Text formats:

Reviewer – bold, non-italic

Answer of the author – non-bold, non-italic

Cited text – non-bold, italic

Anonymous Referee #3

Received and published: 27 September 2016

GENERAL COMMENTS

The presented manuscript applies time-series data from 9 cosmic-ray neutron stations to the land-surface model CLM in the Rur catchment. The authors assimilate the data using an ensemble Kalman filter technique to update states and parameters of their model. The added value of training data in years 2011 to 2012 is assessed (1) by testing the model performance in year 2013, (2) by testing the model adaption capabilities to an invalid soil map, and (3) by jackknifing single stations from the training period. The application of the cosmic-ray neutron method in large-scale models is one of the challenges in state-of-the-art hydrology and thus the present study is worth to be published in the scope of HESS after major revision.

Thank you for your positive evaluation. We sincerely appreciate your time reviewing the manuscript. In the forthcoming revision we will carefully address your suggestions and comments to further strengthen the manuscript.

1 Evaluating the overall quality
Large parts of the manuscript are written in the style of a protocol, by listing lots of other publications who used similar approaches, by mentioning tools that were used, and by reporting every step of the performed analysis. However, I believe that scientific articles should be entitled to challenge their own strategy by discussing alternative methods, by justifying their selection of tools and decisions, and by explaining the corresponding implications. I would thus recommend to rewrite and extend major parts of the introduction and method section. Therein, some literature reviews are unnecessary and probably unrelated to the study and can be omitted or need further explanation (see line-by-line comments). Here, I would suggest to follow the guideline that cited papers should be discussed, and not just mentioned. Other parts concerning the data integration and the SWC model need to be described in more detail. I would further recommend to reduce the detail of the results section, which is hard to follow without proper discussion, and thus to merge it with the discussion section. The results of the study were well structured and decribed, but are not entirely novel, and not sufficient to provide answers to all research questions raised by the authors. For example, to assess the value of a CRNS network of certain density to the performance of a land-surface model, various fractions of the 9 stations should be tested as requested by Referee #1.

We will shorten and sharpen the indicated parts of the manuscript to further increase its readability. Where appropriate, we will add additional discussion to the mentioned references. In the revised version we will state more clearly, how the research questions were answered by this study. As pointed out in the answer to Referee #1, new simulations will be run with a more limited number of cosmic-ray neutron sensors (4) to emphasize the additional value of a network of CRNS using five CRNS for evaluation.

Furthermore, the study uses some questionable assumptions, like a constant error of soil moisture data (although neutron measurement uncertainty highly varies with wetness condition), or the assimilation of SWC data assuming homogeneous vertical profiles and no changes of seasonal biomass (see line-by-line comments for details). Another questionable approach is to allow static site-specific parameters to be variable in time, e.g., hydraulic conductivity or soil porosity. This is highly counter-intuitive and should be discussed with respect to uncertain data and/or model conceptualization. I agree to most of the general comments made by Referee #1 and will thus compensate the previous reviews by detailed line-by-line comments below.

In the revision we will address the mentioned assumptions either when stated in the experiment set up or in the discussion section, if this has not been done already.

2 Line-by-line comments, scientific questions/issues, and technical corrections
Thank you for the detailed review of the manuscript and the line-by-line suggestions. We will address and change the line-by-line comments accordingly in the revised version of the manuscript.

• Title ("Evaluating the value of a network of cosmic-ray probes for improving land surface modelling"): I'd suggest to remove "the value of" to simplify the title. Furthermore, state observations usually do not improve a model, they rather improve model results, e.g. predictions.
Thank you for this suggestion. We agree and suggest to change the title towards:
"Evaluation of a cosmic-ray neutron sensor network for improved land surface model prediction"

Page 2
• L1: "Land surface models can model": bad phrasing, replace "can model" e.g. by "describe".
This will be changed.

• L3: "CRP", please use the newly accepted abbreviation CRNS (cosmic-ray neutron sensing/sensor) with regards to the recent 5th COSMOS workshop.
We agree and will change this.

• L14: improve readability, split in two sentences.
This will be changed.

• L18: please add a statement about the impact of your findings for the scientific community.
This will be added.

• L21-22: this sentence needs a reference.
This will be added.

• L24: " and is "→ "while it is"
This will be considered.

• L27-28: the given number of references here appears to overwhelm the statement and its low relevance to your paper. Please use only the 1 or 2 most important citations.
This will be addressed.

