

## **Reviewer #1:**

We thank reviewer #1 for taking the time to review our manuscript. Following the suggestions of reviewer #1, we agree that our manuscript needs major revisions before it may be published. We are certain that the reviewer's comments will help to streamline our manuscript, to make it clearer and more relevant for the reader of SOIL. In the following section we will reply to all comments of reviewer #1 with R1-1 (i.e. reviewer 1, comment 1) and A1-1 (i.e. author response to R1-1), respectively.

**R1-1:** The soil moisture network (SMN) represents a valuable data-sets (L131-132). Still, it is explained that could not represent CRNS footprint well (L348). Considering in addition that the deeper soil moisture network (> 40 cm) is based on a lower number of point-scale sensors, it could be questionable its value as benchmark for testing the performance of the depth-extrapolated CRNS-SM. For this reason, it cannot be concluded in my opinion if the difficulties to estimated deeper SM are based on CRNS settings, SMAR settings or an unrepresentativeness of the SMN. To circumvent this issue, I strongly recommend the Authors first to test SMAR settings only to the SMN, i.e., using first layer of SM measured by the SMN to estimate the deeper SM. Then, to apply the best settings of the SMAR model to CRNS. This exercise could be performed to each single soil moisture profile and any combinations. The results should provide a benchmark for testing the depth-extrapolated SM based on CRNS. Please note that you have already quantified some differences between CRNS and SMN. Thus, it should be expected that the depth extrapolated CRNS-SM cannot perform better than that. Finally, it is not clear to me if the depth extrapolated CRNS is compared with the weighted SMN or to the arithmetic average SMN. It should be the former to be consistent with the signal but with the disadvantage to have again weighted soil moisture. This issue seems to not be resolvable and could limit some applications. Discussion should integrate this aspect.

20

**A1-1:** We fully agree that a decreasing number is reference in-situ point-scale soil moisture sensors is likely to reduce the representativeness of the SMN in these depths. This indeed adds to the uncertainties coming with the surface soil moisture time series derived from CRNS as well as the SMAR model and the defined soil physical parameters used in SMAR. Different potential reasons for the poor performance are already given in the manuscript but they are difficult to disentangle. We agree that further analyses are necessary and agree on testing the original and modified version of SMAR on the SMN only prior using it with CRNS-derived surface soil moisture time series. For this reason, we will conduct the following additional analyses and add them to the revised manuscript:

1) We perform a comparison of the original and modified SMAR using the soil moisture in 0-20 cm depth for two individual soil moisture profiles which comprise sensors in all depths (i.e. 10, 20, 30, 50, 70, 130, 200, 300 and 450 cm). The SMAR-derived second layer soil moisture time series is then compared to the soil moisture information in the respective depths of the two sensor profiles.

2) We perform a comparison of the original and modified SMAR using the average soil moisture in 0-20 cm depth from the entire SMN. The SMAR-derived second layer soil moisture time series is then compared to the average soil moisture information in the respective depths of the SMN.

5

These additional analyses are performed on daily time steps and with the literature-based soil physical parameters shown in Tab. 1. This means that for the original SMAR, only the  $V_2$  water loss term is calibrated while the modified SMAR is not calibrated at all. This will give us insights on the performance of the original and modified SMAR model against our reference SMN, e.g. how the SMAR model performs on profile/plot scale and if the use on the entire SMN already provides results with lower accuracy. This will provide an additional benchmark for the evaluation of the results which is more robust than the arbitrary threshold of  $0.06 \text{ cm}^3/\text{cm}^3$  which has been used in previous studies to evaluate the SMAR performance.

10

The CRNS-derived soil moisture time series always represents a weighted soil moisture average because by the nature of the geophysical method, it is more sensitive to soil moisture changes close to the neutron detector. When the performance (goodness of fit) of the CRNS-derived surface soil moisture time series is assessed, it should always show a better performance when the reference average soil moisture derived from e.g. the SMN is weighted according to the sensitivity of CRNS. Consequently, the assessment of the different neutron-to-soil moisture transfer functions is done accordingly and our analyses underline the improved performance when the reference soil moisture from the SMN is weighted.

