Reply to Reviewer Comments

Reviewer #1

Comment 1) This paper deals with an innovative calibration framework that combines temporally aggregated observed spatial patterns with a new spatial performance metric and a flexible spatial parameterisation scheme. An application of the mesoscale Hydrologic Model (mHM) to the Skjern River Basin is used as an example show the effectiveness of the presented calibration framework. This is a very timely topic that fits very well to the scope of HESS.

However, the spatial model parameterizing methodology is not well described. Both the root fraction coefficient and the PET correction factor parameterizations should be presented in more detail including graphical presentations of the underlying relationships. For these reasons, my recommendation is to accept this manuscript with minor revisions.

I have provided specific comments and suggestions for improvement below.

Reply from authors: The authors thank the Dr Heye Bogena (reviewer #1) for his constructive comments on the manuscript. We have replied to each comment below.

Specific comments:

Comment 2) P3L7: Which models?

Reply from authors: "Immerzeel and Droogers (2008) showed that the models can be constrained by using spatially distributed observations with a monthly temporal resolution" is replaced by "Immerzeel and Droogers (2008) showed how a semi distributed model of a basin in Southern India could be constrained by using spatially distributed observations with a monthly temporal resolution"

Comment 3) P3L16-32: This section is redundant with the method section and should be shortened.

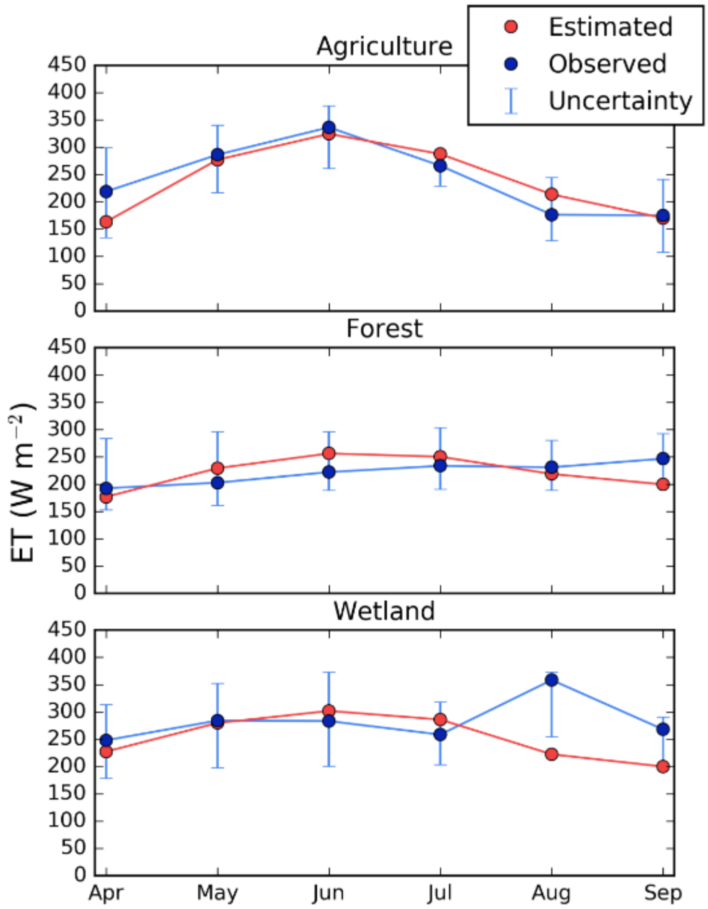
Reply from authors: We agree with the comment. This section will be shortened in the revised version of the manuscript.

Comment 4) P5L6 Delete “s”

Reply from authors: “s“ is removed.

Comment 5) P5L9-11: How accurate are the AET maps, e.g. in relation to continuous EC-measurements?

Reply from authors: The accuracy of the AET maps has not been quantified in this paper. However, a comparison to observed AET for three EC sites has been included in a previous paper Mendiguren et al 2017 HESS. This comparison focussed on the monthly mean AET estimates for three land-cover types, which were well in agreement with the measurements, although it has to be noted that the measurements themselves were subject to uncertainty due to energy balance closure issues. Below is the figure that appears in the Mendiguren et al. 2017 paper



However, in the current study, the accuracy of the AET maps cannot be completely assessed by a comparison to EC measurements, because only the bias insensitive pattern information is utilized, and the uncertainty of this pattern cannot be fully described by just three EC sites. The *current study therefore relies on an assumption that when performing well for monthly means at three different sites and being mainly driven by remote sensing observations of LST and NDVI/LAI, the satellite based AET estimates contain spatial pattern information that is suitable for constraining the spatial pattern simulations of our distributed model.*

Comment 6) P5L30-31: Why should this procedure accelerate model runs?

