Response to Reviewer 1:
We thank Dr. Markus Köhli for the prompt, detailed feedback on our manuscript. We are providing a response to each comment in blue.

Quality:

- The manuscript is written in a straight forward way and can be easily followed. Some sentences might appear to be rather long and complex. The figure quality is good. The references are well organized.

- One shortcoming of this study is that soil moisture was not consistently measured throughout the reference period. Sect. 3.2, citing Woodley et al., 2024, implies that in-situ samples were collected only once. With only considering snow heterogeneity the whole approach of the analysis of different influences on the CRNS signal has a synthetic character. The authors do not hide that fact, but they also do not clearly discuss it.

It would be completely out of scope to integrate soil moisture distributions here as well, yet, the findings from the authors may by such either be increased or smeared out. Soil moisture is most often correlated with the patches of snow, which means if areas with shallow snow pack additionally are more wet, the heterogeneity increases. On the other hand, all relative (SWE) signal variations depend on the underlying soil moisture. This can either decrease or increase the effect of snow. Such should clearly be discussed in order to provide a guide for non-CRNS experts for how to interpret the results.

We agree that in these cases, especially with shallow, heterogeneous snowpacks, soil moisture will impact the CRNS signal. We were able to collect in-situ soil samples once at the end of the winter of 2020/21, but these data provided the spatial distribution of soil moisture only at a single point in time. In addition, soil moisture sensors were installed in nine locations throughout the study area during winter 2020/21, with sensors at 3-4 depths between 0-50 cm in each location. However, these data were not informative because the ground froze to below 0.5 m depth for nearly the entire winter. We highlighted this drawback in Woodley et al. (2024); however, we acknowledge that this should have been made clear in this manuscript as well. In summary, we do not feel that we have sufficient information about the spatiotemporal distribution of soil moisture during the study period to include it in our simulations, so we chose to take the soil moisture estimate we had and apply it to our research domain. We will add discussion of the potential influence of soil moisture spatial heterogeneity to the revised manuscript.

- how is SWE calculated from the LIDAR measurements? Do the authors assume constant snow density?

Snow density measurements from manual snow pits were applied to the lidar snow depth measurements in order to calculate SWE. These snow pits were conducted in the north-south snow drift in the western portion of the study area. Based on the snow pit observations, we adopted a 2-layer density scheme (a lighter snow layer atop a denser, basal snow layer) where the density of each layer was determined from the snow pits and the variation in depth of each layer across the study domain was determined by differencing lidar snow depth maps on different dates.

A detailed process was outlined in Supplementary Information for Woodley et al., 2024, which this manuscript also uses. We will highlight how SWE was calculated by adding a summary of these methods to the revised manuscript and link the appropriate methods and data sources to this manuscript.

- Is simplified weighting function (B1) from the appendix of DOI 10.5194/hess-21-5009-2017 not working for the SWE distribution (it does not have, to, the question is in fact more related to the Woodley publication).

For Woodley et al. (2024), we found that the simplified weighting function (B1) was a decent fit to the distribution of neutrons from our URANOS simulations. However, we were not entirely satisfied with the match to our data; specifically B1 slightly underpredicted neutron counts at short radii (20-60m) relative to our URANOS model output. As a result, we opted to fit our own weighting function to the model output, based on the B1 formulation, and found that our method provided the best fit out of several weighting function types that we tested.

- The comparison analysis is somehow confusing in the way it is structured. Like for soil moisture, you can assume, that there is a horizontal weighting function for SWE. There is no such of a systematic analysis existing in literature, the authors presented one solution in Woodley et al., 2024. The authors compare uniformly assumed SWE with the actual SWE distribution patters (and Eq.(1)). They compare equal weighting with simulations of detectors at 25 virtual locations. The authors find in their analysis that inhomogeneities in the SWE distribution lead to deviations from the assumed equal weighting. That part is discussed quantitatively in detail. They then switch from count rate considerations to spatial representativeness, where they compare a lidar SWE evaluation with an analytically weighted SWE. Here they find that the CRNS weighted SWE is more close to the area average.

Why are those results or evaluations not combined into one discussion? First, the neutron simulation also directly provides the neutron density over the study site. This density can be transformed to a SWE value can easily be compared to other ones. The virtual detectors you only need to describe where neutrons come from, to trace for example the origins individually. For just the count rate the entire detector layer in URANOS provides the values at each pixel in the domain. Secondly, after the probably a little bit too detailed quantitative discussion you repeat the comparison of SWE average vs. heterogeneous by only changing the averaging radius from 171 m to the whole domain. Whereas that is a part to be considered it takes a lot of room, especially as the results are not so different. Thirdly, there is switch from counts and comparison to averages to which points in their domain correspond to the average and therefore are more representative by now including another method which is analytical weighting.