15

However, this indeed means that a (CRNS-derived) weighted field-scale soil moisture average is used as the input for the SMAR model while we compare the modelled second-layer soil moisture time series with arithmetic averages of the SMN in the second layer. We agree that this point should be added to the discussion and add a sentence in a revised version of the manuscript.

20

**R1-2:** First, I found misleading the arguments to support the use of SMAR. The Authors argue several times that “most approaches require a site-specific calibration using depth-profiles of in-situ soil moisture data, which are often not available.” (L6-7). For this reason, you use SMAR. However, this approach also requires calibration. So your arguments do not support your decision. Similarly, it is stated (L9) that “SMAR is usually also calibrated to sensor data, but could be applied without calibration if all its parameters were known” and L90. “This method does not require calibration if the environmental parameters are known.” Well this is valid for any models and not only for SMAR. The problem is that it is often the case that we do not know the parameters, and especially at the scale of model application when soil is under investigation. The results of this study also confirm this statement. So your arguments do not support your decision to use SMAR as many other approaches have similar settings. Overall, I’m fine with the decision to test SMAR (L86). But I’m against the argument that this is physically based and it does not require adjustment in comparison to other approaches. The results of the present study

25

30

also confirm my doubt. Comparison to other approaches, e.g., the exponential filter (Franz et al., 2020; Wagner et al., 1999) could strengthen this study.

Finally, the Authors tested several SMAR settings. Despite the overall exercises are remarkable, I'm wondering why testing approaches that already sound not appropriate. I.e., I found the assumptions behind the original SMAR model very arguable.

- 5 E.g., a constant flux  $V_2$  (P and ET) is already debated when applied for estimating monthly groundwater recharge and it should be considered unrealistic over a long time series (years) for hourly to daily resolutions. The results show how  $V_2$  moves from, e.g., 50 to 300 mm h<sup>-1</sup> in Table 3. These results are not consistent and support my doubt of using standard approach. So I'm not surprised if the Authors tried to modify the approach and I would remove the scenario with the constant  $V_2$ . Similarly, the CRNS effective depth seems to be easily estimated (L279). So why testing a guess of 35 cm? I would remove this scenario.
- 10 Overfitting a model calibrating all the parameters also seems to me a crash test and it does not seem to me a good approach for testing if a model is reliable. Several parameters are not in a physically consistent range and what was argued to be a physical approach has been destroyed. I would also remove this scenario.

- A1-2:** We agree that our formulations may have been partly misleading and inconsistent. The original SMAR model introduced  
15 by Manfreda et al. (2014) comprises different parameters which have a direct physical meaning such as the porosity, soil moisture at wilting point and field capacity. The  $V_2$  parameter also has a physical meaning as it represents a water loss term comprising evapotranspiration and percolation losses from the second soil layer. Given the physical meaning of all parameters in the SMAR model, they are easier to be estimated compared to bulk fitting or calibration parameters. For this reason, we stated that the SMAR model allows for the estimation of soil moisture in a second, deeper soil layer without the necessity of  
20 in-situ reference information for calibration. Yet, we acknowledge that despite its physical meaning, the  $V_2$  parameter is difficult estimate and is usually calibrated.

- The reviewer mentions the soil water index which is often referred to as the exponential filter (Wagner et al. 1999). This very simple approach has been applied in numerous studies and only includes a single, bulk fitting parameter, namely the characteristic time length  $T$ . With its calibration against in-situ reference measurements, many processes are tried to be  
25 accounted for. Estimating  $T$  for a specific study site, e.g. from parameters such as soil depth, soil physical parameters proved difficult (e.g. Zhang et al. 2017, Wang et al. 2017). This makes the use of the exponential difficult when no in-situ reference soil moisture measurements for calibrating  $T$  are available.

- However, we agree to the reviewer that the additional testing of the exponential filter approach would strengthen the outcome of this study. Consequently, we will add the depth-extrapolation with the exponential filter method to our analyses, repeat all  
30 scenarios also with the exponential filter method and add these scenarios to the revised manuscript.