Reply from authors: The file size of the 1x1km inputs is much larger than 10 or 20 km. Therefore, reading these files to the model at each timestep takes very long time as compared to the discretisation algorithm defined inside the Fortran code. This is a technicality related to either pre-processing the LAI-based input on a daily scale prior to model execution or reading the monthly data and disaggregate to daily data inside mHM. We suggest skipping this line altogether.

Comment 7) P6L8: “stretch the spatial contrast“ of what?

Reply from authors: Corrected as “spatial contrast of simulated actual evapotranspiration” in the revised version of the manuscript.

Comment 8) P6L8: “: : :based on soil and vegetation properties: : :”

Reply from authors: Corrected.

Comment 9) P6L11: Please change “domain-specific” into “site-specific” or “local”

Reply from authors: Corrected throughout the manuscript.

Comment 10) P6L11-12: According Feddes et al. (2001) and others, the root fraction coefficient is a vegetation dependent coefficient. Please add more information on why your assumption that this coefficient can be explained by field capacity is justified.

Reply from authors: It is true that generally the literature suggests that root fraction coefficient is mainly dependant on vegetation type. We also separate root fraction coefficient in two main categories: forest and non-forest. However, for the non-forest (in this study agricultural cropland, since this is the only major land cover type) we have introduced an option to distribute root fraction coefficient based on soil type. This is based on several studies carried out by Danish soil scientists who showed that for the very sandy soils of Western Denmark the effective root depth is smaller relative to soils with similar crop types, but higher clay content, although this dependency seizes at a certain clay content. Instead of basing the root fraction *coefficient on soil type, we utilized the FC which in mHM is a function of sand and clay content through the internal pedo transfer functions. We admit that the relation might be specific to the Danish case, where even very poor (very sandy) soils are utilized for crop production and where the relation between root depth and soil type has been established for very sandy to loamy soils. A similar relation is used in the parametrisation of the Danish National Water Resources model and in National crop growth models. In the revised paper we will elaborate on this and make it clear that root depth is also a function of main landcover/vegetation types.*

Comment 11) P6L14-15: At this point it unclear why these parameterizations increase the model freedom. Please reorganise the text in a more comprehensively way.

Reply from authors: We will rephrase and organise this paragraph more thoroughly in the revised version of the manuscript.

Comment 12) P6L19-24: It took me some time to understand your method, also because the assumption that the root fraction is linked to FC is counterintuitive. For a better understanding of your method, it would be helpful to graphically show the relationship between FC and rooting depth based on the soil database.

Reply from authors: We will elaborate on this, but the relation builds on previous work by others, so we can refer to their studies and illustrate the relation graphically but not plot their actual data.

Comment 13) P6L22: Why do you restrict your method to pervious non-forested areas?

Reply from authors: Basically mHM operates with three land cover types: Forest, pervious and impervious (urban). The urban areas in our catchment is very limited, so we ignore that category, leaving just forest and non-forest. As explained above we acknowledge that the overall vegetation type (forest or cropland) is the dominating controlling factor for root depth, therefore we allow separate root fraction coefficients for forest and non-forest and only apply the FC dependency for non-forest.

Comment 14) P6L22: Please change “domain-specific” into “site-specific” or “local”

Reply from authors: Corrected.

Comment 15) P7L2: Why should soil properties of sandy soils impede root development?

Reply from authors: Impede is the word used in the literature we refer to on the root depth dependency on soil type.

Comment 16) P7L2-3: Which parameters and why is a fine vertical discretization more effective?

Reply from authors: This is related to the way the root fraction coefficient effects AET simulations. When a given soil layer dries out (SM falls below FC) the simulated AET is reduced linearly until it reaches zero at wilting point. Fine vertical discretization of the upper soil layers will allow for a finer representation of the root fraction variation with depth and result in more frequent occasions where SM content reduces AET. This might be a technicality, and the sentence could be omitted, since we used the same vertical discretization through all model runs in this study and we have not examined the impact of vertical discretization.

