

Interactive comment on “Using StorAge Selection functions to quantify ecohydrological controls on the time-variant age of evapotranspiration, soil water, and recharge” by Aaron A. Smith et al.

Anonymous Referee #1

Received and published: 31 March 2018

[12pt]article [margin=1in]geometry amsmath,amsthm,amssymb,amsfonts

C1

Review of "Using StorAge Selection functions to quantify ecohydrological controls on the time-variant age of evapotranspiration, soil water, and recharge"

Smith et al.

March 31, 2018

General Comments

The paper "Using StorAge Selection functions to quantify ecohydrological controls on the time-variant age of evapotranspiration, soil water, and recharge" introduces a novel approach to modeling vertical soil water movement accounting for root-water uptake, evapotranspiration, and transfer between a slow and fast moving soil domain. The novelty derives from the use of a "feed-forward" StorAge Selection (SAS) function model. Here, the soil column is conceptualized as a stack of control volumes with different SAS function and flux parameters. For each CV, the model simulates fluxes, water ages, and isotopic concentration (accounting for fractionation) in fluxes and storage. The model was calibrated at two field sites with data including observations of xylem water isotopic concentrations and depth profiles of soil moisture and isotopic concentration over a roughly one year period. The simulation results suggest that the water ages are consistent with previous studies, and that there is a generally

C2

strong preference for transport of younger water during wet conditions, which is also consistent with previous studies. Strengths of the paper include the clever and potentially broadly applicable new modeling approach and the informative and artful figures.

I would recommend, however, major revisions to improve confidence in the modeling results that form the basis of the paper's efforts to "investigate water flow in soils and identify soil evaporation and root-water uptake sources from depth" (P1, L12). The revisions should address two separate but related issues.

First, the evaluation of the model is not very compelling and needs to be improved to make the model results more credible. The model seems to have a large number of parameters calibrated to a relatively small number of isotopic measurements with high within-day scatter. The main indicator of model skill shown in the manuscript is an ability to roughly reproduce a seasonal signal in isotopic concentration that dampens with depth (Figure 3). The model also, presumably, simulates soil moisture, but this was not compared to data in the manuscript. The calibration keeps the 100 "best" parameter sets out of 50,000 random samples, which seems to be an arbitrary standard that does not consider the absolute quality of fit between observations and simulations. The final values of the model parameters are not reported, making model performance more difficult to interpret and potentially impossible to reproduce (given the stochastic nature of the calibration). The NSE_{adj} values for the isotopic concentrations are adequate (0.34-0.75), but this is not necessarily compelling given the high number of free parameters. No sensitivity analysis is done to show the importance of different model components in capturing the data. I was left wondering, for example, if a large change in one of the outflow SAS functions (say, in the CV at 10cm) would have an appreciable effect on model performance. If not, then the calibrated values might have a lot of uncertainty that is not presented, and the trends observed in the flux ages might not be

C3

significant. I was also left wondering, for example, if the difference between Site A and Site B SAS functions were significant, or within expected modeling uncertainty bounds.

Some potential ways to improve model evaluation are listed here. (1) The model parameters could be clearly listed with their calibrated values and ranges, to give the reader a sense of the uncertainties. (2) The manuscript could start with a much simpler model and build up to the complex model presented, showing at each step how additional model complexity is justified by the data. (3) The manuscript could report a sensitivity analysis to show how each aspect of the model structure is necessary to describe the data.

Second, the description of the model and underlying theory is at times confusing and seemingly imprecise. For example, the same variable is apparently used for age ranked storage in the slow and fast domain (see equations 1 and 2), some equations seem to be dimensional incorrect (see equation 3), and the CDF and PDF of the SAS function are seemingly confused (see equation 5).

If the authors can make substantial improvements to better describe and evaluate the model, then the results presented in the paper (e.g., the relative ages of different ecohydrological flows, the shape of the different SAS functions and their storage dependence at contrasting sites, the approach to simulating fractionation) could be significant contributions that merit publication.

Many of the the issues described above are listed in more detail with page references in the Specific Comments section.

Specific Comments

C4

P1, L16-17: Why do dominant young water fluxes lead to stable soil water ages? I would have thought that would make soil water sensitive to inputs.

