Response to Interactive comment by H. Bogena (Referee)

Comments from referee are printed in black. Authors' responses are printed in red.

Résumé

This study concerns the application of the cosmic-ray neutron method to monitor soil moisture in a mixed forest in the lowlands of north-eastern Germany. The authors tested several calibration procedures using soil samples taken 10 times within one year inside the footprint of the sensor. A two-point calibration is assumed to be adequate to correctly define the shape of the N0-calibration function with adjusted parameters when calibration points were taken during both dry and wet conditions covering at least 50% of the total range of soil moisture.

This paper is an interesting presentation of a forest application of the cosmic-ray neutron method giving some new insights into the calibration process. It is also very well written and fits well to the scope of HESS. However, some methodological improvements need to be undertaken as outlined in my comments. In addition, the results should be better discussed in the light of existing publications on the calibration of the cosmic-ray neutron probe.

We want to thank Heye Bogena for a detailed and thoughtful review that helped in significantly improving our manuscript.

General comments

The calibration results might have been affected by the unfavourable locations of the sampling locations. As demonstrated by Köhli et al. (2015), the highest contribution comes from the first 10 m radius, whereas the nearest sampling locations are still 25 m away from the CRNS probe. Please add a discussion on how thus soil moisture differences between the close-up area and the sampling locations might have affected the calibration results.

Fortunately, the conditions within the closest 30 m around the CRS are quite homogenous at our field site since the sensor is located within a pure beech stand. So we expect a similar range of soil moisture conditions at our sampling locations at a distance of 25 m away as compared to those at locations closer to the sensor. Thus, the near-field sampling design should not influence the calibration results. Our aim was to make the sensor work for our purposes in our environment. However, in the future we have to think about what it really is that we are measuring. Is the CRS a measuring device that can only give us reliable information on soil moisture within a radius of 10 m? Maybe the change in the shape of the calibration function that we found is also a result of trying to counteract/compensate for spatial heterogeneity. We will add some of this discussion to the manuscript.

The vegetation within the footprint of the CRS probe is quite heterogeneous (please add a table of the landuse contributions), which complicates the spatial averaging of soil moisture. For instance, as shown by the authors, the coniferous sites are consistently dryer compared to the areas covered by beech. Thus, the limited number of sampling locations could be an additional reason for the differences in the CRS calibrations.

This is true. We state in the manuscript that this variability within the footprint is a possible cause for the differences in our calibration results. We also tested whether an area-weighting function based on the different tree stands significantly changes our results. We found that it did not make a difference. We will add a table of land use as recommended.

Iwema et al. (2015) already showed that using only one calibration data can lead to large uncertainties especially in humid regions with large hydrogen pools. They showed that the best trade-off between number of calibration dates and calibration accuracy can be achieved by using 6 calibration dates. Please discuss your results in the light of the results of this recent publication.
As far as we understand, Iwema et al. state that using more than 6 random sampling dates does not further improve the calibration results. They also state that using sampling dates with dissimilar moisture conditions ('appropriate soil wetness conditions') can reduce the required number of sampling dates. We found that using only 2 sampling dates with a large enough difference in soil water content suffices to achieve a good calibration. We will add this to the discussion.

This study also considers the vegetation correction developed by Baatz et al. (2015). However, since the method considers only linear scaling of the neutron counts, its application should not alter the calibration accuracy in case of temporally stable above ground biomass (as in this study). Therefore the application makes only sense where temporal biomass dynamics are expected and temporal information on biomass changes are available or in case of cosmic rover applications.
This is true. Once we had analyzed this scenario, the expected outcome (i.e. that it makes no difference) became obvious. We will remove this approach from the entire manuscript. We also tried a different approach by Baroni & Oswald which accounts for seasonally changing biomass in the weighting function. This approach also did not improve our results (probably because the amount of seasonally variable biomass in between the sensor and the soil is not large enough at our filed site).

