Anonymous Referee #3

Authors estimated field scale evapotranspiration (ET) by calibrating a 1D unsaturated zone model (HYDRUS-1D) using soil water content measurements, and compared simulated ET with observed ET from an eddy covariance tower. The HYDRUS-1D soil hydraulic parameters were calibrated using daily soil water content measurements from four theta monitoring probes at multiple depths and one cosmic ray neutron probe. While this is an interesting study, the novelty of the current study is not clear. Based on presented results, large differences exist between simulated ET and eddy covariance data and results of soil moisture simulations are not entirely satisfactory given the negative NSE during calibration and small coefficient of determination for soil moisture simulations at certain depths. In particular, authors have not discussed the implications of their results and what can be done to improve model estimation. While the focus of the inverse modelling was on soil hydraulic parameters estimation, the study can benefit from a detailed model sensitivity experiment to soil hydraulic and root growth function model parameters. I suggest authors to perform a detailed uncertainty estimation approach to identify the sources of errors (model input, parameters, or model structure) in ET and soil water content estimates. This can help to identify why the model did not perform well in some cases and how authors can improve their results.

Thank you for your comments. The central theme of the paper was to employ a standard publicity available model to test our hypothesis, not to devise new algorithms for inversion, and that is why we did not get into the inverse modeling details in great depth. As it was mentioned in the paper, more description about inverse modeling can be found in Mualem (1976), van Genuchten (1980), and Turkeltaub et al. (2015).

Moreover, Wang et al (2009) have done a detailed sensitivity analysis of groundwater recharge and evapotranspiration for soil hydraulic parameters in a single layer. We respect your concerns and have undertaken a sensitivity analysis of all 4 layers (24 parameters) extending the original work of Wang et al (2009). A full sensitivity analysis of the root model parameters is beyond the current scope of the paper and we refer the reviewer to Guswa (2012).

Guswa, A. J. 2012. Canopy vs. Roots: Production and Destruction of Variability in Soil Moisture and Hydrologic Fluxes. Vadose Zone Journal 11:3. doi:10.2136/vzj2011.0159.

Wang, T., V. A. Zlotnik, J. Simunek, and M. G. Schaap. 2009. Using pedotransfer functions in vadose zone models for estimating groundwater recharge in semiarid regions. Water Resources Research 45: 12. doi:10.1029/2008wr006903.

1. Introduction, the rational and implications of the current study are not entirely clear. I suggest authors outline the main objectives of their study and discuss how their results advance our understanding of ET estimation using unsaturated zone models. It is not clear whether authors try to develop a benchmark for soil moisture or ET estimation or how their soil hydraulic parameter estimation can help parametrize hyper-resolution land surface

models? These are the ideas that are discussed in the Introduction but their links with the current study are not clear.

Thank you for the comments. We will seek to improve rational of manuscript in the introduction.

2. Section 2.2.1. It seems authors have used a different growth root model compared to the HYDRUS-1D root growth model for annual vegetation. Have authors performed any experiments to assess how the results of the two root growth models compare?

   Since we had annual cultivation rotation between soybean and maize we had to introduce the root depth to the model and we could not use the default values inside the model. Likewise, as default values were constant and cannot be changed for different type of crops in different years during the simulation, we were not able to compare the models. This parameterization is not available in the standard HYDRUS package and a limitation of using it with crop rotations. We wanted to keep intact the cropping history to minimize impact on SWC between years. Clearly the topic of root water uptake deserves more investigation.

3. Section 2.2.1. It will be very useful if authors can report Kc parameters and root growth model parameters as they can impact the results of ET estimation.

   As it was mentioned in the manuscripts the suggested Kc values by Allen et al. (1998) for maize and soybean were used. For root growth model the maximum root depth assumed equal to 150 cm for maize and 120 cm for soybean. In addition, GDD was calculated by mentioned equation using Tmax, Tmin, and Tbase. We will try and clarify in revisions.

4. Section 2.2.2. Additional details regarding the inverse modelling algorithm and an objective function that is used for parameter estimation are required.