P1, L20-21: "More variable" water ages? Meaning 50-65 is more variable than 56-79? The two ranges are not very different.

P2, L14-23: As pointed out, SAS functions have not been used to recover soil water ages at different depths. But there are other "physically-based" models that could be modified to do that (CATHY, ParFlow, etc). Why focus on SAS functions? A better justification would strengthen the manuscript.

P2, L28-34: It would be helpful to outline the structure of the paper to come: theoretical development followed by case study.

P3, L6-7: The phrase "since the time of rainfall" is a bit vague. Consider rephrasing definition of S_T .

P3, L11-13: The parenthetical phrases "exponential distribution", "random mixing", and "piston flow" are apples and oranges. One is a distribution and two are concepts. Consider clarifying.

P4, L14: The text refers to a "distribution of inflow ages (ω_j)...". But the notation ω is already being used for the pdf form of the SAS function (line 11), and this is a distribution of age-ranked storage, not age, with different units. This is either confusing notation or a conceptual mistake, and should be fixed.

C5

p4, L15: The ζ is described as being a relative age which presumably has units of time (in p5, L12) but is set equal to the PDF form of the rSAS function ω_Q in p4,L15, which has units of inverse storage (as shown, for example, in Harman 2015 equation 5). This should be clarified. In general, the proof would be easier to follow if the units (e.g, length, inverse time) were identified when parameters are introduced.

P4, L7-9: The age ranked storage can't be the "cumulative sum of the time", since it has units of storage. It is the volume of storage with age $\leq T$. Also, since this is in terms of "absolute age of water", should it be the time since it entered the vertical modeling domain, and not just the CV?

P4, L16-17 and Equation 1 and 2: The nature of the slow and fast domains was not immediately clear. A few more sentences of explanation would be helpful. Do they represent different conceptual storage volumes with different age ranked storages? Can they be illustrated in Figure 1? Are the left hand sides of equations (1) and (2) really identical? Assuming that they are, then we can set the right hand side of equations (1) and (2) equal, which simplifies to $2 * D * \Omega_D = Q * \Omega_Q + E * \Omega_E$. This suggests that during times when Q and E are zero, then D must be zero. Why so?

P5, Equation (3): It is confusing that S_z is described here as function of two variables (T,t), one variable (ζ), and three variables (T+ ζ , t, and z- Δz). More notation consistency is needed. In addition, the equation does not seem to be dimensionally correct: the LHS has units of length, and the RHS has units of 1/L times L times T times T, or T^2 . I think I understand what the authors' are trying to express, but it needs to be more precise.

C6

P5, Equation (5): Again, doesn't seem dimensionally correct. Seems like the CDF (not the PDF) of the rSAS function is needed on the RHS (Ω_z as defined in line 11)

P6, L18: What is h_s and in which equation is it used? It does not appear in equation 7.

P7, L10: The text states that "under free draining conditions θ_0 approaches zero." Does this mean that θ_0 is time-varying? If so, this should be explained more clearly, since the reader is likely to assume model "parameters" (as it was called in L9) are fixed.

P7, L16: how is V_F calculated? Also, it would be helpful to include an equation for the slow domain volume V_S . If there is a unique volume of storage associated with the slow and fast domain, then the volumes have separate age-ranked storages? Equations (1) and (2) suggest they are the same.

P8, L8: What is $SM(t)$?

Equation 15: The paper defines ω as the PDF SAS function in line 11, P3. Shouldn't the CDF form be used here?

P8, L23: The variable p is the normalized kernel density probability of what?

P9, L10: Please provide a reference for the kernel density estimation technique.

P9, L3-4: One additional sentence on how the "best" calibration was selected would be helpful, with the understanding the reader can refer to the citation for more details.

C7

P9, L7: The phrase "...by estimating xylem through root-uptake..." is confusing. What does it mean to estimate xylem?

P9, L1-4: The manuscript should describe why this calibration approach it thought to produce meaningful "confidence bounds" (as shown in Figure 3), and explain what the bounds mean. As I understand it, the range of the confidence bounds will approach zero as the number of Monte Carlo simulations goes to infinity (i.e. there will be >100 essentially identical "best" calibrations at a single optimal point in the parameter space), which makes the confidence bounds seem arbitrary and difficult to interpret. In other calibration techniques such as GLUE, the confidence bounds approach finite values as the number of MC simulations gets large, which makes the outcome more easy to interpret. I reviewed the Aho-Aho et al 2017 reference, and did not see this issue addressed.

