<u>Reviewer #2:</u>

We thank reviewer #2 for taking the time to review our manuscript and their valuable suggestions regarding our manuscript. We are certain that these comments greatly improve our manuscript and they will be incorporated in a revised version of the manuscript. In the following section we will reply to all comments of reviewer #2 with R2-1 (i.e. reviewer 2, comment 1) and A2-1 (i.e. author response to R2-1), respectively.

R2-1: "The more complex response of the thermal neutrons rise my main concern of the present study. The potential use of thermal and epithermal neutrons has been pointed since the first CRNS publications (Desilets et al., 2010; Rivera Villarreyes et al., 2011). Further attempts have also been performed later by dedicated studies (Bogena et al., 2020; Jakobi et al., 2018; Tian et al., 2016). As far as I understood, difficulties to handle these two signals are related to the non-unique response of the thermal neutrons, i.e., in contrast to epithermal, they depend on chemistry and the thermal intensity also increases during the wetting of initially dry soils (Desilets et al., 2010; Zweck et al., 2013). The present study quantifies the different footprint of thermal and epithermal neutrons. As such, it sheds lights on the understanding on the processes. However, few is discussed on the possibility to generalize the correlation found in the present study between epithermal and thermal in other conditions. As such I see the high risk of this study to be very limited. In addition, if the Authors are really interested on disentangling footprint variability, I rather believe that the use of side-shielded detector (Zreda et al., 2021) could be easier and more promising than the use of the thermal detector."

A2-1: As the reviewer pointed out and as we stated in the manuscript, the response of thermal neutrons to soil water changes is more complex and does also depend on soil chemistry. As pointed out by Weimar et al. (2020), apart from the moderation optimum there is a general decrease of the thermal neutron intensity with increasing soil moisture but with a less steep slope compared to epithermal neutrons. At the study site described in the present study, the neutron response of epithermal neutrons is too steep for soil moisture changes observed in mineral soils close to the neutron detector. For this reason and at our study site, we can make use of thermal neutrons as a proxy for a different relationship between neutron intensity and soil water changes. In this case we agree with the reviewer, that the methodological approach for combining both neutron energies through the rescaling approach presented, is limited to the conditions of the study site and study sites with a similar setting. For this reason, we also successfully tested and presented another approach for improving the calibration against near-field reference measurements which does not involve thermal neutrons.

In contrast, using the relationship between thermal and epithermal neutrons for the identification footprint heterogeneity between the near-field and the far-field of the neutron detector is not limited by the more complex behaviour and non-unique response of thermal neutron to soil water changes because at study sites with homogeneous soil water states and dynamics, the relationship should not change. On the contrary, if soil moisture contents and patterns differ, differences should become visible as illustrated in Figure 8. This approach as a first indicator for footprint heterogeneity in terms of differences between the near-field and far-field of the instrument and could be applied at other study sites. Here, we would like to refer to the comment of reviewer #1 (R1-3) suggesting centre pivot irrigation experiments to test this indicator under more controlled conditions. Directional neutron detectors pose a great potential to the scientific community as they allow for reducing the effect of the soil, when a thick moderator shield is placed below the detector or intensify the soil moisture signal when the moderator shield is placed above the detector. Side-looking devices need to be tested in future studies. A potential drawback could be the fact that neutrons scatter and change directions several times before actually entering the detector. As a consequence, for example, a neutron being detected by a north-ward looking detector does not necessarily originate from this direction or had most of its elastic scattering interactions in this direction. This is different for a directional downward-looking neutron detector placed above the soil. Here, most neutrons are directly reflected from the soil and thus carry information of soil water contents. Nevertheless, the potential of directional neutron detectors for side-looking applications as well as CRNS in general should be explored in future and we will mention this in the discussion of our revised manuscript. Again, irrigation experiments and modelling efforts could reveal interesting insights and illustrate potentials and limitations.