As the reviewer points out, when the  $V_2$  parameter is calibrated against in-situ reference measurements, its absolute calibrated value increases as the second layer thickness increases. This is related to the nature of the SMAR model being a version of a storage model. For example, when the field capacity of the second layer is exceeded by 10 cm<sup>3</sup>/cm<sup>3</sup> or 10 vol.-%, the entire 10

vol.-% percolate from the second layer to larger depths. Expressed in mm, this value is higher for a second soil layer with a thickness of 1000 mm compared to 500 mm. This explains higher values for  $V_2$  in units of mm when the thickness of the second soil layer increases.

Nevertheless, we do agree that the constant  $V_2$  is problematic but in our opinion keeping this scenario in the manuscript is important to compare our introduced modified SMAR (estimating a dynamic  $V_2$ ) with the original SMAR (calibrating a constant  $V_2$ ).

The reviewer also argues that some scenarios (first layer depth of 35 cm) and fitting all soil physical parameters should be removed. We are thankful about this comment as we think that this will help to streamline the story of our manuscript, to make it more digestible for the reader and to highlight the key points of this study. In a revised version of the manuscript, we will therefore remove all scenarios where soil physical parameters are calibrated (except  $V_2$ ). We will also remove the scenarios for hourly time steps and only keep scenarios with daily time steps as this is the temporal resolution for which both, the SMAR model and the exponential filter have been designed for.

15

**R1-3:** L3-5. deeper than what? I would rephrase otherwise it is not clear to what you refer when saying “deeper”, e.g., moving the sentence “Many of these applications require information on soil water dynamics in deeper soil layers.” After “Cosmic-ray neutron sensing (CRNS) allows for non-invasive monitoring of field-scale soil moisture across several hectares around the instrument but only for the first few tens of centimeters of the soil.”

20

**A1-3:** We will rearrange the sentences accordingly:

Original: “Many of these applications require information on soil water dynamics in deeper soil layers. Cosmic-ray neutron sensing (CRNS) allows for non-invasive monitoring of field-scale soil moisture across several hectares around the instrument but only for the first few tens of centimeters of the soil.”

25

Modification: “Cosmic-ray neutron sensing (CRNS) allows for non-invasive monitoring of field-scale soil moisture across several hectares around the instrument but only for the first few tens of centimeters of the soil. Many of these applications require information on soil water dynamics in deeper soil layers.”

30

**R1-4:** L8. Personally, I’m not too much on the philosophical discussion about process-based or empirical approaches (e.g., (Hrachowitz and Clark, 2017)). But to call a bucket model with several empirical assumptions a physical-based approach could be highly criticized. Please consider rewording.

**A1-4:** We will rephrase the sentence in the following way:

5 Original: “The physically-based soil moisture analytical relationship SMAR is usually also calibrated to sensor data, but could be applied without calibration if all its parameters were known. However, in particular its water loss parameter is difficult to estimate.”

10 Modification: “The soil moisture analytical relationship SMAR is usually also calibrated to sensor data, but due to the physical meaning of each model parameter, it could be applied without calibration if all its parameters were known. However, in particular its water loss parameter is difficult to estimate.”

**R1-5:** L36. Despite, I agree that establishing an extensive point sensor network requires a lot of effort, the international soil moisture network ISMN (Dorigo et al., 2021) provides a good example that it is worth citing.

15

**A1-5:** We will add this reference to our manuscript:

20 Original: “As a consequence, extensive point sensor networks which allow for the estimation of field-scale soil moisture are often restricted to a rather small number of research related monitoring sites such as the Terrestrial Environmental Observatories (TERENO, [www.tereno.net](http://www.tereno.net)) in Germany (e.g., Zacharias et al., 2011; Bogenia et al., 2018; Kiese et al., 2018; Heinrich et al., 2018).”

25 Modification: “As a consequence, extensive point sensor networks which allow for the estimation of field-scale soil moisture are often restricted to a rather small number of research related monitoring sites such as the Terrestrial Environmental Observatories (TERENO, [www.tereno.net](http://www.tereno.net)) in Germany (e.g., Zacharias et al., 2011; Bogenia et al., 2018; Kiese et al., 2018; Heinrich et al., 2018) or the International Soil Moisture Monitoring Network (ISMN, Dorigo et al. 2021) covering sites around the globe.”