In summary, that part should be condensed and restructured.

We agree that some of our results can be condensed and restructured for greater clarity. We will attempt to do this in the revised version of the manuscript.

With regards to your proposed alternate approach to this analysis, we think that we understand what you are suggesting, but we are not completely certain about all aspects. Your suggestion seems like a good strategy for understanding the spatial distribution of neutron density across the study area, which can then be used to evaluate the spatial representativeness of individual pixels compared to the areal average. Our approach in this analysis was different, but we feel that it is equally effective. The goal of the analysis is to directly compare how a CRNS instrument will behave in the conditions of a synthetic uniform and natural snow distribution. In our analysis, we also wanted to understand how representative a CRNS estimate is of a large region, to inform comparison to remote sensing or gridded data, or provide information to applied CRNS users that can be used to understand the effective representativeness of CRNS of a large area and for siting CRNS instruments. We felt that this approach required the use of virtual CRNS detectors, given that we were uncertain if a direct analogy could be made between counts from the opaque virtual CRNS detector and a transmissive detector layer in URANOS. However, we feel that both of these approaches are viable ways to address the goals of the analysis.

Specific comments:

Comments to the figures:

- Labels A and B in the figures are referred to in the caption as (a) and (b).

Thank you, Prof. Köhli, for catching these mistakes in our figures. We will edit both the figures and the captions to be consistent throughout the manuscript.

- Caption Fig. 6: "surrounding the virtual" -> "surrounding the virtual detector".

Thank you for catching the typo in Fig. 6. This will be edited in the revised manuscript.

- Fig. 7 seems to not add a lot more to what can already be seen in Fig. 6

We somewhat agree that Fig. 7 does not differ significantly from Fig. 6. While we believe that the point of the analysis in Fig. 7 still has some value in demonstrating the representativeness of the CRNS of large areas, we see that the similar results do not add enough to justify another figure. Thus, we will remove Fig. 7 from the manuscript.

- Fig. 8: in a) and b) the dashed line could receive a label in order make more clear what you want to compare. The label "1 m" in the legend is not very descriptive.

We agree that the dashed lines and the labels for Fig. 8 could be confusing. We will add labels to the dashed lines Fig. 8 and change the legend label to “Snow Scale" to hopefully better communicate the analogy represented in the simulation.

- Fig. 9: convert the axis to normal dates. DOWY is used nowhere else in the manuscript.

We will edit the x-axis to normal dates in the revised manuscript.

- Fig. 9: The "CRNS timeseries" does not match anything which is plotted in the reference Woodley et al., 2024.

The “CRNS timeseries” corresponds to the daily averaged SWE values from the actual CRNS counts shown in Fig. 5a of Woodley et al. (2024). The axes in our Fig. 9 differ slightly from the Woodley et al. (2024) figure. The y-axis of Fig. 9 is in mm instead of cm, and our x-axis in Fig. 9 includes a shorter time span.

- Fig. 9: The authors include the "CRNS timeseries" - do they mean SWE derived from CRNS?

For Fig. 9, we wanted to provide context to what the CRNS instrument installed at the CARC measured and compare them to the gridded products. The “CRNS timeseries” corresponds to the daily averaged SWE values from the actual CRNS counts measured by a CRNS instrument at our research site from Woodley et al. (2024). We will edit the figure and description to include more clarity regarding what the CRNS time series is.

Title:

- the reviewer thinks that the title "Influence of Snow Spatial Variability on Cosmic Ray Neutron SWE" overstates the results as the title is too general, although the research topic presented is a very broad basis for the SWE analysis. It should include that it primarily conceptualizes SWE analysis as derived from their site.

We agree that our title is overly broad and general. We propose to change our title to the following: Influence of Snow Spatial Variability on Cosmic Ray Neutron SWE in Northern Prairies

Comments to the text body:
Introduction
- The overview is very well organized and includes the relevant literature. The different parts are, however, somehow disconnected. It begins with agricultural land-use changes, moves into dryland cropping techniques and then shifts to snow without clear linkage.

Thank you, Prof. Köhli, for this comment on the text. We will edit the introduction and attempt to make better connections between the several aspects of the introduction in the revised manuscript.