R2-2: As far as I have understood, the comparison between neutrons simulations and neutron measurements is not consistent. Simulations are based on theoretical detectors sensitive only to thermal or epithermal ranges. In contrast, measurements have been collected with bare and moderated detectors that are contaminated by epithermal and thermal neutrons, respectively, as highlighted by the Authors. Previous studies showed clear discrepancy between simulations and measurements when this contamination effect was not properly account for (Andreasen et al., 2016; McJannet et al., 2014). As such, I'm surprised about this setting. Either the detectors should be improved to remove the contamination from thermal and epithermal. If this is not possible within the present study, why not repeating the simulations with the real detectors? Despite more rigorous understanding of thermal and epithermal, you are allowed to compare the simulations and the experimental data.

A2-2: We partly agree with the reviewer that the contribution of epithermal neutrons to the thermal neutron detector and vice versa influence the observed neutron intensities which hampers a direct comparison between observed neutron intensities and those obtained from neutron transport simulations. An option could be the use of detector response functions which mimic the sensitivity of a real neutron detector. However, we are more interested in accounting for the actual thermal and epithermal neutron signals at the detector location in the simplified scenario for deriving a general understanding of the expected neutron intensities under heterogeneous site conditions than in reproducing the signal that our real-world sensor would have measured. As a consequence, an ideal virtual neutron detector is more useful for our objective. Furthermore, due to missing soil moisture information from peatland areas as well as a number of simplifications made in the model which, for instance, include the identical chemical composition of organic peatland soils and mineral soils it is not possible to generate a reasonable fit between simulation and observation. For example, the additional hydrogen and carbon stored in peatland soil need to be considered when trying to simulate real-world data and was not possible with the model version used. Similarly, missing trace elements such as gadolinium and boron in mineral soils will have an influence on the simulated and observed neutron intensities in the thermal domain. Due to these model simplifications we decided to use the energy window option in order to understand the neutron flux at the detector location in the thermal and epithermal energy range and draw more general conclusions for such heterogeneous soil water distributions. We will explain this reasoning in the revised manuscript.

R2-3: Point scale soil moisture observations are very limited and they represent only short distance. This has been pointed as main limitation of the present study but very late in the manuscript and without explaining the consequences of that. Please note that the use of limited number of soil moisture locations have been highly criticized in former studies (see discussion for (Rivera Villarreyes et al., 2013). Despite I'm personally do not against comparisons with relative few points sensors, it should be noted that the present study concluded that the use of thermal signal improved the performance. However, I see a strong bias if we consider that thermal has a smaller footprint and point scale soil moisture used for the comparisons are located in the near field. As such, I would rather presume to have worst results in case the point soil moisture sensors would have been distributed also at larger distance.

A2-3: The number of permanently installed in-situ references soil moisture sensors is limited and an assumption we have to make is that the average soil water content observed by the limited number of reference sensors is valid for the part of the footprint covered with mineral soils. A supporting indicator for this assumption is that the two different sensor types installed show similar absolute average values and soil water dynamics. Deriving an aerial average soil moisture from calibration against several in-situ sensors distributed throughout the footprint was not possible due to the low number of sensors. Thus, the aim of this study was to improve the calibration against reference sensors in the near-field in order to make a step forward in deriving a spatially differentiated soil moisture time series from CRNS. Therefore, we conducted the neutron transport simulations in order to understand the footprint sizes of thermal and epithermal neutrons as well as the influence of soil moisture variations in the simulated neutron intensities of both energy ranges.

The smaller footprint of thermal neutrons as well as the weaker increase of thermal neutrons with decreasing soil moisture can be seen in the model results. Thus, thermal neutrons can be used as a proxy to produce a rescaled neutron count rate which improves the calibration. The applicability of this method may be indeed site-specific and may not produce improved results at other study sites with different soil water distributions and environmental settings. Therefore, we also tested the approach of adjusting the parameters of the standard transfer function as a more transferrable method. Although it might be site-specific, the combination of thermal and epithermal neutrons presented here can illustrate the potential different neutron energies and of having a closer look on thermal neutrons for the estimation of soil water contents in the scope of CRNS.

R2-4: L2: I suggest the term estimations instead of measurements, i.e., the sensors measures neutrons and, based on that, estimate soil moisture.

A2-4: We agree. We will change "measurements" to "estimations" in line 2.