30 **R1-6:** “L80 and L82. To my knowledge, (Wagner et al., 1999) proposed the use of an exponential filter approach to estimate soil water content at deeper soil layers based on surface soil moisture. Instead the term SWI refers to a quantity of water content between 0-100%. Thus, I would not call the extrapolation approach “soil water index” but rather “exponential filter approach”.

**A1-6:** We will exchange the term “soil water index” with “exponential filter approach” throughout a revised version of this manuscript.

5 **R1-7:** L128. How field capacity and wilting point have been estimated is missing and should be added.

**A1-7:** This information is already given in the caption of Tab. 1: “[...]. The soil moisture content at field capacity and wilting point were taken from tabulated values in Sponagel et al. (2005) according to the respective soil grain size class (medium-fine sand) and the soil bulk density of the individual layers.”

10

**R1-8:** L134. The use of the manufacture’s calibration function could be another source of errors that should be discussed in the result section.

15 **A1-8:** This is correct. Adding the additional analyses and the new results requested by reviewer #1 in R1-1 and R1-2 will lead to severe changes in content and structure of the manuscript. We will add this point to an appropriate part of the results and discussion section of a revised manuscript.

20 **R1-9:** L154. The fact that you use a 25h moving average could be a reason to not be able to capture fast soil moisture changes? This could be discussed in the result section.

**A1-9:** The smoothing of the corrected CRNS neutron signal prior transferring it soil moisture may have some effect on capturing fast soil moisture changes. Although we think that this effect is not responsible for the poor performance of SMAR,

25 we will add a sentence on potential effects in the discussion of a revised version of our manuscript.

**R1-10:** Eq.5. I guess h is air humidity. Please clarify.

30 **A1-10:** This is correct. We will add this information accordingly:

Original: “The universal transport solution (UTS), eq. (5) – eq. (6), (Köhli et al., 2021) accounts for the changing relationship between neutrons and soil moisture under different conditions of absolute humidity.”

Modification: “The universal transport solution (UTS), eq. (5) – eq. (6), (Köhli et al., 2021) accounts for the changing relationship between neutrons and soil moisture under different conditions of absolute humidity  $h$  in  $\text{g m}^{-3}$ .”

5 **R1-11:** L179. Please clarify how look-up-table approach works. What are also the assumptions/approximations in using this approach?

**A1-11:** The look-up table approach is necessary because the UTS equation cannot be simply rearranged to be solved for soil moisture. Instead, one approach is to calculate the neutron intensity from the UTS for each soil moisture value in a range from 0.0001 to 0.5  $\text{cm}^3/\text{cm}^3$  in 0.0001  $\text{cm}^3/\text{cm}^3$  steps and compare the derived neutron intensity with the observed neutron intensity at each time step. The soil moisture value which produces the smallest absolute difference between the predicted neutron intensity and observed neutron intensity is then used for each time step. A second approach is also possible. We will add this information:

15 Original: “Soil moisture can be derived from the UTS using numerical inversion or a look-up table approach which is used in this study.”

Modification: “Soil moisture can be derived from the UTS using numerical inversion or a look-up table approach which is used in this study. In the look-up table approach, soil moisture values in the range from 0.0001 to 0.5  $\text{cm}^3 \text{ cm}^{-3}$  in steps of 0.0001  $\text{cm}^3 \text{ cm}^{-3}$  are used to predict the neutron intensity using the UTS for each time step. For each time step, the soil moisture value producing the smallest absolute difference between the observed and predicted neutron intensity is then assigned as the CRNS-derived soil moisture value.”

25 **R1-12:** L181. So how  $N_D$  has been actually calculated? Is it related to  $N_0$ ? Please clarify.

**A1-12:**  $N_D$  has been calculated just like the look-up table approach described in the response to the previous comment. For the hours of the calibration campaign in February 2019, soil moisture from independent soil samples and absolute humidity from local temperature and relative humidity measurements were used to predict a neutron intensity with the UTS for a range of  $N_D$  values. The  $N_D$  value with the smallest RMSE between observed and predicted neutron intensities was then used. The  $N_D$  value needs to be calibrated for each version of the UTS (MNCP thl, MCNP drf, URANOS thl, URANOS drf).