- l72: Hydrogen 'trap' (the technical term would be 'absorb') free neutrons if neutrons are thermalized. The CRNS signal attenuation is mainly due to slowing down of neutrons by hydrogen, i.e. losing energy through elastic collisions.

Thank you for correcting the language on our CRNS explanation. We will make the corrections to our revised manuscript to better explain the CRNS signal attenuation.

Data and Methods

- The description of the NASA SnowEx field campaign at CARC is informative, but the connection between different measurement techniques (InSAR, UAV lidar, SfM, snow pits, etc.) is not well worked out given the specificities of each. Considering as the comparison of CRNS SWE with various gridded datasets (UCLA-re, SNODAS, and UA) is in general an important validation step and also undertaken in this study.
Results and Discussion

Thank you for the comments. We wanted to explain what was conducted during the NASA SnowEx campaign, but didn't make clear connections on how this field campaign intersected with this analysis. We will clarify this section and try to highlight the various gridded datasets in our revised manuscript.

- l198: "N_theta is the calibration neutron count, from the "snow-off" reference date of 15 January 2021." That is of course not a real calibration, it is an approximation for the calibration, which is associated with significant systematic uncertainties. One would need to for example look at the entire CRNS timeseries in order to try to understand (or justify) this choice. Even if you think, that this is a 'good enough' choice this needs to be discussed very clearly, as later on in the text you make very detailed numerical comparisons of which some directly relate to the choice of this value.

We agree that our choice of N_theta is an approximation of a calibration.  We made this decision because it was not possible for us to collect calibration soil samples prior to winter snow accumulation due to delayed delivery of the CRNS instrument. We included this discussion in Woodley et al., 2024, but we will add discussion in this manuscript as well.

- l263: How significant are the relative differences mentioned here and shown in Fig. 5? The reviewer assumes, for example based on Fig. 6, that the number of neutrons counted rather low given the small relative differences. So that analysis might not be precise enough for sub-percent statements (i.e. 3.16%)?

You are correct that the relative differences are rather low, and the degree of specificity in the percentiles is a bit overstated. We will round the percentages to the nearest percent in the revised manuscript.

- l288: As stated above, the overestimation might simply be a result of not using any weighting function, but there might other effects playing a role, too.

Respectfully, Prof. Köhli, we are somewhat uncertain what you mean in the comment. Specifically, we are not sure which weighting function the reviewer is referring to. The results discussed in this section were modeled outputs, where we directly compared model runs with a uniform versus heterogeneous snow layer. No spatial weighting function was applied to the model outputs.  With everything else being equal in the model simulations, these “overestimations” are likely from the differences in the snow layers.  We think that the “overestimation” may result from the fact that a large portion of the study area had low or zero snow accumulation, making it more likely that the virtual CRNS detectors would experience higher count rates in the heterogeneous snow simulation due to a large percentage of the neutrons originating close to the instrument. Additionally, in this manuscript (and Woodley et al., 2024) we have not applied a spatial weighting function to the model outputs

If we misinterpreted or did not adequately respond to your comment, we would welcome clarification.

- l353: you mention 624 simulations but not how you arrive at that number, given that, the reader is left wondering what that means.

Thank you, Prof. Köhli, for pointing out this omission. We will include a breakdown in the methods of how we arrived at the 624 number (26 points x 3 snow layer scenarios x 8 dates) in the revised manuscript.

Conclusion:
- l487: "a naively sited CRNS instrument (i.e., with no knowledge of the snow distribution) is still 2 to 5 times more likely to be representative of the large-scale average SWE than a more conventional (...)" - that statement is not found in the discussion before.

Respectfully, this statement was included in the end of Section 4.3 (l408). We will make these results more visible in the text, since this is a major takeaway from this analysis that we want to highlight.

Technical comments:
Typography:
- Equations are part of the text body. Therefore, they follow the interpunctuation. That means a dot after Eq. (1), not before, and a comma after (2) and (3).

- Tables: Try using the same number of decimals and align numbers either left or, preferably, right.

Thank you, Prof. Köhli, for the technical comments. We will make the appropriate edits to the punctuation around the equations and the table.

References:
- This manuscript refers to for the (previously already published) data. However, the DOI which is provided in Woodley et al., 2024 for the CRNS SWE data (https://doi.org/10.5067/NJR0AMMOHZ4E) does not work.

Thank you for pointing out the DOI is broken. This dataset is still in the process of publication with NSIDC since the Woodley et al., 2024 paper. The data will be available at this link once the process is complete.