R2-5: I think it should better phrased. 1) soil homogeneous conditions are unlikely and from my understanding 2) the added value of CRNS emerges exactly in case of heterogeneous conditions. The key assumption is in my opinion to sense a representative volume where soil moisture shows a relative short correlation length. In this case neutrons well mix within the footprint. In case of longer correlation length and spatial patterns, empirical data deviates from theoretical functions and hysteresis behaviour could also emerge. Similar consideration has been detected in snow patches conditions (Schattan et al., 2019)

A2-5: We agree with the reviewer in saying that soil homogeneous conditions are unlikely, and we did not claim this in the manuscript. Nevertheless, there is no approach for translating observed neutron intensities into soil moisture yet that explicitly considers sub-scale heterogeneity in the CRNS footprint. Usually it is assumed that soil moisture is homogenous or that its spatial correlation length is smaller than the CRNS footprint. However, given the non-linear relationship between neutron counts and soil moisture, this assumption my not be applicable for calculating CRSN footprint soil moisture. We agree that this should be clarified in the revised manuscript.

Original: “Most approaches and processing techniques for observed neutron intensities are based on the assumption of homogeneous site conditions within the measurement footprint of the neutron detector.”

Adjustment: “Most approaches and processing techniques for observed neutron intensities are based on the assumption of homogeneous site conditions, or of soil moisture patterns with correlation lengths shorter than the measurement footprint of the neutron detector. Nevertheless, in view of the non-linear relationship between neutron intensities and soil moisture it is questionable whether these assumptions are applicable.”

R2-6: L39. If I’m not wrong, some papers refer to the threshold 0.5 for thermal neutrons. Could you provide reasoning for this value?

A2-6: The threshold of 0.5 eV in some previous studies refers to the cutoff-threshold for cadmium as an additional shielding material to reduce the contribution of thermal neutrons to the moderated (epithermal) counter tube of the neutron detector. In this study we did not use a second shielding (cadmium or gadolinium) in addition to the polyethylene shielding of the moderated (epithermal) counter. Due to the absence of this shielding in neutron observations made on-site, we decided to use the physical energy threshold of thermal neutrons in the neutron simulations; i.e. the energy were thermal neutrons are going to be in equilibrium with the surrounding atomic nuclei energetically and neutron absorption (capture) becomes a relevant process.
This will be explained in the revised manuscript.

R2-7: L43. The more complex response of the thermal neutrons rise my main concern of the present study, i.e., the results are very site specific (see general comment above)

A2-7: The more complex response of both energy ranges needs to be indeed considered and requires further and more general investigations to enable further applications of thermal neutrons in CRNS. A combination of thermal and epithermal neutrons leads to an improved estimation of near-field soil moisture at our study site (see our responses above). Although the rescaling approach presented might be rather situation-specific and the combination of thermal and epithermal neutrons presented in this study may not lead to improvements at study sites with different soil moisture conditions and spatial patterns, the second alternative calibration approach does allow improving the calibration against near-field reference measurements without the use of thermal neutrons. Thus, our study is able to present a more general approach for the improved calibration of near-field soil moisture as well as illustrating the potential of thermal neutrons which should be explored in greater detail in future studies. Here, it would be inevitable to consider and investigate the impact of the non-unique behaviour of thermal neutrons in order to estimate more general implications for the use of thermal neutrons in the scope of CRNS. Although this is beyond the scope of the present field study, further research investigating the thermal neutron characteristics in greater detail are underway.

R2-8: L65. I'm very surprised if most studies with stationary CRNS assume homogeneous site conditions. Please rephrased as previously discussed.

A2-8: As done in A2-5, this sentence will be rephrased as well.

Original: "This may be of particular importance as most studies with stationary CRNS assume homogeneous site conditions."

Adjustment: "This may be of particular importance as most studies with stationary CRNS assume quasi-homogeneous site conditions or spatial patterns of different soil moisture states and dynamics with correlation lengths smaller than the CRNS footprint."

R2-9: L105. Please add if possible the 21 random locations on figure 1.

A2-9: We will add the sampling point locations to figure 1.

R2-10: L125. but why to simulate something that it does not represent the real detector? See general comment above

A2-10: Please see to our detailed response A2-2.

