Reply to Reviewer Comments

Reviewer #3

Review of the Article

Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model by Mehmet C. Demirel, Juliane Mai, Gorka Mendiguren, Julian Koch, Luis Samaniego, Simon Stisen

The paper addresses the relevant topic of using spatial data for the calibration of distributed hydrological models. The paper explores the tradeoffs between streamflow-based and spatial-based calibrations, illustrating the benefits of combining separate observation types and objective functions. Given the increasing availability of spatial data, to understand the value of a spatial calibration versus a more traditional calibration is a timing and relevant research topic. Spatial pattern of actual evapotranspiration (AET), obtained with the mesoscale Hydrologic Model (mHM), are calibrated against spatial patterns of AET estimated trough remote sensing (TSEB model). The major novel point of the paper is to present novel bias-insensitive spatial pattern metric, which exploits the key information contained in the observed patterns, useful for spatially distributed models optimization. I found the paper valuable and worth to be published. However, I believe some methodological aspects should be further clarified, as detailed in the specific comments.

I suggest a minor revision for the paper. Please revise also carefully English language.

Reply from authors: We wish to thank Dr. Giacomo Bertoldi (Reviewer #3) for the very constructive comments and suggestions.

General comments:

1.In the paper formulates the hypothesis that "The current calibration framework builds on the assumption that the satellite based estimate of AET patterns approximate an observed pattern that is suitable for model optimization". In my advice, this hypothesis should be better motivated and the implications better discussed.

At the end, six monthly mean "climatological" AET maps have been produced. But spatial patterns are time consistent? If the spatial structure of AET changes a lot from time to time, or there is a great inter-annual variability, then it makes little sense to calibrate the model to such patterns. A preliminary analysis on the temporal variability of the observed AET patterns should be done. Only if spatial AET patterns are quite constant over the different years, the proposed calibration approach could be used. Moreover, the fact that an AET monthly "climatology" has been used, could be one reason of the limited tradeoff found with the discharge. Could you discuss about this? In this way, you cannot for example identify from AET observation a drought period, which can have an effect on discharge.

I see later that this choice is explicit in the paper (see 4.1 Objective functions). “Since all three terms are bias-insensitive, the spatial efficiency only constrains the model simulations with the pattern information in the satellite data while leaving the water balance (bias) to be constrained by streamflow”. However, this implies a hypothesis of spatial invariance of AET patterns. You need to support and discuss this assumption in the paper.

Reply from authors: We will elaborate on the assumption that the RS based AET data contains valuable pattern information that can help constrain the hydrological model regarding simulation of spatial patterns. We did a preliminary analysis of how consistent the spatial patterns were in time. They turned out to be quite consistent, this we will include in the final reply and potentially add to the manuscript as either a discussion or added figure/table. The overall idea is to utilize the most informative part of the observation data (spatial patterns from RS AET maps and water balance and temporal pattern from discharge). In addition, we are aiming at calibrating mainly the spatial parametrization of soil and vegetation parameters using the bias-insensitive spatial pattern metric, these parameters typically do not vary in time (expect in some cases vegetation, like in our case vegetation dynamics change, but the vegetation related calibration parameters are time-invariant). Therefore, we believe that the temporally averaged spatial patterns are a reasonable way of constraining the spatial model parametrization. In addition, a spatially consistent drought would be captured by the discharge observations; however, admittedly our approach is not designed to capture spatial variability in drought patterns (this would especially be relevant for larger catchments).

2. Another assumption is that “Bias and temporal variability of satellite derived AET estimates could be useful for model optimization, however, in this study, we deliberately limited the information content of the satellite data to address the spatial patterns.” You clearly separate in your calibration framework the source of temporal information (discharge) from the source of spatial information (AET). It seems at the end that this second part is not really a calibration framework, since you even introduce some modifications in the model structure to properly incorporate spatial information in modelled AET. It seems to me that the main outcome of the paper is an interesting way to integrate spatially distributed observations in a modeling framework. It is a kind of spatial data assimilation, or better spatial data integration approach, more than a traditional model´s parameters calibration. I suggest in the Introduction and in the Discussion to place your results in the broader perspective of the data integration techniques and not only model calibration literature, and underline this potential of your proposed approach.

Reply from authors: I see the point; however, I still regard this as a classical model parameter calibration approach different from data assimilation schemes. The reason is that we do not update states, fluxes or parameters directly to match the observed pattern, but only use the observed spatial patterns to formulate a classical objective function that is then minimized by changing global parameter values. Some of these global parameters are then part of an upscaling operator or transfer function, but they are essentially still model parameters. In the revised manuscript, we will make a distinction between our approach and data assimilation.

Specific comments:

Introduction

It is important to underline that actual evapotranspiration (AET) estimates from satellites are not observations, but the results of a model (in this case the TSEB). This model shares part of the input information used by the hydrological model (i.e. LAI, fractional vegetation cover, meteorological data as temperature). This does not influence paper´s results, but it is important to clearly state that AET patterns are not observations, but the results of a remote-sensing based model, with a lot of uncertainties.

Reply from authors: True, this should be made very clear in the revised manuscript.

2.2 Satellite based data

Canopy fraction is a key input information of the TSEB model. How it has been estimated? AET estimations of the TSEB model are relative to the instant of the thermal remote sensing image used. How AET estimations have been extended to the daily and monthly time scale?

