Author Response to Reviewer #2 from

Improving Calibration and Validation of Cosmic-Ray Neutron Sensors in the Light of Spatial Sensitivity – Theory and Evidence

M. Schrön, M. Köhli, L. Scheiffele, J. Iwema, H. R. Bogena, L. Lv, E. Martini, G. Baroni, R. Rosolem, J. Weimar, J. Mai, M. Cuntz, C. Rebmann, S. E. Oswald, P. Dietrich, U. Schmidt, S. Zacharias
Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-148

RC: ***Reviewer Comment***, AR: *Author Response*, ☐ Manuscript text

Dear Anonymuous Reviewer #2,

we are greatful for the positive review and the helpful comments and suggestions. Please find below our response to the main points. Every single specific comment will be addressed in the revision.

1. Reviewer #2

1.1. Main concern #1

RC: ***The recommended calibration scheme calls for about half of the calibration samples to be taken inside the 20 m radius. This will over-characterize the small area (<20 m) in comparison to under-characterized wider field (>20 m). I would like to see some discussion of how these over- and under-characterization will affect the calibration.***

AR: *Thank you for pointing this out, which is indeed a big challenge related to sampling strategies and interpolation in the CRNS footprint. The recommendation that we aimed to give is that samples should be taken according to the local conditions and correclation lengths. Thus, a large amount of samples below 20 m only makes sense if the correlation length of soil moisture is much smaller than 20m. Our argumentation is, that point measurements should have highest accuracy in regions where the CRNS is most sensitive. Therefore, taking less samples far afield would be justified by the very low sensitivity to those regions. Measurement uncertainties can be higher at large distances, as the influence on the weighted average would be marginal.*

A universal quantification of the effect on the calibration cannot be provided, because it highly depends on local structures and conditions, as has been argued in the text. Instead, a dedicated study would be necessary to quantify the effect of the number of sampling points in different distances from the probe. We will further soften and clarify the statements in the revision.

2. Main concern #2

RC: ***The concern about time series stems from the fact that distributed sensor data typically do not include the top 10 cm of soil. And because the top 10 cm is critical to cosmic-ray signal, particularly after a precipitation event, these data are not comparable to cosmic-ray derived soil moisture. Placing point sensors close to the surface (0-2 cm) will change your comparison between point sensors and cosmic-ray data. If these shallow depth data are not available, your conclusion regarding the comparison, and by extension regarding the calibration, suffer diminished credibility. I would like to see an assessment of the effect of the lack of near-surface point data on the comparison. Perhaps you can develop an assessment***

that could add to the uncertainty of the point data, and this will affect the comparison with cosmic-ray data.

AR: *This is a good point that indeed influences the calibration/validation performance with CRNS. One of our coauthors, Ling Lv, investigated exactly this issue at their site in the T.W.E. Daniel Forest. They burried the shallowest sensors at 10 cm depth and extrapolated data towards 0–9cm using hydrophysical simulations with the soil hydrologic model `HYDRUS`. This approach significantly improved the comparison of the time series to CRNS. In our study, we made use of these interpolated and extrapolated data. We will add this detail to the site description.*

At the other sites with time series data, the shallowest soil moisture sensors were burried at 5 *cm, while their electromagnetic sensitivity covers distances of up to* 5 *cm. We are thus assuming that the measurement is representative for the first layer of the soil. However, we fully agree that surface water at 0–*1 *cm is probably not tracked by conventional sensors. This is why we argue that deviations of CRNS and soil moisture sensors are often related to intercepted or ponded surface water, especially during and shortly after rain events.*

In fact, this issue and the detailed assessment suggested by the reviewer is currently part of a separate research project (in which some of the authors are involved), which aims to quantify the uncertainties related to such lacks of information in soil profiles.

