"Probabilistic inference of ecohydrological parameters using observations from point to satellite scales" by Maoya Bassiouni et al.

Response to David Dralle (Referee #3)

The authors pair in situ and remotely sensed soil moisture data with a Bayesian approach to infer parameters in a 1-d analytical model for soil moisture dynamics.

Thank you for your thorough review and constructive suggestions. We have provided responses and preliminary corrections below.

General Comments:

1) My primary concern is that the authors frequently claim "accurate" results, yet the study does not include any comparison between predicted and measured soil moisture thresholds. I would say that the study is more accurately described as an exercise in Bayesian model calibration. The novelty of the study, in my opinion, lies in comparing parameters of calibrated PDFs across observation scales. This is a useful exercise, though it's not fully explored in the study.

We agree that this study is primarily an exercise in Bayesian calibration of the commonly used stochastic soil water balance model. We explore whether a Bayesian inversion of the model can provide plausible estimates of ecohydrological parameters that are generally not directly measured. It is therefore challenging to compare estimated parameter values to site-specific observations and determine their accuracy because these are not directly available. We can however, as suggested in your minor comments, relate estimated parameters to calibration efforts of comparable parameters from a few previous studies. We will revise section 2.3.2. to explicitly state the evaluation goals and metrics used.

Optimal analytical soil saturation pdfs are evaluated by the following criteria.

(1) The Bayesian inversion converges and the Gelman-Rubin diagnostic approaches 1 for each estimated parameter (<1.1).
(2) There is goodness of fit between the optimum analytical pdf derived from the mean parameter estimates and the empirical pdfs derived from observations using the Kolmogorov-Smirnov (KS) statistic and the quantile level Nash-Sutcliffe efficiency (NSE) (Müller et al., 2016).
(3) Posterior distributions of parameter estimates are physically plausible and have low coefficients of variations.

A range of different sites was selected to develop and demonstrate methods in varying environmental conditions. However, the purpose of this study is not to compare estimated values at these sites. We limit the scope of this paper to presenting the model inversion methods and deriving criteria to obtain meaningful parameter estimates. A comparison of estimated parameters can be the focus of a future study in which a larger number of sites are considered and provide more insights on the variability of these ecohydrologic parameters. We will amend the statement of objectives in the introduction and our conclusions to clarify this scope and propose potential future applications.

The authors only go so far as to say that "spatial heterogeneity" explains shifting parameter values across scales. The significance of the study would be greatly increased if the authors worked to address some of these scaling effects. A couple questions include: How transferrable are inferred parameter values between scales? How might the optimal form of the PDF change across scales if heterogeneity is the culprit? And, are there simple in silico exercises that could be performed to explore these questions? For example, if the authors generate spatially correlated fields of soil moisture parameters and solve the 1-d model at each point, can aggregation explain (even qualitatively) observed trends in the inferred parameters? What are the implications for applications in sparsely monitored areas, or for making useful predictions at a point using remotely sensed data?

These are interesting questions that would require a different dataset. We will add a few sentences describing the potential of the proposed methods to address these questions in the discussion section and relate to recent references on scaling of the stochastic soil water balance model.

2) While I appreciate the authors' thoroughness, the inclusion of 6 distinct models for soil moisture dynamics somewhat obscures the paper's results. What intuition does this degree of added complexity provide, other than "model performance *increases when there are more parameters to tune"? Could some of these results be relegated to Supporting Information, keeping the two most illustrative models?*

We agree that this section has some information that may be obscuring the main message. We will keep all model alternative and revise the associated figure to only show the evaluation criteria defined section 2.3.2 (KS, NSE, convergence and parameter coefficient of variation). We will revise the description of these results to highlight the importance of the objective of this sensitivity analysis and clarify their relevance to the overall conclusions.

While model performance increases with model complexity, the risk of equifinality is also greater and the number of converging independent runs in the MH-MCMC is reduced when the number of parameters to tune is increased. The primary reason to evaluate models of increasing complexity was to detect the maximum number of parameters that can be estimated without the risk of equifinality and which are the minimum parameters that need to be fit to have acceptable goodness of fit between the empirical and analytical pdfs. We found that the 3 parameters s_w, s^*, and E_{max} were necessary and fixing E_w to a small value (5% of E_{max}) equifinality was removed. We will amend the discussion of this point in the revision.

3) The authors assume steady-state conditions for application of the stochastic models. While this may be appropriate for MMS and ARM, soil moisture dynamics at the seasonally dry sites Tonzi Ranch and Metolius are highly non-stationary during the dry season study months April – September. One can see this in the bi-modality of the soil moisture PDFs in Figure 3. At the very least, it is important for the authors to address or test the effects of this non-stationarity on inferred parameter values. How might strong non-stationarity affect the interpretability of parameter inferences? Perhaps more appropriately, the authors could consider related models that can accommodate seasonally dry soil moisture dynamics. In particular, Dralle et al. (2016, doi: 10.1002/2015wr017813) develop a seasonal stochastic soil moisture model and apply the model at Tonzi Ranch. The calibrated parameter values in that study are exactly comparable to inferred values in the present study. Similarly, Viola et al. (2008, doi: 10.1029/2007WR006371) present a transient formulation of the same stochastic soil moisture model.