Section 3.1 and Figure 3: The similarity in the isotopic concentrations observed at site A and B across time and depth (shown in Figure 3) is striking. The values and trends in isotopic concentration seem to be the same at both sites, even if individual values vary a bit. This similarity is unexpected given that Site B is described as more freely draining and has a different soil moisture profile (Figure 4, top panels). It seems important for the manuscript to comment on the similarity and whether there are any significant differences in the data collected from the two sites, since the models are calibrated to this data. Related to this, it also seems important to comment on why the difference in the drainability of the soil at site A and B does not seem to affect the measured isotopic concentrations.

Figure 2: What are the two dotted lines that split the Slow and Fast Domain in figure

C8

(a) and (c)? Also, consider adding political boundaries and labels to the map of the UK, to orientate the reader.

Figure 4: The upper soil moisture plots should be included in the figure description.

P14, L8-9: Here it states the "CVs for each site" are shown in Figure 4. This is the first reference I see to the actual number of CVs modeled. Is the bottom of each CV the number shown in the Y axes of Figure 4? This ambiguity speaks to a wider problem, which is that the number of final model parameters and the calibrated parameter values is not reported. Given the stochastic nature of the calibration, the calibrated parameter values used in the manuscript are needed to reproduce the results.

P17, L5-7: The sentence "The selection of deeper soil water at Site A relative to Site B resulted in slightly resulted in..." raises some of the same concerns described above. First, is the difference between the sites considered significant because the confidence intervals don't overlap? If the calibration used 10,000 monte carlo simulations instead of 50,000, would the confidence intervals be different in a way that could affect the significance test? Also, if the difference really is significant, is the effect of this difference apparent in the measured isotopic data? To convince the reader that these small differences in performance between Site A and B are greater than model uncertainty, it is helpful to show how they arise from the calibration data.

P18, L7-10: "It is notable though, that of the five xylem sample days, one (June 2016, Fig 3i, 3e) showed isotopic compositions different from either the simulated fast or slow domain isotopic concentrations." I was confused by this, because I do not see a uniquely bad fit between the simulation confidence interval and the observations on June 2016 in Figures 3i and 3e. I see that the simulation is not very good that day, but

C9

it is also not very good in 10/15 site B. Also, I don't see any differentiation in this plot between the fast and slow domain. Consider clarifying.

P18, L24-26: The meaning of $P_T(> 0.5)$ is not clear.

P18, L28: The phrase "one of the difficulties of identifying SAS functions at catchment scales is the shape of the SAS function..." seems to be circular logic. Also in general it was relatively difficult to follow the logic of this paragraph. Consider reviewing and clarifying.

P19, L32: Why is it that the median water ages were similar to previous estimates *despite* the similarities of the derived SAS function? I would have thought the median water ages would be similar *because of* similarities in the derived SAS function. A bit confusing.

P19, L2: The "Figs. 2b, 2c" do not show simulated isotopic enrichment. Is this the right figure reference? Also, in general, the sentence starting with "Notably," is confusing. Consider clarifying.

P19, L23: Not clear what is meant by the phrase "with the selection of young water" in the context of the sentence. Consider clarifying.

P20, L11-12: How was "a general reduction in the uncertainty of the SAS function" observed in Figures 4 and 3? As best I could tell, the certainty of the SAS function was not explicitly shown in the figures.

C10

P20, L30: "relatively simple framework"... relative to what? The approach seems fairly complex, even without accounting for lateral fluxes.

Technical corrections

P1, L29: "has infer a" is a possible typo.

P9, L22: Possible typo: "results".

P10, L5: Missing an open parenthesis.

P12, L5: possible type / extra word: "the"

P15, L8: possible typo: "through" should be "though"

P19, L21: possible typo

P19, L25: possible typo

Supplemental materials: There seems to be a typo or confusing phrase in the first sentence: "...using Eq. Soil fluxes...".