Specific comments
L143-145: On the web-site of the sensor manufacturer no specification of the measurement technique is given. I suspect that these sensors are actually based on an oscillator-ring as described in Qu et al. (2014) and not on the time-domain transmission technique. In addition, a problem of these sensors could be the top shielding, influencing the soil water content below. Since these kind of sensors only measure soil moisture at a very small volume (only very few centimeters around the sensor blade) this might lead to systematic underestimations of soil moisture.
According to our sensor manual the sensors are indeed based on time-domain transmission. It is hard to find other online information: http://www.ibot.cas.cz/en/kalibrace_stanice_projekt. When calibrating the sensors in the field, we actually found a systematic overestimation at low soil water contents that we corrected for. We did not detect a shielding effect which would cause an underestimation of soil water content.

L147-148: Is the sensor blade actually 15 cm long (at the web site there is no information on the seize and the pictures suggest that the sensor blade to be much shorter)

Yes, the sensor blade is actually 16 cm long.

L149: Why didn't you use all data for the calibration?
The sensors were installed in May 2014 only, so we could not use the data for the first 5 calibration campaigns.

L205-208: No scaling needed since this correction considers the relative changes in incoming neutron flux. However, the cutoff rigidity of the Jungfraujoch Station is somewhat different from the study site given is lower latitude. An good choice for the neutron monitor is be the Lomnicky station, Slovakia (LMKS).
We agree that the scaling is unnecessary. Therefore we will shorten the paragraph and move the figure to the supplements. We will still use the scaled and gap-filled time series we computed since we do not want to redo all the analysis (it should not make any difference, as you say).

L206: The correct unit for incoming neutrons is "counts/sec"
We will change all occurrences of 'n h^{-1}' to 'counts h^{-1}'. To stay consistent and avoid confusion we will also use hours for the incoming flux corrections.

L224: The methods to determine soil organic carbon and root biomass water equivalents are not presented.
A description of the methods will be added.

L230: This statement is too vague. I think what you meant here is that the objective performance measure is minimized, right?
True, the sentence will be changed accordingly.

L315: What about the other tree species?
We did not conduct surveys on the other tree species. The beech stand covers 55.5% of the footprint around the CRS (when assuming the exponential distance-weighting from Zreda et al. (2008)). Pine covers 16%, spruce covers13%, oak covers 8%. With the new distance weighting function of Köhli et al. (2015), the numbers of the other tree species will decrease even further. Also, the seasonal variation in spruce and pine above-ground biomass is very small and thus we consider it to be constant in this study.

L317: On which grounds did you assume these values?
These values are taken from the literature (papers by Gravano et al. (1999) and Bouriaud et al. (2004)). See line 320.

L359: The correct unit for incoming neutrons is "counts/sec"
This will be corrected in the revised manuscript.

L380: Change to "the same value for the N0 calibration parameter"
Will be changed to: 'Ideally, we would have ended up with the same value for the N0-calibration parameter for each of the 10 calibrations.'

L381: The correct unit for the CRS measured neutron intensity is "counts/h"
This will be corrected in the revised manuscript.

L382-387: According to Zreda et al. (2012) the presence of other hydrogen pools than soil moisture increases the stopping power of the soil, which leads to a change in the slope of the calibration function. Thus, calibration has to be performed using the total hydrogen pool, and soil moisture is then computed by subtracting other hydrogen pools than soil moisture from the measured neutron-derived soil moisture. It is unclear whether this procedure was applied in this study. If not, this would partly explain the differences in soil moisture estimates.
As outlined in the description of our four approaches, we used the total hydrogen pool for calibration before subtracting other than soil moisture contributions according to Eq. 4.

L388: The term "new calibration function" is misleading. Changing the "a" parameters of the N0 calibration function is not new and was already presented by Iwema et al. (2015). They called this more adequately "modified N0 method". However, they only calibrated 3 parameters (the N0 parameter was omitted), because of strong correlations between the parameters leading to ambiguous calibration results (equifinality problem). Did you check for this calibration issue?
You are right that 'new calibration function' is misleading. We will change all occurrences to 'modified calibration function'. It is true that the N0 parameter has a very similar influence on the shape of the calibration function as the a0 parameter. Still we don't think that in this case equifinality (we agree that it exists) is a real problem for this application because the goal of the calibration simply is to find an efficient function that represents the calibration points. There is no reason to consider any potential adverse implications of the adjusted parameter values or combinations of them. As a side note: we also tested whether simple exponential or gamma functions would perform as well as the 4 parameter calibration function and we found that they in fact did not. So there seems to be justification for the specific set of shapes that is described by the N0-calibration function.