2.1. P1 L8

RC: ***Clarify the meaning. Is it two calibrations per time series (or per location)?***

AR: *Thank you. The approach is extensively tested at six distinct monitoring sites: two sites with multiple calibration datasets and four sites with continuous time series datasets. We will change it as suggested.*

2.2. P1 L9

RC: ***clarify the meaning. or drop "robustly"***

AR: *We mean that the improvement is robust across different sites, conditions, and performance measures. However, we agree to drop the word as it would require further explanations which is too detailed for an abstract.*

2.3. P1 L10

RC: ***what is CRNS product? define it***

AR: *We mentioned in line 2: "The sensor provides continuous soil moisture data", we thus feel that the term "CRNS product" is adequately exlained and would thus suggest not to change the wording.*

2.4. P1 L10

RC: ***unclear from the previous sentences what the "presented approach" is; state clearly what the approach is and then expand on it by providing some insights.***

AR: *Thank you, we will clarify that we mean the weighting approach and how it differs from the conventional function.*

2.5. P2 L2: 10 to 10^4 m

RC: ***why this range? it doesn't follow from previous text; explain, for example by saying that this is the scale at which many relevant hydrological processes operate at the land surface (and give some supporting references)***

AR: *Thank you, we will change the wording to make clear that we refer to the models cited in the preceding sentence, which operate at that scale.*

2.6. P4 L1

RC: ***I wonder if adding another term to represent "other" hydrogen would be good here; "other" could be defined loosely as any known and unknown hydrogen not represented by the previous terms***

AR: *Thank you for the suggestion, we will introduce θ_{other} as suggested.*

2.7. P4 L19

RC: ***this sentence is self-contradictory; if they are minor, why recommend future investigation? please, rewrite to make a clear point.***

AR: *Thank you, we will clarify that those factors are irrelevant for the investigated sites and are thus of minor importance for the conclusions in this work. We will omit to suggest "further studies" here despite the fact that we still think those studies are important in cases where non-conventional sensor locations are used or sensors placed in extraordinarily complex terrain.*

2.8. P4 L21-23

RC: ***in homogeneous sites all approaches should give the same result, shouldn't they? please rewrite to make this clear.***

AR: *You are right, this can be considered as a basic principle of weighted averaging. We tried to emphasize that some previous studies applied the equal average only at homogenous sites, such that potential flaws of that approach did not became evident. However, we will skip this part of the sentence as it created confusion and does not add significant information.*

2.9. P6 L16

RC: ***is this a proportionality sign? if so, put everything to the right of it in brackets to make it clear***

AR: *Yes. We do not see the need for brackets enclosing a full side of an (proportionality) equation.*

$$\boxed{e^{-r/127} \approx W^{\text{conv}}_{r\leq 300} = \partial_r \text{CFoC}(r) \propto 1 - a_1\, r + a_2\, r^2 - a_3\, r^3 + a_4\, r^4}$$

2.10. P7 Figure 2

RC: ***I wonder if the two graphs combined would show the weights better.***

AR: *We decided to split the graph in order to resolve the short-range and the long-range features of this highly non-linear function. If combined into a single graph, the peak below 5 m would become unresolvable for the* *human eye.*

2.11. P8 L7-8

RC: ***it shouldn't matter whether water contents are gravimetric or volumetric; but male it clear which on you used, and why you would convert to gravimetric, for example.***

AR: *Thank you. We used the volumetric approach as described in the text and will rephrase the text to clarify. We agree that the difference may not be significant, however, the gravimetric approach would contradict with the foundation of the weighting functions.*

2.12. P9 L1-3

RC: ***does this mean that the fancy weighting may not be necessary? or am I misunderstanding this statement?***

AR: *It is not necessary to integrate over volumes around the point samples as long as they are distributed reasonably across the relevant area or soil profile. The corresponding errors have been observed to be far below significance.*

2.13. P7 L7

RC: ***up to this equation the equations were numbered; can you continue numbering them?***