Comment 17) P7L5-7: By changing the root fraction parameters for maximum and minimum FC, the relationship between FC and rooting depth will be changed. In the extreme case, both parameters have the same value, which means that there is no relationship at all. Did you check whether the optimised model still provides realistic relationships between FC and rooting depth? In addition, it is unclear, how you derived FC values from your digital soil map. Typically, one would sum up the horizontal FC values down to the certain soil depth (e.g. rooting depth). Please provide additional information.

Reply from authors: For the latter question, The FC is estimated in the pre-processing within mHM based on pedo transfer functions. It is a fundamental part of the model concept that only soil texture data is specified and the soil physical properties are derived. This insures a spatially consistent parametrization (given the soil texture data is good) and reduces the calibration parameters to the global pedo transfer functions parameters. We used the same soil texture parametrization (and therefore FC) for the entire soil column. We have looked at the variation in both the derived FC and WP maps and the range of root fraction parameters and they all look reasonable. In addition, we have tied the minimum and maximum root fraction coefficient so they can approach each other but never reverse. If they approach each other through calibration and become identical, that would essentially indicate that the approach did not benefit the spatial pattern performance, but that is not what we see. It clearly improves the spatial pattern performance to separate root fractions based on FC.

Comment 18) P7L16-17: You should not mention equations that have not been already introduced.

Reply from authors: Corrected.

Comment 19) P7L21-25: The section needs to be rewritten in a more comprehensible way.

Reply from authors: We will reorganise this section in the revised version of the manuscript.

Comment 20) P7L26: It is how this equation was derived and why it should be "physically meaningful". In addition, it is unclear how the DSF parameter is used to correct ET_ref.

Reply from authors: We didn't derive this equation. It is a time-space variable implementation of the crop coefficient concept suggested by Allen et.al. The basis is that the climate dependant reference evapotranspiration is given at a coarse scale of 20 km grid for a reference crop. In reality, the vegetation does not reflect a reference crop everywhere. In order to correct the reference ET a scaling factor is applied which is above 1 if the evaporative potential of the vegetation is higher than for the reference crop and below 1 if it is lower. Typical values of crop coefficients are between 0.8 and 1.2. Allen et al. 1998 and others (eg. Hunink et al. (2017)) suggested using LAI or NDVI to estimate the crop coefficient.

Our implementation is simply using the same equation in combination with remote sensing based LAI to create a time-space variable correction factor to convert ETref to ETpot.

*The DFS is simply a multiplication factor. ETpot = DFS*ETref*

mHM originally contains an even simpler correction factor, which is spatially and temporally uniform (although it also has an option to include aspect in mountainous terrain). We have omitted this correction and implicitly included it in the calibration of the DFS, because the average DFS does not necessarily add up to 1. We will be more accurate in the description of the DFS in the revised manuscript.

Comment 21) P7L28: The previous sections also belong to "Methods".

Reply from authors: We don't regard the parametrization of the mHM model as a methodology as such, we are not changing the process descriptions of the model but merely adding spatial flexibility to the parametrization of exiting model parameters. Therefore we feel that it is more appropriate to limit the methodology section to the more novel parts of the manuscript, namely the development of a new spatial performance metric and a multi-objective calibration framework based on a complimentary principle.

Comment 22) P9L1: Only accepted paper should be used a reference.

Reply from authors: We agree with the comment.

Comment 23) P9L4-9: This section can be omitted.

Reply from authors: will be corrected

Comment 24) P9L24-26: Is this statement relevant for this work or Koch et al., 2017a

Reply from authors: This statement is relevant only for Koch et al (2017) and will be omitted.

Comment 25) P9L34: Why are you using the "same" cloud-free days?

Reply from authors: This perhaps should be rephrased; we want to explain that we are making the monthly average AET maps based on simulated daily AET by averaging only the days that were also available for estimating the RS based maps (cloud free days).

Comment 26) P10L5: Either use the Greek symbol for phi or “phi” (here and elsewhere)

Reply from authors: Corrected.

Comment 27) P10L22: Delete “and”

Reply from authors: Corrected.

Comment 28) P10L33-34: I wonder why the parameter “root fraction” for forest area is listed and not for the impervious non-forest areas, since the latter was used for the spatial parameterization.

Reply from authors: Impervious areas are very small and not represented in the land use map for root fraction distribution. Therefore, root fraction for impervious areas is not listed in Table 2.