• L30-31: Please discuss the alternatives in more detail to strengthen your decision to use CRNS technology. Were space-borne remote-sensing products assimilated to LSMs before? Why wasn't it successful? What about the use of airborne products with higher resolution and depth? You could also mention point-scale or large-scale soil moisture monitoring networks which have been used for evaluation of land surface models.
This will be addressed and we will discuss in more detail the limitation of assimilating remotely sensed soil moisture products.

• L31: "not reliable for areas with dense vegetation": a paper by the same first author recently found that CRNS is also influenced by dense vegetation. Is it more reliable?
Yes. CRNS is influenced by vegetation and more reliable. This will be addressed.

• L33: the selection of citations for this statement appears to be random/unrelated. If you want to provide references for the "intermediate scale", Zreda 2008 and Köhli 2015 might be appropriate.
This will be addressed.

• L34: "desired application scale of land surface models": please make the reader happy by finally providing concrete information. What is the scale? Are you talking about centimeters or lightyears? Please do not use citations inflationary and do not keep them untouched. How do the three citations help you to support your argumentation?
We will provide the concrete information earlier in the revision.

Page 3
• L1: omit "fast" as it repeats with the next sentence.
Okay.

• L3: add "fast" to make clear that the sensor measures the non-moderated neutrons.
We will change this-

• L4: "15 ha", your SWC range seems to be 10 to 40%, which leads to an approximate CRNS footprint of 7 to 14 ha following Köhli 2015, excluding vegetation and altitude influence. You could write "maximum area of 15 ha" to circumvent mentioning this variability.
Thank you for the suggestion. We will change this.

• L6: omit "Desilets and Zreda, 2013" as it does only marginally address heterogeneous averaging. Franz 2013a is already a great reference to this topic, Köhli 2015 also touched this.

We will address this comment.

• L8: Bogena et al. 2013 did not perform simulations to the penetration depth. Instead, Franz et al. 2012 (doi:10.1029/2012WR01187) and Köhli et al. 2015 provided simulations that both support these values.

We will address this comment.

• L11: add a reference for COSMOS-UK, Evans et al. 2016, 10.1002/hyp.10929

Thank you. We will add this new reference.

• L13-15: please rephrase to make clear what data assimilation is and is not.

We will change this.

• L15: It is not clear why you choose EnKF. Please at least mention other techniques and provide reasons for your choice. The sentence further should be moved to the end of the paragraph after you have introduced the history of DA.

Thank you, we will add this.

• L16-34: This historical overview appears to be unnecessary in the context of your study. Neither do you explain what things like " four-dimensional variational DA" are, nor is the relation to your work described. Furthermore, citations are used inflationary again. Please reduce this paragraph to the key publications which support your study. Also think about moving certain studies about ensemble size, multiple time steps, and other filtering approaches to the methodology section, where you need justification for your approach.

Thank you. We will add motivation and consider moving the overview to the methodology section. This is not clear yet.

• L30-32: Just to emphasize the previous comment, these lines particularly carry no information for non-experts due to the lack of explanation.

Thank you. Same as above. We will make this more clear in the resubmission.

Page 4

• L1-19: As stated before, the whole literature review appears to be random and irrelevant to your work. Or at least the relations are not explained. For example, work from Montzka 2011;2013 and Han 2014b appear to be of some relevance for you, prior to others.

We will restructure the literature review and explain this better. Nevertheless, we disagree that the literature review is random or irrelevant and think that the large majority of the cited papers is relevant. This will be motivated better in the revision.

• **L23-24: "Its capability to propagate surface soil moisture information into the deeper soil column was analyzed by Rosolem et al. (2014)", what does this sentence mean?**
We will clarify this in the resubmission.

• **L26: "The COSMIC operator", third repetition as a sentence starter.**
We will reformulate sentences here.

• **L27-29: combine those sentences: "neutron observations have been used to update states (. . .) and hydraulic parameters (. . .)"**
The sentences will be combined.

• **L29: "showed" → "demonstrated"**
We will revise the use.

• **L29-30: be more correct in phrasing. Rephrase that Villarreyes 2014 used a different model, but also estimated hydr. parameters by inversion. Han 2016 did so too, using support from neutron data, but neutron assimilation alone does not "update" a hydraulic parameter.**
Thank you. We will rephrase these sentences to be more precise.

• **L31: "This work further explores", omit "further". Until now it is not clear what this work does, you only told stories about work of others. Please summarize which of the presented approaches you are picking up and what scientific novelty you add.**
We will clarify this in the revision.