AR: *We think that equation numbers should not be used inflationary. Only equations that are actually referred to in the text exhibit numbers. If it is a convention of HESS to number every equation, we are open to add these numbers.*

2.14. P13 L14

RC: ***clarify and expand this***

AR: *Thank you, we will add explanations. If a sensor is dedicated to measure soil moisture in a certain field, it should be ideally placed in that field. Also CRNS stations at the field border can be biased by different local characteristics, such as land use or soil properties.*

2.15. P14 L1-4

RC: ***This is unclear (and contradictory); how do the orange plots show reduced hydrogen? by associating the measurement with a particular line? even so, the difference is small, so perhaps you could discuss this in some more detail.***

AR: *When the same number of neutrons corresponds to different field soil moisture (i.e., different values on the line* $\theta(N)$*), then other hydrogen sources must have changed that were not statically included in* θ*. Considering results from the conventional approach at* $N \approx 1150$ *cph, the calibration line from May indicates 14 % soil moisture, while the line from July indicates 17 %. Therefore, other hydrogen pools must have decreased to achieve the same neutron count for increasing soil moisture. This can lead to the interpretation of reduced biomass from May to July, which contradicts experience that growing maize increases its biomass. We will rephrase the text to make this more clear.*

2.16. Fig 6 N_0 values

RC: ***the three N0 values seem close to one another; can you include standard deviations? I bet they will make all three N0 the same, statistically.***

AR: *Although the range of N_0 from 1139 cph to 1172 cph gives the impression of insignificant deviation, this parameter is highly sensitive to the performance of the CRNS soil moisture product. As N_0 is determined from N and θ by calibration (and not by observation or averaging), we cannot provide standard deviations. However, we agree that the Gaussian propagation of uncertainties for N and θ leads to uncertainties of N_0 that are comparable with the mentioned range. The involved uncertainties were addressed implicitely in the text:*

> Therefore, calibration results that resulted in an RMSE of $\approx 4\,\%_v$ (Fig. 6) are not meaningful under wet conditions (where $\Delta\theta(N) \geq 4\,\%_v$), and are still uncertain under dry conditions (where $\Delta\theta(N) \leq 4\,\%_v$). Consequently, the partial coverage of the CRNS footprint by the irregularly distributed SoilNet hampers the proper evaluation of the CRNS data, and especially of the weighting strategies.

2.17. Fig 6

RC: ***this right panel is misleading: of course the CRNS with revised calibration will show better agreement with revised data; can you make it clear and include a statement or two to discuss this artefact?***

AR: *Thank you, we will clarify this in the caption. The average soil moisture using the conventional weighting approach (orange) exhibits poor performance against the CRNS signal (not shown). The revised approach improves four performance measures of the averaged soil moisture (blue) and the CRNS signal (light blue), although the SoilNet probes are unevenly distributed in the CRNS footprint.*

2.18. P15 L5-6

RC: ***this would be true if the remaining (not covered) area has soil moisture that does not show any correlation with soil moisture in the covered area; as this cannot possibly be true, this statement is also untrue.***

AR: *Thank you, we will clarify this statement, saying that 36–51 % of the neutron variability does not directly respond to the wetness conditions monitored by the irregularly distributed network, and in most cases the soil moisture of the outer area can be assumed to correlate with the inner area.*

3. P16 L12-13

RC: ***I think this statement must be justified somehow; how about producing "cow water equivalent" (CWE) and using it to correct the data?***

AR: *This procedure would be ideal to investigate the influence of cows. However, no data is available on which days the cows were present at the field and near the CRNS probe. We will soften the statement slightly. In order to justify the potential effect of cow equivalent water, we can at least argue with preliminary results from recent experiments using "Scientist water equivalent". The following figure demonstrates that (1) the influence on the neutron signal depends on the distance of the scientists (or cows), which would require advanced cow-tracking technology at our site, and (2) the sensitivity to scientists at distance r confirms the shape of the revised sensitivity function W_r used in this manuscript.*

