Response to Interactive comment by Anonymous Referee #3

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

The manuscript describes an improved calibration of cosmic-ray soil moisture probes. The problem of calibration has surfaced many times before and different scientists have proposed improvements. But even with those improvements, some problems have remained. This paper adds a valuable component to the growing body of knowledge of how cosmic-ray neutrons respond to surface moisture and how best to calibrate the response functions. The work described in the paper is thoughtfully designed and carefully executed and analyzed. The conclusion is simple, but important: calibration function is different than that originally proposed, and that difference can be captured by doing only two calibrations at two different levels of soil moisture. The good quality of work and the important conclusion make the paper important. I recommend that it be published with minor revisions suggested on the annotated copy (attached).
Thanks to reviewer #3 for a considerate and helpful review that helped in improving the paper.

Line 9: this is instrument, not technique. The technique is "measuring soil moisture with cosmic-ray neutrons".
Will be changed to: 'Measuring soil moisture with cosmic ray neutrons is a promising technique for intermediate spatial scales.'

Line 12: Desilets' function was computed using MCNPX
Will be changed to: 'The calibration is based on soil water content derived directly from soil samples taken within the footprint of the sensor.'

Line 21: this conclusion is based on ten calibration data sets, but these ten sets span a narrow range of average soil moisture values between 7% and 16% (Fig. 6) - so how solid is this conclusion?
Soil moisture in our very sandy soils only ranges between 5% and 25%. So although 16% still appears to be on the dry end (when comparing it to more silty/clayey soils) we are covering 50% of the total range of possible soil moisture. Moreover, it will be difficult to calibrate at even higher soil water contents since they usually only appear for a short period of time after large precipitation events. At that time, however, conditions for calibration are not optimal since there will still also be intercepted water in the canopy and litter layer. Also, keep in mind that the values on new Fig. 3 (old Fig. 6) are non-weighted averages of all 108 soil samples. The weighted averages used for calibration range from 8% to 22% (see Table 3, column: θ_{depthW}). We will add some of this information to the manuscript.

Line 32: meaning? seems unnecessary
Agreed. Will be removed.

Line 34: the largest spatial variance is usually at intermediate moisture levels, and decreases towards both dry and wet ends
Will be changed to: '…especially under intermediate conditions…'

Line 37: delete "very". perhaps rephrase this to provide more specific information, for example: "this is time-consuming and expensive"
Will be changed to: ‘This can be time-consuming and expensive.’

Line 45: not from the sun (not energetic enough to produce cascade neutrons); say something like "Cosmic-ray neutrons on Earth are formed by high-energy protons coming from galactic sources, such as supernovae." cascade neutrons are high-energy neutrons, not fast neutrons. Fast neutrons are produced by secondary neutrons via the process of evaporation. So the fast neutrons are tertiary neutrons, by that term is not used; term evaporation neutrons is used instead.
Thanks for this clarification. This is quite complicated. Will be changed to: ‘Cosmic ray neutrons on Earth are formed when high-energy protons deriving from galactic sources (such as supernovae) enter the Earth’s atmosphere. Once in the atmosphere, the protons start interacting with atomic nuclei (mainly nitrogen and oxygen) producing cascades of secondary neutrons (that are also called high-energy neutrons) that travel towards the Earth’s surface and into the soils. When secondary neutrons interact with air or soil they trigger the release (evaporation) of fast (low-energy) neutrons.’

Line 61: specify that Kodama's detector was buried in the soil.
We will add: ‘…when burying a neutron sensor in the soil.’

Line 63: larget than what? perhaps say "large" or specify that area (~30 ha).
We will add: ‘…(~30 ha)…’.

Line 68: this is not appropriate reference; the range 10 cm - 70 cm was given in Zreda et al., 2008. Also, the dependence on soil water content was specified, but not dependence on soil type and distance from sensor. Koehli et al., 2015 worked out the dependence on distance.
We will exchange the references.

Line 102: why unwanted? perhaps better to say "adds additional signal" that complicates analysis (or complicates extraction of soil moisture signal).
Will be changed to: ‘…and adds additional temporal variability to the CRS time series complicating the extraction of the soil moisture signal.’

Line 108: please make sure that this is true; I think the idea is to fix the calibration parameter at the high end (water), rather than low end (dry soil, in Desilet's calibration function), but that is a regression parameter, not measured value.
In the 2013 paper they say that this parameter can be easily retrieved from measurements over a large water body. We think you are right in noting that it is in fact a regression parameter. So we will change the wording to make this clearer.

Line 135: Are they applicable also at the site? 1.6 km seems close, but precipitation and humidity can vary at smaller scales.
We compared the climate data with data we collected ca. 400 m away from our site and the differences were marginal. We decided to use the data from 1.6 km since it was a longer data set (we only installed the 400 m sensors in the middle of 2014).

Line 175: and assuming the sensitivity decrease with distance given in Zreda et al., 2008, for the footprint radius of ca 300 m. The new footprint estimate by Koehli et al., 2015, gives smaller footprint, so the sampling distances for equal weights will be smaller than your 25 m, 75 m and 200 m.
This is true. We will discuss this in more detail later in the manuscript.

Line 204: this correction uses shielding depths (g/cm2), not pressure (hPa); however, the difference in this case would be cosmetic (or zero, if the value of 133.3 hPa was obtained from the equivalent shielding depth), so it does not matter for the results; on the other hand, if you want to be consistent with cosmic-ray literature, please make the change to shielding depth.
Since we do not know how to convert pressure to equivalent shielding depth and vice versa, we will not change anything here. Is there a conversion formula you could provide us with?