Page 5
• **L3-4: "the soil moisture characterization at the larger catchment scale", what exactly is meant by these terms, and how do you measure improvement?**
We will specify this in the revision.

• **L4: "how dense the CRP network should be", do you answer this question?**
We will specify to which extent this question is answered and what are the limitations to further answer this question.

• **L7-8: "soil maps and atmospheric forcings show spatial correlations over larger distances", this is an interesting point, please provide reference. Isn't the largescale heterogeneity of soil maps only an artefact of soil data scarcity?**
We will provide a reference and think the large scale correlation is not an artefact.

• **L9: "10 stations", do you assimilate all 10, or just 9?**
In fact we assimilate 9 stations. This will be clarified/corrected.

• **L15: "feasibility of the updated large scale soil hydraulic parameters", how can a parameter be feasible? Please clarify your novel research question.**

We will rephrase the sentence to become clearer.

• L18-19: The sentences can be omitted as being obvious.
The sentence introduces the topics of the following paragraphs. As such, the sentence is not redundant, but may seem obvious for some readers.

Page 6
• L6: correct wording, a "process" can not be "solved"
This will be corrected.

• L10: "Oleson et al. (2013) provide further details on CLM4.5", redundant information with regard to L5-6.
We will remove this sentence.

• L10-12: provide reasons why you artificially limit the scope and complexity of your study. What process would a "biogeochemical module" have added and why are they not important here compared to a prescribed LAI?
This would be beyond the scope of this module as invoking the biogeochemical module would require a mode spin-up of 1000 years for the catchment.

• L14: please finally (after lots of references in the introduction) provide concrete information about the grid size in your study (the reader is still lost between centimeters and lightyears)
We will explicitly state the grid cell size. It is however common practice to introduce the grid cell size after the introduction and after the methodology.

• L23: use standard format for functions, $k[z] \rightarrow k(z)$
We will change this.

• L24: format $z \rightarrow z$
This will be modified.

• L24: what is the difference between "soil moisture" and SWC? Why are you using the expression θ here, while SWC is used elsewhere (e.g., eqs. 24 and 25)?
Thank you. Agreed. We will change the term SWC in Eqs. 24 and 25.

• L25: use the more convenient expression $k_{sat}(z)$,
We will do this.

• L26 (eq. 1):
– format $k[z] \rightarrow k(z)$,
– rewrite $k_{sat,z} \rightarrow k_{sat}(z)$ as this is a functional relationship. In contrast, indexing a state variable θ_i is ok.
– omit occurrences of 0.5 since $\frac{0.5}{0.5} \approx 1$,

– case conditions (e.g., *1 < i < N . . .*) are usually preceded by a comma ineach line

– the curly bracket on the right is not common in multi-case equations.

Thank you. We will recheck the formulations.

Page 7

• eq. 3 and 5: reformat *sand* → sand**, same for clay.**

We will format the terms as suggested.

• L6: "whereas", split sentence here.

Thank you. We will split the sentence.

• eqs. 9 and 10: this is a single equation, requiring only a single equation number, and a multi-case alignment using a curly bracket

We will modify this as suggested by the reviewer.

• L12: reformat *mm* → mm**,**

This will be corrected

• whole page: please motivate the reader why these details are important for your research question. Also provide information where all these empirical (fixed) parameters (or regression coefficients) are coming from. Is the underlying theory so well understood that no uncertainties or further dependencies are required?

This is important and we will formulate the motivation more clearly to inform the reader on the use of these pedotransfer functions. Here we will point out uncertainties and further readings which are related to the uncertainty of these particular pedotransfer functions.

Page 8

• L4: "COSMIC parameterizes interactions". The interactions are parameterized by the underlying physical cross-section data. COSMIC rather parameterizes the neutron transport.

Thank you. We will be more precise in the description of COSMIC.

• L7-8: Repetition from the introduction.

We will remove the sentence.

• L10: "high energy neutrons are reduced" → "the number of high energy neutrons is reduced"

We will modify this as suggested.

• L11: "with less energy in each soil layer", misleading/unphysical. Fast neutrons typically evaporate with constant energy.

We will modify this.

• L12: rewrite "soil interaction", as fast neutrons predominantly interact with the water.

Agreed. It will be changed to something like "soil moisture interaction".

• **L16-22 and eq. 14: this part can be omitted, since it is already well described in papers from Shuttleworth and Baatz, and does not add to the message of this paper. If you decide not to omit it, replace θ in eq. 14 to avoid confusion with soil moisture.**
Thank you. We will replace θ accordingly.