Comment 29) P10L34: “(SPAEF column in Tab. 2)”

Reply from authors: Corrected.

Comment 30) P11L9-10: Why are you using a second model warm-up period (2005-2008)? This is also not mentioned in the text.

Reply from authors: The second warm-up period is just used to ensure that the validation results are not effected by initial conditions. The text will be updated accordingly in the revised version of the manuscript.

Comment 31) P12L18-20: However, it should be noted that the uncertainty of the spatial AET information from remote sensing is typically larger than the runoff measurements.

Reply from authors: We completely agree that the average AET estimate from RS is more uncertain that the runoff measurement. That is the reason why we use only the runoff to constrain the water balance, while we only use the bias insensitive pattern information of the

RS AET to constrain the simulated pattern. It is unclear to us exactly what to correct or rephrase here.

Comment 32) P12L28: Delete "that"

Reply from authors: Corrected.

Comment 33) P13L2-3: This indicates that the spatial AET information either has large uncertainties or that it has very limitated information on the subsurface properties due to extensive irrigation in the catchment. Please discuss.

Reply from authors: We disagree, the poor performance on discharge for the spatial-only calibration is a direct consequence of the way the SPAEF objective function is designed. Since it is bias-insensitive and only contains information on the spatial pattern it cannot be used to constrain the discharge and water balance. The Spatial only calibration is (as stated in the manuscript) not meaningful from a water balance or discharge perspective, it is solely included as a benchmark for the spatial pattern performance capability of the modelling framework and to illustrate the limited trade-offs between discharge and SPAEF when applied as proposed in the current study.

Comment 34) P15L13: Does Figure 4 present a certain year or monthly averages of several years?

Reply from authors: As mentioned in Page 5 line 6-10, Figure 4 presents averages of cloud free days for a specific month across all years for the model calibration period (2001-2008).

Comment 35) P15L14: The differences seem to vary also from month to month. You should quantify the differences in AET patters, e.g. by presenting the variances or variograms.

Reply from authors: What was meant by P15L14 was simply that the calibrations including the RS AET maps result in simulated patterns that has a much better representation of the variance compared to the Q-only calibration. Table 3 and 4 present month to month differences in SPAEF which also includes the coefficient of variation. It is unclear what a specific presentation of variances across months would add to the analysis, we are open to hear more on this from the reviewer.

Comment 36) P17L21: Awkward sentence. Please reformulate.

Reply from authors: Corrected as below.

“This is largely attributed to the nature of the metric as the spatial performance metric is bias-insensitive whereas the streamflow metrics have very little sensitivity to spatial redistribution of AET patterns as long as the spatial averages remain unchanged.”

Comment 37) P17L25: “This was: : :”

Reply from authors: Corrected as below.

“This was done because even though the satellite based 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.”

Comment 38) P617L31-32: Unclear why the sandy soil texture should restrict the model parametrisation method.

Reply from authors: What is meant here is that the root fraction coefficient dependency on soil type might be site-specific due to the uniform land use (agricultural cropland) across soils ranging from very coarse sandy soil to loamier soils. In other words, we do not claim that this parametrisation is generic, but it reflects knowledge of the particular area and fits well with the observed patterns in soil properties and RS AET as illustrated in Figure 1. We will rephrase this sentence in a revised manuscript.

Comment 39) P617L32: Please change “domain-specific” into “site-specific” or “local”

Reply from authors: Corrected.

Comment 40) P18L23: “has proven”

Reply from authors: Corrected.

Figures and Tables:

Comment 41) Figure 1: The RS-AET scale should be reversed (red should indicate high AET values)

Reply from authors: We disagree, we have used green/blue colours for high AET and red/brown for low throughout the manuscript. To us this is the most logical colouring scheme, which intuitively symbolizes wet and dry conditions.

Comment 42) Table 1: Columns 3, 4 and 5 should be removed as they provide only limited information, which is already presented in the text.

Reply from authors: Corrected.

Additional literature

Feddes, R.A., H. Hoff, M. Bruen, T. Dawson, P. de Rosnay, P. Dirmeyer, R.B. Jackson, P. Kabat, A. Kleidon, A. Lilly, and A.J. Pitman, 2001: Modeling Root Water Uptake in Hydrological and Climate Models. Bull. Amer. Meteor. Soc., 82, 2797–2809.