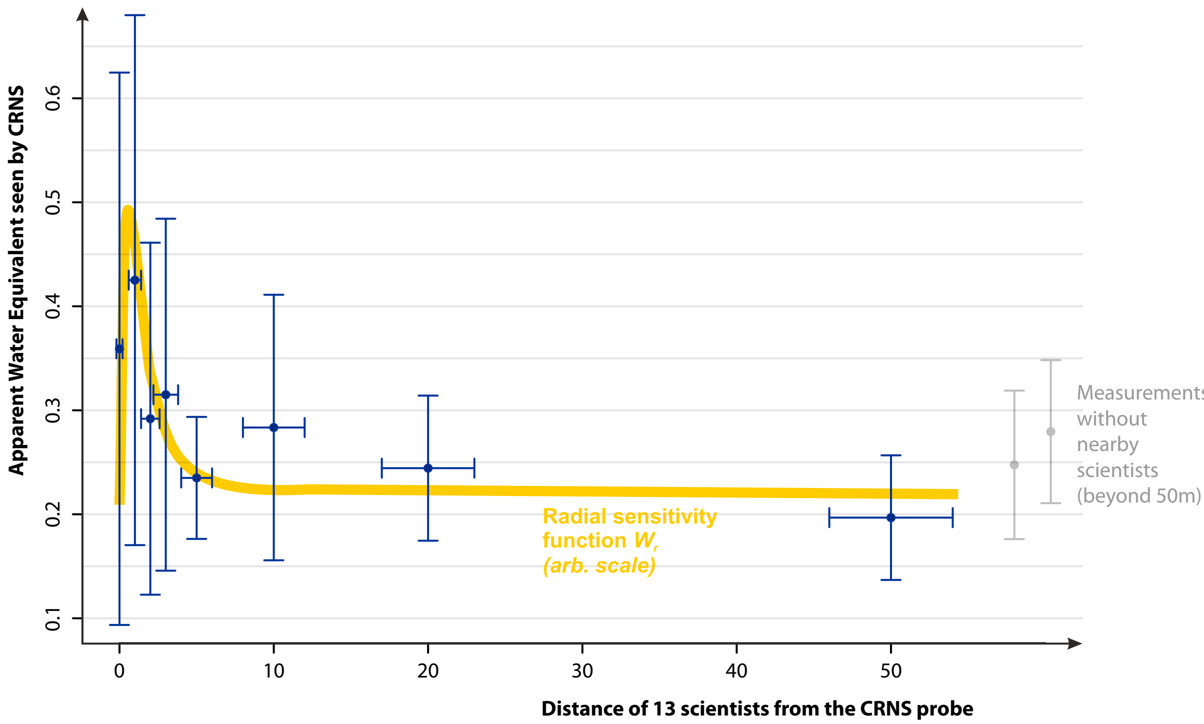

Figure 1: (not in the paper) An experiment to investigate the sensitivity of neutrons to scientists at various distances from the CRNS probe (preliminary). The experiment indicates that also cows could have substantial influence on the neutron count rate (and thus performance) in *Großes Bruch*, where cows are able to approach the CRNS probe as close as 20 centimeters.

3.1. P15 L20-22

RC: ***so what's the point of this exercise?***

AR: *Since the CRNS signal is influenced by intercepted and ponded water, calibration against soil moisture is an uncertain venture. Thus, we repeated the calibration only in periods were soil moisture can be assumed to have the main influence on CRNS. This experiment helps to find out whether the performance gain was just an artefact of the dynamics unrelated to soil moisture. We will clarify this in the text.*

3.2. P15 L24-25

RC: ***or moving infiltration front, as you mentioned before; or is this taken care of by the proper weighting?***

AR: *The proper weighting should take care of inhomogeneous soil moisture profiles, although uncertainty may be involved due to the lack of continuous data in the profiles. Detailed investigation of the corresponding uncertainties will be dedicated to another, currently active study.*