Reply from authors: Details of this are given in: Mendiguren, G., Koch, J., and Stisen, S.: Spatial pattern evaluation of a calibrated national hydrological model – a remote-sensing-based diagnostic approach, Hydrol. Earth Syst. Sci., 21, 5987-6005, https://doi.org/10.5194/hess-21-5987-2017, 2017. We will make sure reference is given in the manuscript.

Figure 1 TSEB results are usually highly dependent on LST and canopy fraction maps. It is very interesting how the RS-AET map reflects also well the soil type map, since no soil or soil water content information is given to the TSEB. Your hypothesis to link soil type to root density and therefore to root water extraction is interesting. However other hypotheses could be done. The simplest is just the areal with lower sand content are wetter, because the soil has a higher water holding capacity, and therefore there is a higher AET. Could you comment on this?

Reply from authors: Yes, our initial thought was that simply soil physical properties could explain the differences. However, this was not enough to explain the differences seen in the RS AET maps. In addition, we had specific knowledge from the literature about the root density variations in this region. Also the sensitivity of root density to AET is very high in the model (and in many other PET based AET models as well), so getting the soil influence on AET right while preserving the original uniform root density for all non-forest areas would not be sufficient or correct.

Distributed root fraction coefficient. The assumed model could be better described. Why land cover does not influence root fraction? I expect that different vegetation types are the main source of the different root fractions. Is the root depth a relevant parameter in the mHM model or only root fraction is used?

Reply from authors: mHM uses root fraction distribution, and land cover is the main driver for differences in root fraction, but it is simplified to forest, non-forest vegetation and impervious (the later not being used in the current study due to very minor area of urbanization). The soil dependency we introduce is only to further distribute the root density across non-forest vegetation (agricultural cropland). The main differences in root density is still between forest and non-forest. This will be made a bit clearer in the manuscript.

4.1 Objective functions. The application of the new SPAEF index is one of the main elements of novelty in the paper. A schematic figure explaining the concept of “histogram intersection” could help. It is not obvious for me.

Reply from authors: The histogram intersect will be explained better.

Figure 3. Why including AET patterns reduces model performances in the summer months? This is the time when I expect AET counts more for the water budget. Is this a drawback of the choice of not considering biases in AET? What would happen if you include also as calibration target the temporal evolution of the spatial average of AET?

Reply from authors: This is probably a drawback, and the loss in performance in discharge during summer months is seen as a tradeoff resulting from adding another objective function.

Figure 4. This figure shows the potential of this technique of spatial calibration more in terms of remote sensing data integration, than in terms of calibration. In fact, in many cases distributed hydrological models cannot produce detailed spatial patterns, because the coarse spatial patterns of models input data (except topography, soil and land cover types are usually known at a coarse spatial scale). Could you comment in this? (see also general comments)

Reply from authors: As replied above, we regard this as a classical parameter calibration exercise. However, the flexibility in the spatial model parametrization and the use of distributed input (with *same resolutions as the model grid) of soil texture and vegetation enables the model to approach a detailed spatial pattern. It should be noted that the results of the Q-only calibration in figure 4 has exactly the same parametrization scheme and free model parameters as the two other calibrations, so the only reason that it fails in reproducing the spatial pattern is that it is not informed about what the "true" pattern is.*

Page 16 Results show how combining Q and AET results in a more robust model parametrization for the validation phase. Could this also contribute to reduce models equifinality?

Reply from authors: We hope so; however, we have not addressed equifinality as such.

Page 16 "It is recognized that traditional downstream discharge measurements contain much more accurate and robust information on the overall water balance compared to the non-continuous remotely sensed estimates". Please provide some references to support this.

Reply from authors: Not sure there are specific references for this quite specific statement. We are here arguing that a continuous time-series of total catchment discharge contains more information on the overall water balance compared to a series instantaneous snapshots of AET for only cloud free conditions. We will rephrase to make this our own statement, unless we find an appropriate reference.

Discussion

Page 17 "Here, we ignore the temporal aspect and focus only on the consistent spatial patterns for each month of the growing season" "The current calibration framework builds on the assumption that the satellite based estimate of AET patterns approximate an observed patterns that is suitable for model optimization". Please motivates better this assumption. See general comments above.

Reply from authors: We will motivate this better in the revised manuscript.

The choice of not including temporal information from the AET data is an assumption quite well justified in the discussion. However, this could reduce the information amount coming from AET.

Reply from authors: true, we could miss information actually contained in the AET data, however, as explained throughout the manuscript, we focus on the information content of the RS AET that supplements the discharge time series the best.

“AET estimate is validated against eddy covariance stations (Mendiguren et al., 2017) they only represent specific cloud-free days limiting their value to assess the long term water balance of the catchment” OK, but if AET patters in cloudy days are different? What happens if AET patterns are not stationary in time? From my experience with distributed models, looking to daily model´s AET output maps, I´ve been ever surprised how such patterns are changing from day to day. See also general comments. A preliminary analysis of the time series of the RS-AET map should be done to assess how much time-consistent are the patterns.

Reply from authors: This is definitely a limitation to our approach. But again, we have decided to utilize what we assume is the best information content available from the AET data. This is the spatial pattern on the days where we have LST observations. We could have interpolated these cloud-free “observations” to cloudy and rainy days, however this would not increase our number of “true” observations. In addition, for the current catchment, the impact of soil and vegetation parameters (which we calibrate) on AET on very cloudy and rainy days will be very limited, because in this region, potential Evapotranspiration will occur as soon as the potential evapotranspiration becomes low (as will happen on cloudy and rainy days).