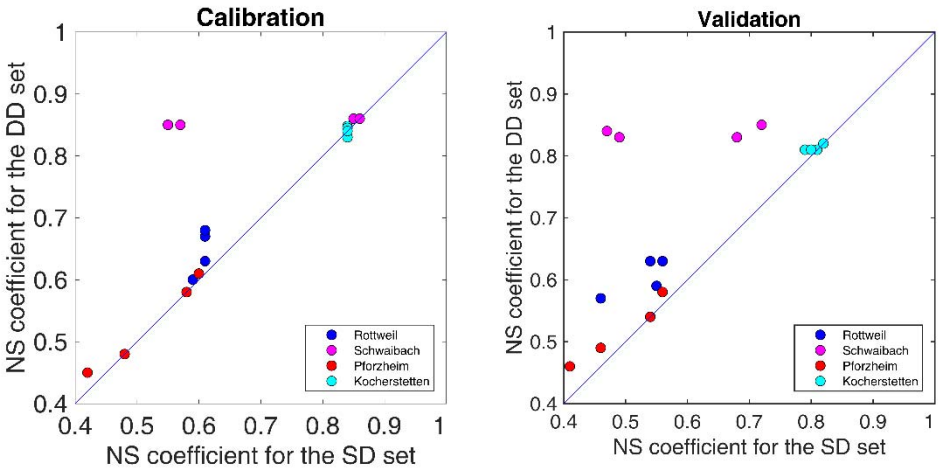

Figure 9. Comparison of model performance for different density of rainfall observation network, models were simulated based on daily time step.

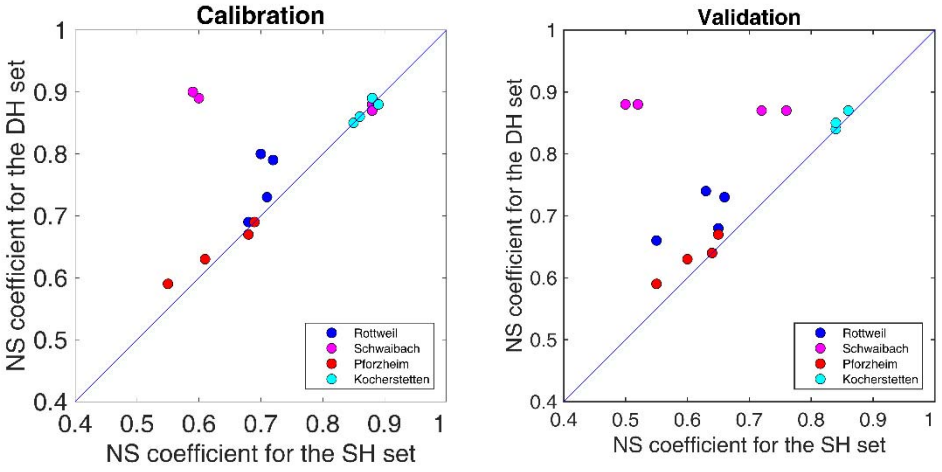

Figure 10. Comparison of model performance for different density of rainfall observation network, models were simulated based on hourly time step.

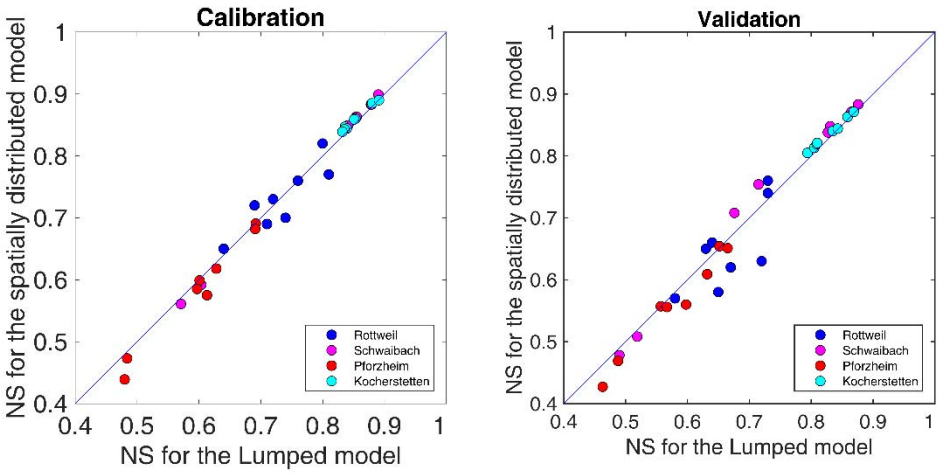

Figure 12. Comparison of model performance for different spatial resolution of model structure.

References:

Bárdossy, A., Huang, Y., and Wagener, T.: Simultaneous calibration of hydrological models in geographical space, Hydrology & Earth System Sciences Discussions, 12, 11223-11268, 2016.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80-91, 2009.