**Response to short comment of Dr. Julian Koch on the manuscript "Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products"**

**Patricia López López, Edwin H. Sutanudjaja, Jaap Schellekens, Geert Sterk and Marc F. P. Bierkens**

The authors would like to thank Dr. Julian Koch for his time and useful comment on the manuscript. His suggestion will help us to improve the quality of our manuscript. We have included detailed responses to his comment in the supplementary .pdf file. We have also included a modified version of the original manuscript. Please note the supplement to this comment and the modified manuscript, with modifications in blue.

## Short comment:

*This study caught my interest because it aims at improving discharge predictions in data scare areas using remote sensing (RS) data of evapotranspiration (ET) and soil moisture (SM) as calibration targets to inform model parameters of a distributed model.*
*I think the work is nicely presented and a very relevant contribution to HESS. However, I would like to discuss one aspect, which the authors did not touch upon. This may inspire the authors when revising the manuscript.*
*I would like to ask the authors if they regard spatial averaging of RS data as the optimal way to utilize these data? If I understand correctly from Figures 8 & 9, the authors used timeseries of averaged ET and SM as calibration targets. This neglects the valuable spatial information which is contained in the RS data. One could imaging that an additional metric, which is targeted at the spatial patterns of ET and SM may improve the discharge performance from calibration scenarios 2-4. A pattern oriented calibration can potentially yield more realistic parameter fields which are necessary to be able to simulate the spatial variability of runoff generation within the catchment in a more realistic manner.*
*Ultimately, if the spatial variability of hydrological processes is not of concern and the model is simply calibrated, evaluated and used at aggregated scale, then why not just use a simple lumped model? I think the discussion of the manuscript at hand would be significantly improved by showing and discussing simulated and observed spatial patterns of ET and SM.*

## Answer:

As Dr. Koch pointed out in his comment, we used average actual evapotranspiration and surface soil moisture over the entire Oum Er Rbia basin to calibrate model parameters under five different scenarios. With this approach, we take into account the spatial variability of the catchment indirectly through the model parameters that are different for each grid cell. However, we agree with Dr. Koch that other calibration approaches could be applied to better use the spatial information from GLEAM and ESA CCI products. Therefore, we carried out a new calibration scenario based on ESA CCI surface soil moisture, but instead of comparing the basin average (calibration scenario S3), we compared soil moisture estimates and observations per grid cell (calibration scenario S3 pixel). This calibration approach allows the identifiability of the optimal prefactors set for each grid cell. This means that there is not a single spatially uniform prefactor (e.g. $f_w$ = 1.25), but a map of different prefactor values depending on the location / grid cell (e.g. $f_w$ ($lat_1$, $lon_1$) = 1.25, $f_w$ ($lat_2$, $lon_2$) = 1, etc.). A similar calibration strategy was described in Sutanudaja et al. (2014).

Below, we included two figures (Figure RC1 and RC2) to summarize the results obtained with this new calibration scenario based on the pixel comparison of soil moisture model estimates and ESA CCI observations.

In Figure RC1 we plotted the prefactor values identified over the Oum Er Rbia basin in different columns (1st column: $f_e$, 2nd column: $f_j$, 3rd colum: $f_k$ and 4th column: $f_w$) for the three global precipitation products in different rows (top: EI, bottom: WFDEI and middle: MSWEP).

From Figure RC1, there is not a clear spatial pattern for each prefactor values. For example, $f_k$ values vary non-uniformly over the basin, with values of 0.25 in a location next to others with values of -0.25 and 0. This raises the question of how to regionalize the prefactor values and therefore, the model parameters. Further work would be needed to investigate novel strategies of regionalization based on land use, soil characteristics, climatic zones, etc.

In Figure RC2 we plotted discharge estimates and observations at Mechra Eddahk when MSWEP is used as model forcing. The red dashed lines represent discharge estimates from calibration scenario S0 (a) and S1 (b) and the purple dashed lines represent the calibrated time series from calibration scenario S3 pixel.

Similar to results obtained with the other calibration scenarios, discharge estimates reproduce ground observations quite well. Using ESA CCI soil moisture pixel by pixel for calibration (S3 pixel) improves the discharge estimates from 0.648 (S0, reference run) to 0.661 $KGE_q$ values (Figure RC2a). However, calibrating based on in-situ discharge observations (S1) leads to a further improvement to 0.680 (Figure RC2b). If we compare results of Figure RC2 with results of Figures 7 and 8 in the manuscript, $KGE_q$ values of 0.710 are reached when ESA CCI soil moisture and GLEAM actual evapotranspiration are used in the step-wise calibration approach.

With this new calibration scenario (S3 pixel), we are limiting the comparison to soil moisture estimates and observations pixel by pixel and problems of over-parameterization may occur, i.e. obtaining similar model results with more than one parameters combination. A possible route to overcome this problem could be using new spatial performance metrics to select the best parameters set (Koch et al., 2017).

We believe that the present experiment is useful and gives further insight into how to incorporate spatial information of variables such as ET and SM for calibration of spatially distributed hydrological models. However, we think that including these results in the manuscript may hamper the comprehension of the overall work. Therefore, we will discuss these aspects in section 5. Discussion and conclusions as follows:

P17L5: "… The spatial information of these satellite-based products could be used in a different way than the one explained in this study. For example, a calibration scenario based on a pixel by pixel, instead of basin average, comparison of surface soil moisture and actual evapotranspiration model estimates and observations could further improve discharge estimates. This calibration approach would have into account the spatial variability of the variables over the basin. Previous studies investigate how to incorporate spatial information into hydrological models using innovative spatial performance metrics to analyse the spatial sensitivity of simulated land-surface patterns (Koch et al., 2017). …"

## References:


- Sutanudjaja, E. H., Van Beek, L. P. H., De Jong, S. M., Van Geer, F. C., & Bierkens, M. F. P. (2014). Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data. *Water Resources Research*, *50*(1), 687-705.