30  $N_0$  is derived equally for the standard transfer function. However, the physical meaning of the two parameters differs and a direct, unique relationship between the calibration parameters does not exist as the absolute humidity term in the UTS changes this relationship. We will add the following sentence to line 201:

Original: “Reference soil moisture information from the soil sampling campaign in February 2019 was weighted accordingly and used for calibrating both transfer functions.”

5 Modification: “Reference soil moisture information from the soil sampling campaign in February 2019 was weighted accordingly and used for calibrating both transfer functions.  $N_0$  and  $N_D$  were iteratively calibrated. For  $N_0$ , the value producing the smallest RMSE between soil moisture from soil samples and the one predicted with eq. 1-4 was chosen. For  $N_D$  soil moisture information derived from soil samples for the hours of the sampling campaign were used to predict neutron intensities with eq. 5-6. The  $N_D$  producing the smallest RMSE between predicted and observed neutron intensities was chosen.”

10

**R1-13**: L187. What are the differences between URANOS type and MCNP type equation?

15 **A1-23**: The UTS is based on neutron transport simulations conducted with two different neutron transport model codes: URANOS and MCNP. As expected two different models produce slightly different results and hence, different fitting parameters for  $p_1$ - $p_{10}$  have been obtained. We tested all available parameter sets for the UTS to determine which should be used and our results are in line with the initial findings of Köhli et al. (2021) who introduced the UTS.

20 **R1-14**: L211. Are you sure you have to move back and forward from volumetric soil moisture and relative saturation? Could you not try directly expanding eq. 7 directly for volumetric soil moisture?

25 **A1-14**: In case of the SMAR model we have to stick to the procedures and units defined. This would be different for the exponential filter method which can be applied to soil moisture data in units of relative saturation and  $\text{cm}^3/\text{cm}^3$  (Bouaziz et al. 2020). However, the calibrated characteristic time length parameter  $T$  is likely to be different for a soil moisture time series that have been previously transformed to relative saturation and one in units of  $\text{cm}^3/\text{cm}^3$ . This should be considered when comparing different studies reporting calibrated values for  $T$ .

30 **R1-15**: L248. I guess the estimation would be very sensitive to the length of the time series. Did you use the entire time series for estimating  $ET$ ? Please clarify.

**A1-15**: The average  $ET_2$  for the entire time series determined based on eq. 12-13 would indeed be dependent on the length of the surface soil moisture time series. However, we are not determining an average  $ET_2$  and  $P_2$  for the entire time series and



then use the subsequently derived time series average  $V_2$  to run the SMAR model for deriving soil moisture in the second layer.

Instead, as it can be seen in eq. 13, the  $ET_2$  value is calculated and used individually for each time step and so it also done for  $P_2$ . This also means that each time step of in the modified SMAR model receives its individual  $ET_2$  and  $P_2$  value and hence, also and individual varying  $V_2$  water loss. We will add the following modification:

Original: “Here, we estimate the amount of evapotranspiration from the deeper layer  $ET_2$  based on the difference between the current and past value of relative saturation of the first layer, by scaling the value to the dimension (i.e. extent) of second layer and by considering the difference in cumulative root fraction between both layers, assuming that root water uptake for ET is larger in the layer with more roots eq. (13).”

Modification: “Here, we estimate the individual amount of evapotranspiration from the deeper layer  $ET_2$  for each time step based on the difference between the current and past value of relative saturation of the first layer, by scaling the value to the dimension (i.e. extent) of second layer and by considering the difference in cumulative root fraction between both layers, assuming that root water uptake for ET is larger in the layer with more roots eq. (13).”

**R1-16:** L490. Well, the discovery that root water uptake varies with time and depth depending on the water availability goes probably back to the introduction of irrigation practices in the history. I guess Maysonnave et al.,2022 showed more than that. If the Authors remove the scenario with  $V_2$  constant, this comment could also be probably removed. Otherwise please rephrase.