L400-402: Do you have any idea why?
Yes. If you look at the new Fig. 5 you see that at higher water content a smaller change in neutron counts is associated with a larger change in soil water content (the function is steeper). Therefore the uncertainty during the calibration is also larger.

L407: See comment L388

L414-416: Please provide a figure showing the comparison.
We have prepared a figure and will add it as a supplement (Fig S2).

L422-424: Shouldn't the relationships vary with soil moisture content due changing sensor penetration depths?
Yes, that is what we expected and that is also what we found for the first two approaches. When it is wetter, the penetration depth is reduced for the CRS measurements and the wetter shallower layers receive more weight. Therefore, the CRS measurements show higher SWC than the gravimetrically determined SWC. However, it seems that the distance weighting counters this effect. A probable explanation is that the formula used for the distance-depth weighting increases the critical depth. This causes

higher weights for deeper (drier) soil layers even under wet conditions and could counteract the trend.

L429-430: This finding is quite obvious given the insignificant changes in above biomass. Generally, the application of the vegetation correction makes only sense, when temporal biomass dynamics are expected and temporal information on biomass changes are available.
We will remove the whole part on vegetation correction.

L441: This investigation is very similar to Iwema et al. (2015). Please discuss your results in the light of this study.
We will add this to the end of the discussion.

L450-457: The results plotted in Fig. 12 show clearly, that only the most extreme dry and wet samplings result in an acceptable calibration result, whereas sampling at intermediate soil moisture will lead to very uncertain calibration of the modified N0-method. On the other hand, this illustrates the value of the standard N0-method that will also produce stable results in case only one sampling date is available. Please add this to the discussion.
The new Fig. 9 (old Fig. 12) will show that the best 2-point-calibrations are achieved with one sampling point taken under very dry conditions and another sampling point taken either under intermediate or wet conditions. In our case it is hard to see the value of the standard N0-method since it always resulted in too much soil moisture variability no matter whether the calibration was performed during wet, intermediate or dry conditions (because the standard calibration of N0 does not allow a change of the slope of the calibration function).

L458: This chapter belongs to discussion.
We will move parts of this chapter to discussion and only leave the parts that really describe results. We will also add results on other hydrogen pools, so we will rename the chapter.

L471-474: This is only true when assuming that the CRS footprint is completely covered by beech, which is however not the case.
This is true. So for this calculation, we assume an extreme case since the other vegetation types experience smaller seasonal changes in above-ground biomass. In reality we should expect even less variation in neutron counts due to foliation/defoliation. We will add this statement to the revised manuscript.

L484-485: So the whole discussion of this chapter is unimportant and should be reduced to 1-2 sentences.
Would you say that just because our results suggest that seasonally-varying above-ground biomass does not influence the neutron count significantly the discussion of this finding is not important? We think this finding is very important for the use of CRS in forested areas and worth the extended calculation and discussion. (In the end, it makes life much easier when applying CRS in forests).

L488-518: This section is a summary, not a discussion and thus should be omitted.

We will restructure the discussion section.

L520-528: Please discuss your results in the light of the results found by Iwema et al. (2015).
We will add a discussion of the findings of Iwema et al. (2015).

L558-560: This statement is not clear to me. Please explain in more detail.
We will change and add more detail: ‘Following the argumentation of Lv et al. (2014), the fact that distance weighting improved our results can be regarded as an indication that non-homogeneous soil moisture conditions indeed lead to changes in the shape of the calibration function. At our site distance weighting reduced the spatial variability within the footprint of the sensor since it assigned higher weights to the closest sampling sites which were all located in the relatively wet beech forest, while the influence of the drier soils under the coniferous trees was reduced.’