3.3. P15 L26

RC: ***so that installation bias can lead to bias in area-average soil moisture based on averaging sensor data; can you elaborate on this possibility as it is as likely as bias in CRNS data.***

AR: *In general this is correct for soil profile monitoring (see general reviewer comments above). However, the statement in the text refers to the difference between soil profile measurements and water storages at or above the surface. Therefore we see no need to change the present formulation:*

> Ponded water in local hollows, trenches, and the litter layer are not visible in the soil profiles of the monitoring network, which are typically installed in solid and elevated ground.

3.4. P15 L29 "By comparing CRNS data and point measurements, residual information could be used to identify additional processes like biomass dynamics or rainfall interception"

RC: ***if, and only if, the point measurements are correct AND give true average moisture field; neither is assured.***

AR: *We agree on this comment, however this is a general issue (or assumption) for soil moisture monitoring networks. We thus see no need to add this comment to the argumentation.*

3.5. P13 L10-12

RC: ***so that is the same as giving the average within each land use type to each 1 m2 pixel within this type? what if there is a significant variation within land use type?***

AR: *Yes. Spatial interpolation is always site-specific and should be guided by expert knowledge about local conditions. In our case, mean soil moisture in the four parts had significantly distinct values (as shown by dotted lines in Fig. 9) and variations within the land use types were insignificant (not shown). Since the CRNS represents a large-scale average of soil moisture, we think it is reasonable to smooth small-scale variability of point measurements among an area of comparable characteristics.*

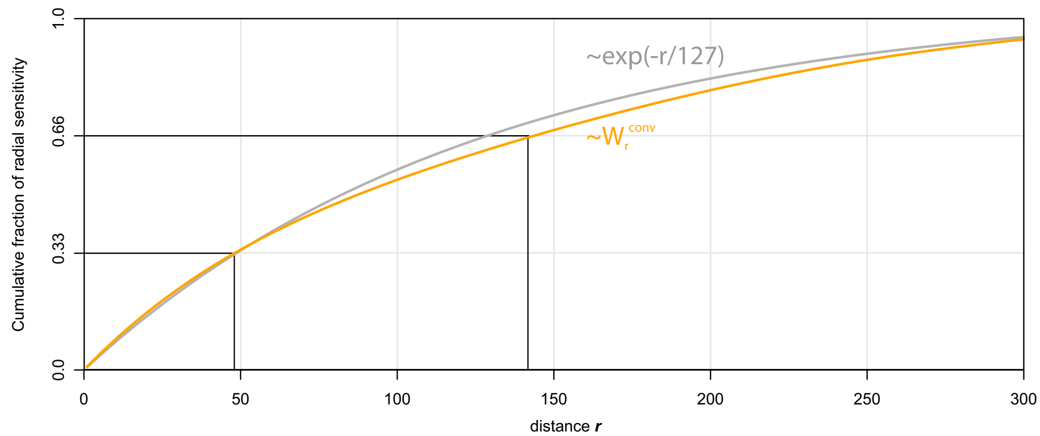

Figure 2: (not in the paper) Cumulative fractions of radial sensitivity, dN/dr, for the conventional weighting functions (exponential, and approximation by Bogena et al. 2013).

3.6. P19 L7 "{25m,75m,200m}"

RC: ***I see this as 25, 75, 175 in Zreda et al. (2012), which suggests rounding to the nearest 25 m; perhaps what Franz et al. (2012) did was to round up the 175 m value to 200? or maybe it is higher because Franz's field site was not at sea level, and thus the footprint was be larger.***

AR: *The latter is true. They roughly adapted the sample radii on the altitude of the individual sites, refering to the dependency of the footprint radius on air pressure as was published by Desilets and Zreda (2013). I guess the authors further rounded the numbers 25, 75, 200, and 175 to the nearest 25 meters for the sake of simplicity. This approach might have had reasons at their individual, almost homogeneous sites, however we think that more accurate sampling and weighting should be performed in any heterogeneous terrain. Which motivated us to write this manuscript.*