The maximum number of parameters that we can be optimized by the Hydrus-1D model is 15. We have followed the same procedure as Turkeltaub et al. (2015) and Wang et al. (2015, 2016). We used RMSE as our objective function and will clarify this more in the manuscript. Finally, a sensitivity analysis of all 24 parameters will be presented in the revisions.

5. Section 2.2.2. Line 206- Can authors provide further details about initial soil hydraulic parameters that they used in the modelling experiment? Did they use soil hydraulic parameters based on soil texture class information? Similarly, authors used the same parameter bounds for model calibration for all soil texture classes. It will be useful if authors can incorporate the soil texture information to define priors and initial parameter values.

The initial values were just the default values in the Hydrus-1D model which are based on the different soil types. Agreed, priors could be used with pedotransfer functions to improve results. Unfortunately, the connection between hydrologic fluxes and soil texture classes is unclear (Groenendyk et al. 2015). This work continues on that disconnection and will be highlighted more in the revisions.

Groenendyk, D. G., T. P. A. Ferre, K. R. Thorp, and A. K. Rice. 2015. Hydrologic-Process-Based Soil Texture Classifications for Improved Visualization of Landscape Function. PLoS One 10:6: 17. doi:10.1371/journal.pone.0131299.

6. Section 2.2.2. Why homogeneous soil type was used for simulating water content for the Cosmos-Ray neutron probe while for the Theta probes variability in vertical hydraulic conductivity is considered?

   As a first cut we used a single layer. Since the CRNP only sees the top 20 cm we wanted to see how well it could or not reproduce ETa values. Clearly more investigation is needed about the use of CRNP to estimate ETa.

7. Why the spin-up period is varied between the inverse modelling approach and the forward model? What criteria authors used to define model spin-up?

   We have followed the same procedure as Wang et al. (2015, 2016) for model spin up. We will clarify this in the text.

8. Table 2-Why negative NSE is obtained during calibration period particularly in deeper soil layers? Even R2 values are pretty small for a VZM model that is calibrated to observations. Can authors describe the reasons for this mismatch? Similarly results of soil moisture simulation are not satisfactory for the CRNP calibration based on Table 3.

   We will add clarification to revisions.

9. Authors indicate that inverse modelling based on CRNP data is most useful during the periods that soil evaporation is dominant. Can authors further explain why that is the case? One would expect that CRNP should provide better estimate of ET as its footprint is likely to overlap the EC tower footprint.

   Since the CRNP only sees the top 20 cm we wanted to see how well it could or not reproduce ETa values. We hypothesize that at roots development into deeper layers and Transpiration becomes more important in the latent energy term the information content in the CRNP would diminish. Clearly this topic requires more investigation. We will add clarification to revisions.

10. Section 3.2. Authors relate variability in performance of the model in ET simulation to variability in soil texture. However, one important information that is missing is vegetation type at the location of the probes and the EC tower footprint scale. Perhaps, authors should combine ET estimates from multiple probes to estimate ET at a field scale.

    Thank you for the suggestion. Based on reviewer 1, we will investigate if upscaling the AET by the SSURGO soil map will be useful. We also note that Professor Franz uses a range of hydrogeophysical mapping techniques (i.e. electromagnetic induction, cosmic-ray neutron rover) to understand soil patterns and properties. We have mapped this site several times and will consider adding some of the maps to this manuscript or a companion manuscript. We note the vegetation will be the same for all locations. Destructive vegetation sampling at each location is available from the site PIs to look at variability in the canopy.

11. It will be useful if authors can provide information about deep drainage from model simulations at multiple locations.

We will consider adding this to the manuscript.

Minor comments: Figures 1 and 2 can be combined in one Figure.

12. Line 166- Extinction

    Thank you, we will correct it.

13. Line 238-Please revise the Figure number to 7.

    Thank you, we will revise it.

14. Figure 5- Can authors describe the reason for large differences between the spatially averaged TP and CRNP by the end of year 2014?

We will consider this comment in the revised manuscript.