**A1-16:** As stated in A1-1, we would like to keep the results for the original SMAR version in this manuscript to compare it with our modified SMAR version and, as requested by reviewer #1, also with the exponential filter in a revised version of our manuscript. We interpret R1-16 in the way that the reference of Maysonnave et al. (2022) is not appropriate here and that information on depth and water availability related root-water uptake is well established. Consequently, we remove this sentence in a revised version of our manuscript.

**R1-17:** The Authors have stated that this study is the first who evaluates the UTS (L515). If this is the case I think is worth clarifying this from the beginning and provide some additional information to better understand how it works and is implemented.

**A1-17:** Yes, to our knowledge this is the first study apart from Köhli et al. (2021) who introduced the UTS which compares the standard transfer function after Desilets et al. (2010) and UTS in terms of their performance in deriving soil moisture from

CRNS. More information on the calibration of ND and the look-up table approach will also be added (see A1-11, A1-12, A1-13). We will replace the following statement in the introduction:

5 Original: “In addition, we apply different neutron-to-soil moisture transfer functions available to derive the surface soil moisture time series. This is done to assess which transfer function performs best and if a better CRNS-derived surface soil moisture time series translates into better estimates of the depth-extrapolated soil moisture.”

10 Modification: “Different approaches exist to derive soil moisture from observed neutron signals. The standard approach after Desilets et al. (2010) is commonly used to derive soil moisture from CRNS but has been found insufficient especially at observation sites with low soil moisture contents. New approaches include the interdependence of the relationship between neutrons and soil moisture (Köhli et al. 2021) and report an improved estimation of soil moisture with CRNS. Therefore, this study aims for a first independent assessment of performance of the standard and the revised approach after Köhli et al. (2021) as well as if a better CRNS-derived surface soil moisture time series translates into better estimates of the depth-extrapolated soil moisture.”

15

#### **Additional references:**

20 Bouaziz, L. J. E., Steele-Dunne, S. C., Schellekens, J., Weerts, A. H., Stam, J., Sprokkereef, E., Winsemius, H. H. C., Savenije, H. H. G., and Hrachowitz, M.: Improved Understanding of the Link Between Catchment-Scale Vegetation Accessible Storage and Satellite-Derived Soil Water Index, *Water Resources Research*, 56, e2019WR026365, <https://doi.org/10.1029/2019WR026365>, 2020.

25 Dorigo, W., Himmelbauer, I., Aberer, D., Schremmer, L., Petrakovic, I., Zappa, L., Preimesberger, W., Xaver, A., Annor, F., Ardö, J., Baldocchi, D., Bitelli, M., Blöschl, G., Bogena, H., Brocca, L., Calvet, J.-C., Camarero, J.J., Capello, G., Choi, M., Cosh, M.C., van de Giesen, N., Hajdu, I., Ikonen, J., Jensen, K.H., Kanniah, K.D., de Kat, I., Kirchengast, G., Kumar Rai, P., Kyrouac, J., Larson, K., Liu, S., Loew, A., Moghaddam, M., Martínez Fernández, J., Mattar Bader, C., Morbidelli, R., Musial, J.P., Osenga, E., Palecki, M.A., Pellarin, T., Petropoulos, G.P., Pfeil, I., Powers, J., Robock, A., Rüdiger, C., Rummel, U., Strobel, M., Su, Z., Sullivan, R., Tagesson, T., Varlagin, A., Vreugdenhil, M., Walker, J., Wen, J., Wenger, F., Wigneron, J.P.,  
30 Woods, M., Yang, K., Zeng, Y., Zhang, X., Zreda, M., Dietrich, S., Gruber, A., van Oevelen, P., Wagner, W., Scipal, K., Drusch, M., and Sabia, R.: The International Soil Moisture Network: serving Earth system science for over a decade. *Hydrology and Earth System Sciences* 25, 5749–5804. <https://doi.org/10.5194/hess-25-5749-2021>, 2021.

Wang, T., Franz, T. E., You, J., Shulski, M. D., and Ray, C.: Evaluating controls of soil properties and climatic conditions on the use of an exponential filter for converting near surface to root zone soil moisture contents, *Journal of Hydrology*, 548, 683-696, <http://dx.doi.org/10.1016/j.jhydrol.2017.03.055>, 2017.