3.7. P19 L9

RC: ***These cutoff values seem little too high for exponential decrease with distance according to exp(-distance/127), which you used before. Can you double check it? Is the value of 127 m computed with the increasing area as the distance increases? (It looks like it to me, but please confirm.)***

AR: *Thanks for finding this detail. It is true that the weighting functions account for the increasing area as the distance increases. The cutoff values are indeed too high, as we calculated them by integrating the conventional function beyond 300m. However, as written in section 2.2, Bogena et al. 2013 only defined it for r<300m. We recalculated the cutoff values and will correct the equation in the revision. The new values (48 and 142) even better match the conventional sampling radii (25,75,175) in the areal centers. Please find also attached a figure that shows the cumulative fraction of neutron count sensitivity over distance. The function from Bogena et al. 2013 (which we used) is only an approximation to the exponential function.*

3.8. P19 L13-15

RC: ***this is a very important point; I reviewed a paper in which most of the the calibration points were more than 100 m from the probe, in effect producing questionable (at best) calibration; can you elaborate on***

this to make it sound stronger/clearer?

AR: *We dedicated the whole section on this topic, tried to recommend to take more samples within the nearest few meters, and concluded that sampling schemes should depend on local site conditions rather than follow a generel convention. We think we emphasized these results in the conclusions section, particularly in item 5:*

> Data points in the first 0 to 10 m radius and 0 to 20 cm depth around the sensor are most important for calibration and validation purposes. It is thus recommended to reduce the uncertainty of those measurements, e.g., by avoiding flints in the samples, or by increasing the number of samples in that area.

3.9. P19 L18-19

RC: ***not sure if this is the only problem; the calibration problem consists of many subproblems, of which two have to do with spatial variability of soil moisture; one is how representative the data points are; the other is the weights; if you collect half of the calibration points within 25 m, the area enclosed within that 25 m will be well (subjective term) characterized; the other half of the points will be, I guess, beyond the 25 m perimeter; regardless of how far you will go, that area will be less well characterized because you will have fewer points per unit area there than you had in the area between 0 m and 25 m. so the way out would be to increase the number of calibration points beyond the 25 m perimeter. How exactly this would work out? I have struggled myself to find out, as it would depend on the degree of heterogeneity of soil moisture. A brief discussion of this issue would make the paper more correct and more relevant.***

AR: *Thank you for pointing this out, this is indeed a challenging task shared by all of the evaluated calibration approaches. But, along with all the aspects discussed in the manuscript, the sampling scheme needs to be feasible. As a result, the number of calibration points is practically limited to a number that can be handled within e.g. one day. As a consequence, we think that there is the clear need to optimize the sampling scheme also under the aspect of feasibility. In general we think it is highly advisable to take more measurements closer to the footprint than far afield for the following reasons: (1) the CRNS probe is most sensitive to the first few meters, hence the measurement uncertainty by the soil samples in that area should be as low as possible; (2) the CRNS probe is less sensitive to remote areas, such that uncertainty due to soil sampling has less impact to the calibration result. Will change the text accordingly.*

3.10. P19 L21 "sampling radii R_i of concentric circles could be calculated"

RC: ***actually, wouldn't it be better not to sample on these circles, but within the areas they encompass? it may or may not make much difference, but it would look better (more representative) when mapped. Or perhaps this is, in effect, what you will accomplish by increasing the number of circles from three (as in Zreda 2012) to more than three, as you discuss below.***

AR: *We actually recommend to sample within the area which the circles of radii r_i encompass. It should not be confused with the radii R_i at which the actual sampling takes place. The sentence to which the reviewer refers however deals with the sampling radii R_i. We think this is clear from the text:*

To give further advice on a reasonable distribution of points for homogeneous terrain, sampling radii R_i of concentric circles could be calculated as follows. First, select a total number of circles n based on prior knowledge about the patterns at the individual site. Since the signal contribution of an area between any radii can be calculated by integrating W_r (compare also Köhli et al., 2015, eq. 1), the n borders of equal areal contribution, $r_i, i \in (1, ..., n)$, can be calculated by solving the integral: ...

