Dear Authors,
Thank you for your contribution to HESS. However, there are still several questions need to be clarified.

(1) As is suggested by reviewer #2, two watersheds, typical though, are far from enough. If this manuscript is designed to only provide an idea or assumption on "short period calibration of hydrological models", it does not contribute a lot to the hydrological community as many previous researches have already done this using different models though (same comments as reviewer #2). Therefore, I think you may conduct few more case studies, test it and make it convincing. Or if this manuscript focuses on short period calibration of "Physically-based distributed hydrological model", a discussion on the calibration of both conceptual models and physical models is expected. And also their differences in calibration should be discussed. In addition, what can we learn from this idea and how can we use this idea in the calibration of hydrological models for data-scarce basins?

(2) In your responses to reviewer #1, as is concluded that the good model performance for the validation period "may come from the fact that, the three validation years (2006, 2007, and 2008) are all wet years. And two of the three calibration years (2003, 2005) are also wet years. Calibration data may contain sufficient information for parameters identification in wet years." If in this case, the model is useful only when the calibration period and validation period both are wet years or both are dry years. What can we do when only one year's data is available for a data-scarce watershed? We cannot tell if this year is as the same condition as the following/previous years. Therefore, the objective of this study "how the use of limited continuous daily streamflow data might support the application of a physically-based distributed model in data-sparse basins" is not well explained.

Technical questions:

1) I don't agree that "However, none of these observations has the capability of streamflow data for constraining hydrological model parameters." The soil moisture, ground water level can also be used in hydrological calibration and constraining model parameters (e.g., Immerzeel and Droogers. 2008. Calibration of a distributed hydrological model based on satellite evapotranspiration. Journal of Hydrology, 349(3): 411-424.).
2) Through I agree with you that "it is difficult to apply measured values to hydrological models directly", I don't think these parameters are conceptual without explicit physical meaning (e.g., CH_K2, SOL_AWC, SOL_K). They have explicit physical meanings.
3) Check the language throughout this manuscript. For example, Line 14: period(s); Line 22: year(s)…
4) What are the criteria to distinguish a good model performance? How to tell the model performance (e.g., calibrated using one-month/three-month streamflow data) is as good as the one calibrated using 3-year? The meaning of U combining the P_factor with the R_factor should be clarified. For example, the model performance of "one year 2003 or 2004" is not superior significantly than that of "one month" in Table 5.
6) About the GLUE equation, your approach calculates the posterior distribution at each point first and then gets the 2.5% and 97.5%, while reviewer's approach computes the 2.5%

and 97.5% directly. I think both methods are correct.

(7) Location of these two watersheds should be identified in Fig. 1. And Figure 1 and Figure 2 can be merged into Figure 1. The hydrological stations used in this study should be included.

(8) Figure 6: the unit should follow “HESS author instruction”.