• **L22: explain to the reader how the 300 soil layers in COSMIC communicate with the 10 soil layers from CLM.**
We will explain this in the revised version.

• **L26: what is a "COSMIC soil surface"?**
We will rephrase or explain this in the revised version.

• **L25ff: it looks like you are not assimilating neutrons, but reiterating SWC from neutron data. The whole paragraph creates a great confusion about what the difference is between SWC, CLM SWC, weighted CLM SWC, and CRP SWC. In contrast to other less relevant paragraphs in this section, this part is highly unclear and simultaneously highly important to understand the most important part of your model. Please rephrase the whole paragraph and clarify to the reader what exactly you do, and why (i.e., why not assimilating N directly?)**
Thank you. Obviously this paragraph needs to be revised. We will rephrase the paragraph to make clearer what is assimilated and how COSMIC and CLM communicate.

Page 9
• **L5-10: Your paper is not a protocol. Again, it is described what you are using and who else used it, but the reader is left with the question why you (and others) made this decision. Shortly explain advantages of your strategy and why it serves your research question better than others.**
Thank you. We will put the motivation and explanation forward.

• **L11: what is f in $\vec{x}^f$?**
We will add an explanation. f marks the model state at which data is assimilated whereas a marks the new model state after assimilation.

• **L19: confusing typesetting. Is it $\vec{H}$ as a function of the COSMIC model, or is $\vec{H}$ identical with the COSMIC operator?**

The sentence will be reformulated to clarify this.

• **eq. 19: do not use T as a symbol for transposition, there is a reserved symbol for this: $\vec{Y}^\top$.**
Thank you. We change T to the reserved symbol.

Page 10

• L18: redundant sentence.
The sentence will be deleted.

• L20: why these values? is it comparable with the catchment-mean texture? If your question is, what impact a rough and uncertain soil map in data scarce region would have, wouldn't it be more reasonable to smooth out the existing soil map to a very rough degree, rather than using a completely arbitrary soil map?
Thank you for your questions. We will do calculations with a third soil map for the revised version of the manuscript (see responses to reviewer #1).

We will also add information on catchment wide average sand and clay content, and what was the motivation to select the biased soil map as initial soil map in part of the simulation experiments:

"*The BK50 soil map provides the initial high resolution soil texture for the catchment and is the most detailed soil map available for the defined region. Average sand and clay content of the catchment are 22.5% and 21.4%, respectively. As an alternative, simulations were also performed for a biased soil texture distribution with a fixed sand content of 80 % and clay content of 10 % (S80 soil map). This represents a large error with respect to the expected soil properties. However, perturbations of 10 and 30 % guarantee some variability in the initial soil properties. The S80 soil map simulations allow evaluating the joint state-parameter estimation approach because given the expected bias, we can evaluate whether and to what extend the soil properties are modified by the data assimilation to be closer to the available high resolution soil map.*"

Page 11
• L3-5: omit physical units (they are irrelevant in this context).
We will omit physical units here.

• L8-11: How do you justify the perturbation of physical soil parameters like porosity and texture? Does the uncertainty of the soil map justify the huge variation ranges applied in this work? Are models allowed to adapt their physical basement to hydrological data (which also show uncertainty)?
Concerning your first question: little information is available, but tests were made with different degrees of uncertainty and the sensitivity of the model outcomes to the imposed perturbation was limited. Concerning your second question: this is a standard procedure in data assimilation and we can refer to numerous papers where this strategy was followed. We will provide additional motivation in the revised version of the manuscript.

• L15: omit "="
We agree.

• How was the CRP SWC uncertainty determined? Assuming a constant CRNS error is not physical and might have substantial influence on the results (to be tested). For example, the error of neutron observations N is pN, while N can almost double from very wet to very dry conditions, which leads to a

variation of the neutron uncertainty by 30%. This can propagate through the non-linear relation to soil moisture in such a way that your observed SWC is significantly more uncertain in wet periods compared to dry periods. Consequently, the DA approach should give more weight to dry periods during assimilation.

This is an interesting point. We will add explanation in the manuscript and acknowledge the approximation that was made here. The error you are referring to is the Poisson noise in the measurement. Averaging measurements over a sufficiently long time windows (24 hours in our case) reduces the error in the measurement substantially. Naturally, not all sources of uncertainty can be identified and ruled out. Previous work has shown that there is still an error in the SWC estimation by CRNS (Baatz et al. 2014). Based on this work we assume a fixed error of 0.03 cm3/cm3.