3.11. P20 L10

RC: ***equal number will not give equal quality of the average - see my comment two comments above; we would need more sites away from the probe***

AR: *This is true for a sensitivity function that is proportional to $1/r$. The sensitivity functions used here are highly non-linear, such that remote areas have a much less influence on the neutron signal. Given the small range of their contribution, we think it is not required to generate the same measurement quality (i.e., precision) in near and remote areas. We will clarify this in the text.*

3.12. P20 L11

RC: ***I think this is bordering on nonsense - what if the <2 m is affected by the installation of the probe or probes, which is quite likely?***

AR: *If this is the case, the corresponding material moisture should be taken into account (e.g., concrete at $\approx 8\,\%_v$ moisture content). However, most CRNS stations known to the authors do not require significant preparation of the ground, as the detector is often mounted on a thin metallic stick burried in the ground. We further like to emphasize again that the reocmmendations given here have an exemplary status. They cannot be generally applicable, since we argued further in this section that the best sampling scheme always depends on local heterogeneity and site conditions.*

RC: ***Also, if I remember correctly, the maximum sensitivity is NOT at the probe, but some meters away (because of the increasing area), so the <2 m is not justified. Please, rethink and provide an updated figure.***

AR: *The maximum sensitivity is found within the first meter (Fig. 2). We therefore think that it is required to achieve high precision of the soil moisture measurements within this area.*

3.13. P20 L16 "compare also Heidbüchel et al., 2016, section 4.2"

RC: ***but Heidbuechel took it straight from Koehli 2015, so it is not independent***

AR: *Correct, but we did not state that the research is independent. In contrast, we mentioned in the Introduction that Heidbüchel et al. 2016 already applied an adaption of the revised weighting functions. Their study serves as a good example here.*

3.14. P21 L11-13

RC: ***I am not sure what you mean here. Can you rewrite to make clear?***

AR: *Thank you, this paragraph presents another approach to select locations of point measurements in the footprint. The approach accounts for the radial shape of the sensitivity function, however it is independent of dynamic variations. We will rewrite the text to clarify this.*

3.15. P22 L24-26

RC: ***this is the most important point from the practical point of view - it means that we do not need 2+ calibration campaigns, as suggested by Heidbuechel for example, but a single one, properly weighted, will do. This is a great result! You should emphasized it more strongly.***

AR: *Thank you, we will emphasize this key point.*

3.16. P22 L30 "Sampling strategies that are based on concentric rings can only be recommended for homogeneous terrain"

RC: ***but this is not known a priori, so this recommendation is not very useful to a field hydrologist.***

AR: *State-of-the-art research about optimal design and sampling strategies recommends to take into account additional information (proxies) about local structures at the individual sites (Diggle and Lophaven, 2006; Nowak et al., 2010; Bramer et al., 2013; among others). We therefore think it is reasonable to recommend* a priori *site assessment in the footprint of the CRNS probe in order to achieve best results for the sensor calibration.*

3.17. P22 L11-13

RC: ***so how is ponding otherwise invisible? and, more importantly, would you be able to see these things if you look at the neutron data alone? would you be able to separate interception from ponding? I have doubts, and should you have similar doubts, please spell it out.***

AR: *Thank you for the clarification, we will rephrase the text to make this clear. Regarding the invisibility: Ponding and other water storages are invisible for conventional soil moisture sensors. If the CRNS is calibrated against the wrong average, the relevant changes of neutron counts could be interpreted as noise, and identification of other water storages could become impossible.*