L564-569: This part is somewhat misleading. Corrections of the neutron count rate (Eqs. 1-3) are essential for any application of the CRS (e.g. Zreda et al., 2012). Vegetation correction is only needed for sites with significant biomass changes. On the other hand, the characterization of the temporal stable hydrogen pools is important for the application of the N0-method. However, the abundance of different pools and the uncertainties in the sensing depth estimation will always lead to uncertainties in the calibration process. As shown by Iwema et al. (2015) and by the results found in this study, this issue can be partly circumvented by the using site specific calibration parameters estimated at using in-situ samples taken during dry and wet conditions. Please reformulate in this sense.
Will be changed to: ‘If it was possible to fully correct for all factors that influence footprint size, depth-weighting and neutron count, a one-time calibration of the CRS would be sufficient. However, the abundance of different hydrogen pools and the uncertainties in the sensing depth estimation will always lead to uncertainties in the calibration process. Therefore we argue that for using the CRS as a simple tool to measure soil water content at intermediate scales, the efforts of measuring all necessary parameters are not justified. As shown by Iwema et al. (2015) and by the results of this study, this issue can be partly circumvented by using site-specific calibration parameters estimated from in-situ samples taken during dry and wet conditions. Hence, we recommend a two-time calibration that – although being empirical in nature – inherently incorporates many of the required corrections.’.

L583-584: Actually, the seasonal changes of the hydrogen pools in this forest site are negligible. Thus vegetation correction can be omitted.
Agreed.

L594-600: This statement is based on Köhli et al. (2015), but not on results of this study and thus should be omitted.
Since this a very important statement and should be considered by everybody using a CRS, we would like to keep it. But we will add the proper reference (Köhli et al.).

L606-608: This step is obvious and should be omitted.

When we first set up the sensor and calibrated it for the first time we brought a battery with us to let the sensor run. After we had finished the soil collection, we took the battery back home with us. Only later it became clear that we should have collected neutron counts for a longer period of time. So we could not use the data from our first calibration effort. This was a hard lesson to learn and we want to make sure that other people do not make the same, admittedly stupid, mistake.

L614: The sampling locations should be adapted to the footprint estimates after Köhli et al. (2015).
Our aim was to find the best way to calibrate the CRS for soil moisture measurements within a radius of ~300 m using the standard calibration procedure. We cannot provide recommendations for a different sampling pattern since we have not tested it. We will add this to the discussion and we will also add a recommendation to have homogeneous conditions within the first 30 m around the sensor.

Figures
Figure 1: This map should be integrated in figure 2.
Will be done.

Figure 2: According to recent results of Köhli et al. (2015) the footprint is considerably smaller than 300 m. Please adapt the figure. In addition, it would be helpful to color the aerial photograph according to the different tree species.
Since this figure is also supposed to illustrate our sampling scheme we would like to keep the 'old' footprint size. To make that clear we will rephrase: 'The yellow circle approximates the footprint of the CRS as it was assumed when sampling took place'. The distribution of different tree species can be seen on new Fig. 3 (old Fig. 6) and it would probably make this Figure too busy adding colors or patterns on top of the photograph.

Figure 3: This figure can be omitted (see comment L205-208)
We will move the Figure to the Supplement (S1).

Figure 4: This schematic figure is wrong in presenting the cosmic-ray neutron intensities as actual rays that are reflected by the soil. The actual processes leading to neutron intensity are far more complex (see e.g. Köhli et al., 2015) and should not be presented in this way in a scientific paper. Also the above ground and below ground footprints are not connected in the simple way as suggested by the schematic drawing. Thus, the figure should be omitted.
We acknowledge the fact that this figure simplifies the actual processes a lot and will add this statement to the caption. We will modify the figure to resolve two of your concerns. The neutron intensities will not any longer be depicted as rays and the above-ground footprint will be removed entirely. We still think that the figure helps to get an overview over the many parameters that have to be accounted for before/during the use of the CRS method.

Figure 12: The Pareto front needs to be discussed in the text as well.
We realized that it is actually not a Pareto front. So we will add to the text: 'The existence of a rather clear front in Fig. 9 indicates that the calibrated neutron count-soil

water content conversion will always perform well if the soil moisture differences between the two calibrations are sufficiently large.’