Line 207: reference Zreda et al., 2012 for this correction
We will add the reference.

Line 216: this difference (309 vs 327) cannot possibly be due to difference in altitude between the monitors (which should be a factor of four or more). I think these data sets are pressure corrected, and the difference idiosyncrasies of the two neutron monitors.
We will delete the mention of altitude.

Line 219: reference Rosolem et al., 2013 for this correction
Reference to Rosolem et al., 2013 will be given.

Line 320: is this 0.6 kg of H2O per kg of dry biomass? or wet biomass? clarify.
It is actually 0.6 kg of H2O per kg of wet biomass.

Line 321: this seems low. does this figure include free water (H2O in xylem) and cellulose-bound water (OH group)?
No, this just includes free water. The cellulose-bound water is calculated in the next step (assuming hydrogen content in dry biomass is 0.0622 kg kg-1).

Line 335: isn't there a seasonality or some other temporal variability (eg, with droughts) in the free water within trunks and branches?
Yes, there is some seasonality also in this regard (there is even daily fluctuations indicating variations in transpiration flux). This, however, was not part of our analysis since it is difficult to determine the exact numbers and it is likely that these variations are too small to influence the neutron count.

Line 342: This section should provide more information, not just the equation. What is KGE? How is it used? What is compared with what? How are these variables computed (eg, r)? What is the significance of the result? Etc...

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2} * \sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

We will add the equation for the correlation coefficient r and some more information on the KGE': 'The KGE' measures the Euclidian distance in a 3-D space where the correlation coefficient r is on one axis, the variability ratio β is on the second axis and the bias ratio γ is on the third axis. KGE' scores range from 1 (representing a perfect fit) to -∞.'

Line 398: this is inaccurate; one time series shown in Fig. 7 has differences of that magnitude, whereas the other nine have considerably smaller deviations; please, state this result correctly, without creating undue alarm.
You are right in that our formulation is unclear and could be interpreted in different ways. So we will modify the statement: 'As a consequence, the 10 computed time series based on the standard N0-calibration function of Desilets et al. (2010) also showed differences in volumetric soil water content (Fig. 4 illustrates results for the DDW approach). In the most extreme case, these differences were larger than 0.1 m3 m-3 (which is equal to 50 % of the total range of soil water content at the site), especially during conditions of high soil moisture.'.

Line 403: It is unclear how this conclusion was reached. Please describe in detail what was done, what were the results and how to interpret them.
We conclude that there is no unique calibration parameter and will describe the calibration procedure in detail in the methods section. The results are presented in Table 4 and new Fig. 4 & 5 (old Fig. 7 & 8). We will add a more detailed interpretation of the results to the discussion section.

Line 420: There is a lot of noise in Figure 8 and the difference between standard N0 and the improved data is not clear. Can you add error bars to the data points?
We are not sure about this comment. In our view, this difference between the standard and the modified calibration functions is clearly represented by the dotted and the solid lines respectively. Since there is already a lot going on in this figure, we decided not to add error bars and we were not sure of the added value they would provide.

Line 553: The absolute count rate has no influence on the shape of the response function, just on the precision of calibration. I would remove this conclusion or reword it to make this conclusion ("it clearly does not at our site") less strong.
Ok. We will remove this conclusion.

Line 564: but on page 9831 (of my copy) you stated that the sensitivity is better when using your calibration. Please, make these two statements consistent.
On page 9831 we stated that the sensitivity of the sensor is essentially higher than it should be (not better). This means that already a small difference in neutron counts indicates a large difference in soil moisture. The modified calibration accounts for this by decreasing the slope of the calibration function and thereby reducing the sensitivity of the sensor (so that now a bigger difference in neutron counts is required to cause differences in the soil water content reading). We will modify our statement on page 9831 to: '…essentially the sensor appears to be more sensitive than one would expect.'

Line 642: Franz et al 2012 is not the correct reference; Franz et al merely repackaged the information given in Zreda et al 2012.

Franz et al. 2012 describe 3 circles with distances around the CRS of 25, 75 and 200 m. Zreda et al. 2012 describe 3 circles with distances around the CRS of 25, 75 and 175 m. Since we used 25, 75 and 200 m the more correct citation to describe our calibration setup would be Franz et al. 2012. However, to recognize the contribution of Zreda et al. we will insert the reference: '…we followed the recommended sampling pattern for the calibration of CRS which was developed by Zreda et al. (2012) and slightly modified and detailed in Franz et al. (2012b).'.

Table 3: Insert line!
I would love to. Actually there was a line when we submitted the manuscript. But it seems that HESS does not allow lines within tables. However, I will try to convince them to leave this line in there because I also think it is necessary.

Table 4: I am not sure how to read these results. Can you clarify in the figure caption or in text (which also glosses over this in one short paragraph at the end of section 4.3).
We will modify the text to make the table clearer: 'All resulted in largely deviating N0-values between the individual calibrations (see means and standard deviations in column 1 and 2 of Table 4). This in turn led to differences in the time series of volumetric soil water content between the individual calibrations (see means and standard deviations in column 3 and 4 of Table 4).'

Fig. 4: can you group these factors into two groups: (1) constant in time (eg, WL), and (2) variable in time (eg, water vapor)? And then perhaps also into another groups: (3) those measured at calibration (eg, WL), and (4) those measured as time series (eg, water vapor)? The significance is that some parameters are easy to handle, others are difficult, and the two should not be commingled.
We will revise the figure according to your suggestions. We will group the factors into two groups and mention which of them need to be monitored continuously and which can be accounted for during calibration.