We agree that seasonality at the Tonzi and Metolius sites affect our ability to inverse the soil water balance model with the selected data. A steady state period could have been better selected for each individual site. For simplicity/consistency we selected a single concurrent period for all sites and scales. Results therefore revealed which sites had poorer goodness of fit statistics and for which the steady-state solution for the analytical soil saturation pdf may not have been most appropriate. We will amend the results and discussion to explain this issue at the Tonzi and Metolius sites and propose practical considerations to address it in future studies.

We will amend the study goals, methods and discussion to include an additional sensitivity aimed at addressing the issues associated with the steady-state assumption. We will use data from the 2012 record periods and compare the goodness of fit of empirical and analytical pdfs using the full year of observations, and dry and wet periods selected specifically for records at each site/scale.

Specific Comments:

Page 1 Lines 8-9: What is a "hydrologically meaningful" scale?

The first sentence of the abstract will be changed to

Vegetation controls on soil moisture dynamics are generally not directly measured directly and not easy to translate into scale and site-specific ecohydrological parameters for simple soil water balance models.

Page 1 Lines 9-10: Passive voice makes the sentence a little confusing; try, "we hypothesize that pdfs of soil saturation encode sufficient information. . ."

The sentence will be changed to:

We hypothesize that empirical probability density functions (pdfs) of relative soil moisture or soil saturation encodes sufficient information to determine these ecohydrological parameters, and that these parameters can be estimated through inverse modeling of the commonly used stochastic soil water balance.

Page 1 Line 12: When the authors refer to soil "saturation", do they mean "water content", or "moisture"? I associate the word "saturation" with a water content equal to porosity.

We will specify : relative soil moisture or soil saturation

Page 1 Line 28: Check spelling of reference.

The spelling will be fixed

Page 1 Line 31: What are the "mean components of the soil water balance"?

Sentence will be changed to:

Given this ecohydrological framework, the probability density function (pdf) of soil moisture and the mean components of the soil water balance (rainfall, runoff, evapotranspiration, and leakage losses) are analytically derived

Page 2 Line 17: Issues with citations

The citation will be fixed

Page 3 Line 18: "interference"?

Interference will be changed to inference

Page 4 Lines 1-2: Usage, "confront pdfs. . .to a commonly used analytical model"?

We will think of a better word

Page 6 Lines 3-4: I do not believe the model specifies that ET occurs at a constant rate Emax.

The word constant will be removed and the sentence changed to:

The rate of evapotranspiration is assumed to occur at a maximum rate (E_{max}), which is independent of the saturation state.

Page 7 Line 12: Do Rawls (1982) list physical soil characteristics for these sites?

We will revise the sentence to:

Physical soil characteristics for soil textures associated with each site, s_h, K_s, and b were taken from Rawls et al. (1982) and are listed for each site in Table 1.

Page 8 Lines 9-10: It's not clear to me why values for Ew/Emax must be tested in a separate (not shown) calibration procedure. See General Comment (2).

Seer response to General Comment (2). Our results showed that *Ew/Emax needs to be smaller than 10% for equifinality to be reduced and that the convergence, goodness of fit and posteriori parameter distributions were not sensitive to values between 1 and 10%. So we picked 5%. We will amend the methods section to make this point more explicit. We are not including Supplementary material in with this manuscript. However the code associated with the analysis will be published.*

Page 12 Lines 6-7: My understanding of Emax is that it quantifies atmospheric moisture demand. Why should it scale with rooting depth? Typically, I've seen this value computed using Penman-Monteith e.g. Viola et al. (2008, doi: 10.1029/2007WR006371).

E_{max} is not exactly the atmospheric moisture demand, it is a fraction of the atmospheric moisture demand that can be withdrawn from the soil layer considered. E_{max} can be equal to the atmospheric moisture demand approximated by potential evapotranspiration (PET) if the full soil column or rooting depth is considered.

In this study we cannot assume that E_{max} = PET because only the surface soil moisture is sensed. In the revised manuscript we will only consider Z equal to the sensing depth and E_{max} is always expected to be lower than PET. We will clarify definitions in section 2.2.2 with the following sentences

The soil depth considered corresponded to the measurement sensing depths of 10, 20, and 5 cm for the point, footprint, and satellite scales, respectively. Because the soil depth Z is more shallow than the rooting depth, E_{max} is only a fraction of the atmospheric moisture demand (or potential evapotranspiration) contributed by that soil depth and therefore unknown.

Page 13 Line 1: I would suggest that model performance at Tonzi and Metolius suffers primarily due to the stationarity assumption, which is likely not valid at these Mediterranean sites.

We agree and will revise the discussion to reflect this comment.