- Koch, J., Mendiguren, G., Mariethoz, G., & Stisen, S. (2017). Spatial sensitivity analysis of simulated land-surface patterns in a catchment model using a set of innovative spatial performance metrics. Journal of Hydrometeorology, (2017).

## Figures:

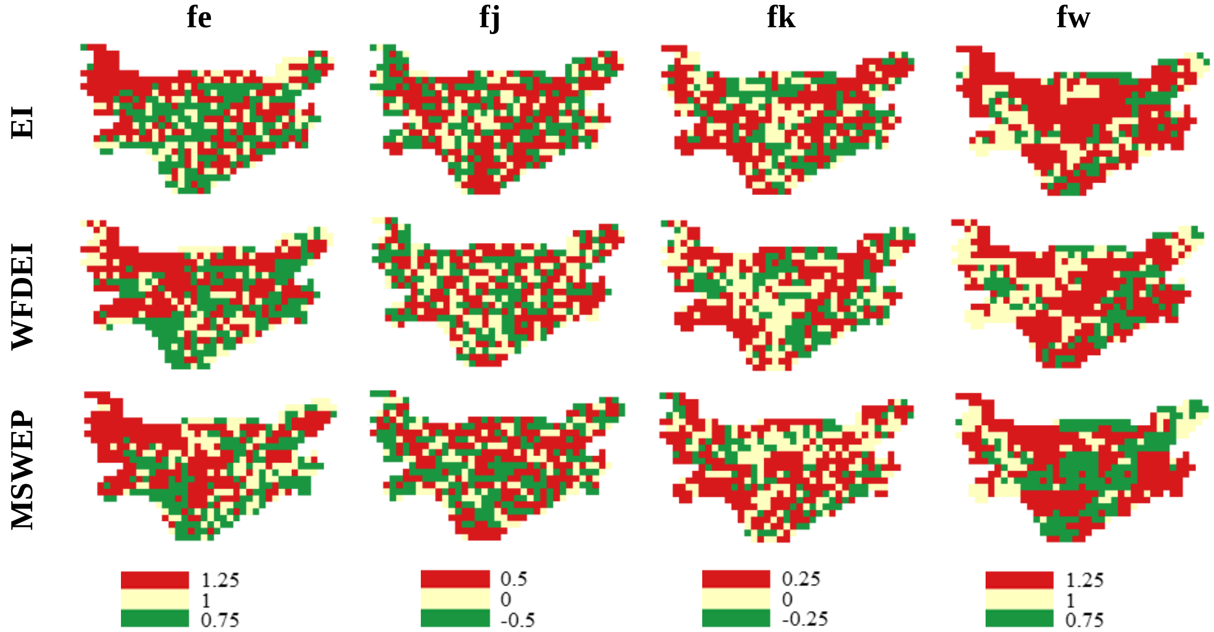


***Figure RC1.*** *Prefactors values identified with calibration scenario S3 pixel. Columns indicate different prefactors and rows indicate different global precipitation products.*

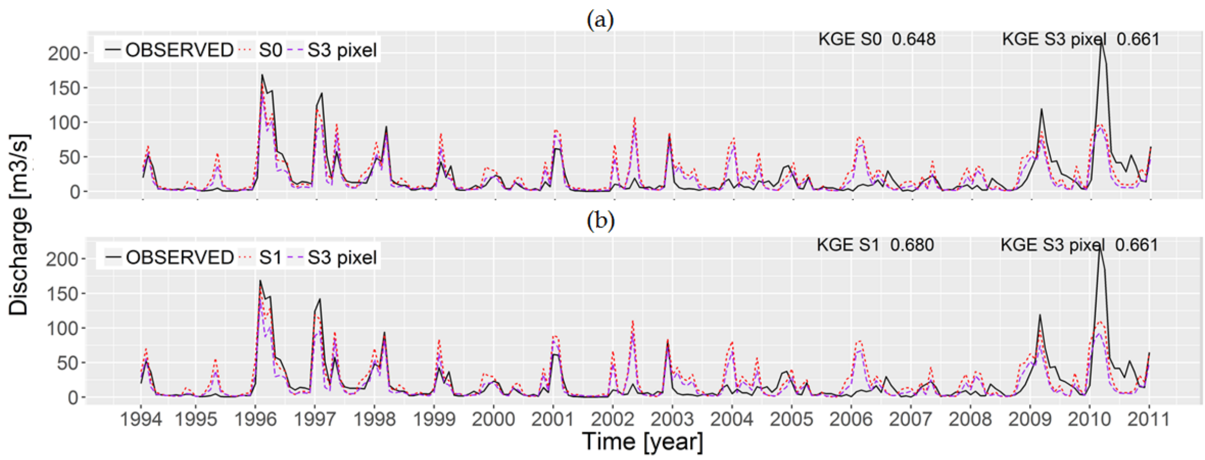


***Figure RC2.*** *Comparisons between monthly observed discharge (black) and estimated discharge (red and purple) time series at Mechra Eddahk when PCR-GLOBWB is forced with MSWEP precipitation. The red dashed lines represent discharge estimates from calibration scenario S0 (a) and S1 (b) and the purple dashed lines represent the calibrated time series from calibration scenario S3 pixel.*