R2-11: L131. point scale soil moisture observations are very limited, they represent only short distance and they are not evenly distributed. The CRNS calibration is strongly biased

A2-11: The reference sensors being installed close to the sensor only as well as the missing reference sensors in far-field peatland soils are indeed important limitations of this study. However, because both sensor types installed show similar dynamics we assume that they are representative for the area covered with mineral soils. The limitation of reference soil moisture sensors that are limited to the mineral soil part is discussed in chapter 4.3.

R2-12: L166-183. This text refers to all simulations and not only to simulation set 1. It should be moved up in section 2.2.

A2-12: We agree. We will move lines 166-183 to section 2.2 starting after line 149.

R2-13: L180. As far I understood from previous studies, D86 is not spatially constant. Please specify if you refers here to the maximum (or average) depth over the footprint

A2-13: This is correct. The integration depth decreases with increasing distance to the neutron detector. In our study D_{86} represents the average depth in the footprint per simulation scenario. We will rephrase this in the following way:

Original: "For thermal neutrons, the measurement depth D_{86} is defined as the 86 percent quantile of either the depth of the thermalization point or the maximum depth along the neutron transport path while for epithermal neutrons we use only the latter."

Adjustment: "For thermal neutrons, the average measurement depth D_{86} is defined as the 86 percent quantile of either the depth of the thermalization point or the maximum depth along the neutron transport path while for epithermal neutrons we use only the latter."

R2-14: Figure 2. ground water level should be reported as depth from soil surface to ground water instead of ground water level above sea level to facilitate the interpretation on the discussion on shallow water table influencing soil moisture detected by CRNS.

A2-14: The groundwater levels displayed refer to the height above sea level. As the peatland areas have different elevations (see figure 1) ranging between e.g. between 59 and 61 meters and mineral soils areas being located slightly higher, we will revise the figure with an elevation of 61 m as the reference value for calculating the groundwater table depth variations. In this case, the displayed groundwater table depth time series refers to the approximate depth beneath the peatland soil surface.

R2-15: Equation 6: you merge thermal and epithermal with different footprints. But you compare the scaled sum with point scale soil moisture weighted based on epithermal footprint. Are you not mixing up the signals? Additionally, what about using only thermal? I expect good or even better results when calibrating with these near field point locations.

A2-15: This is correct. We intended to mix the signals to derive a new rescaled signal. This signal is then more similar to the theoretical epithermal intensity occurring if the entire footprint would have the soil moisture conditions of the near-field of the sensor and thus, better matching the shape of the functional relationship of the standard transfer function. One reason is the smaller footprint of thermal neutrons more likely to less influenced by far-field soil moisture variations. Another more important reason for the improvement is the generally smaller decrease of thermal neutrons with increasing soil moisture. As a consequence, we can make use of thermal neutrons as proxy for a different signal response better matching the response which would occur if the soil moisture conditions of the near-field would cover the entire footprint at our site. As described earlier, this approach may not be directly transferable to sites with different spatial patterns of soil moisture and different dynamics. However, the more general approach by adjusting the transfer function instead of adjusting the neutron signal could be used instead.

We did not consider thermal neutrons alone as this would require to derive a new transfer function. This, however, requires extensive modelling efforts and is beyond the scope of this study. Nevertheless, this should be considered in future studies; especially when thermal neutrons applications evolve and the observed thermal neutron intensities need to be explained in detail. This would then require a closer look on soil chemistry as well as it was pointed out by reviewer #1 and #2.

R2-16: Figure 3b. If you calculate the maximum D86, it should be expected to not changing much the depth by increasing far-field soil moisture. Please clarify

A2-16: As clarified in author response A2-13 the measurement depth D_{86} represents the areal average measurement depth. Under dryer conditions in peatland soils, the footprint radius is larger and a stronger influence of peatland soils on D_{86} can be expected.

References

Weimar, J., Köhli, M., Budach, C., and Schmidt, U.: Large-Scale Boron-Lined Neutron Detection Systems as a 3He Alternative for Cosmic Ray Neutron Sensing, Frontiers in Water, 2, 16, https://doi.org/10.3389/frwa.2020.00016, 2020.