Page 12

• L16: Why do you use RMSE, although many alternative measures are accepted as state-of-the-art measures for time series evaluation, e.g., KGE or NSE, in order to assess bias, deviation, and correlation simultaneously?

RMSE is a standard measure used in the data assimilation community, whereas NSE is not standard.

• eqs. 23 and 24: reformat *SWC* → SWC**, same with** *RMSE* **and** *bias***, as those are single multi-letter variables, not products of multiple single-letter variables. Following this style guide, rewrite** E_{RMS} → E_{RMS}**. You can even omit "**RMS**" since** E **is the only error used in this work. This would improve readability of the results section.**

Thank you for the suggestion. We will reformat the equation and adapt the equation to the style guide.

Page 16

• L17-26: It is argued that changes in SWC states have impact to simulated ET flux. However, only for state-parameter updates (L19). Why is ET not affected by (SWC) state updates only?

This .

• L32: "to E_{RMS}**-values", omit "-"**

Thank you, agreed.

Page 17

• L9: if precipitation data from COSMO_DE was used, why was this information omitted in the method section (only mentioning DWD)?

This information will be added in the revised version of the manuscript.

• L26: replace "fast" with "quickly".

Thank you.

Page 18

• L10-23: This question already needs an answer in the method section, I'd suggest to move the whole paragraph.

We agree and we will integrate this paragraph in the method section.

• L10-23: I cannot follow the argumentation. Baatz et al. 2014 suggested a correction function for neutron counts based on vegetation estimates. In your model, you already have LAI data every month, implementation of the correction functions in the model would probably be straight forward. Furthermore, to convert neutron data to SWC, some vegetation correction would be necessary, too. Third, assimilating CRP SWC assumes homogeneous vertical SWC profiles (before iteration), this assumption would be unnecessary if neutrons would be assimilated directly. I am afraid that this topic is more complex and needs further discussions and tests. It would be most convincing if you could show that neutron assimilation indeed gives different results than CRP SWC assimilation.

We will clarify why leaf area index data is not sufficient to estimate biomass in a catchment. There is no vegetation correction necessary to convert neutron data to SWC. We did on-site calibration which circumvents the need for a vegetation correction if e.g. a rover is applied in a region with high biomass differences (e.g. forest and crops).

• L27: "neutron flux intensity", do you mean flux or intensity or both?
Thank you. We will remove "flux".

• L27: "Although . . . only available at few locations", write more positively. Neutron data was available at up to 9 locations, which was intended to be the amazing novelty compared to other catchments!
Thank you. We will rephrase the sentence.

Figures
1. South → south, same with North.
This will be changed.

2. Please add grid lines
We will add horizontal grid lines.

3. Please add grid lines
We will add horizontal grid lines.

4. Please add grid lines
We will add horizontal grid lines.

5. It is hard to distinguish two black lines with different meaning. Further indication of the expected "true" sand content (given by the soil map or soil samples) would be helpful to evaluate these plots.
We will add the sand content of the BK50 soil map.

6. This figure is not understandable without the text. Please shortly provide information about the B parameter in the caption to understand the message of this figure.
We will add information to the caption.

7. replace k(sat) → k_{sat}**. It would be interesting to also show the evolution of the soil porosity parameter together with an indication of its measured value. Why does hydraulic conductivity (and probably also porosity) vary over time at individual sites? Those are expected to be constant physical parameters of the sites. In my opinion this is a serious flaw of the DA approach used here.**

Thank you. This issue is well known. This is a standard procedure in data assimilation and we can refer to numerous papers where this strategy was followed. We will provide additional motivation in the revised version of the manuscript. Notice also that simulations were made for a verification period with constant parameters (estimated in the assimilation period).

8. The purpose of this figure is not clear, as no observation data is provided to evaluate the model performance with respect to simulated latent heat.

Thank you. The issue you point out is in agreement with the desire of referee #1 for additional latent heat flux observations. We will add or modify Figure 8 in the revision with some LE observations in comparison to modeled ET.

Tables

1. what is C3? Replace "non arctic" with "non-arctic", probably add a citation to the caption for plant functional types.

Thank you. We will add information, a reference for plant functional types and replace "non arctic".

3. improve readability by increasing font weight (boldness) for particularly good cases below an RMSE threshold, which is a common strategy in many journals.

Thank you for the suggestion. We will mark particularly good cases with bold fonts.

4. same as 3. Rephrase the last sentence.

Thank you. We will rephrase the sentence.

5. same as 4.

Thank you. We will rephrase the sentence.