Responses to reviewer's comments

We thank the editor and the reviewers for their time and effort to review our manuscript, which helped to further increase the quality of the paper. All comments have been addressed carefully.

Below, reviewer comments are marked in red.
Responses to the comments are marked in blue.
Cited changes that have been made in the manuscript are marked in *italic*.

Reviewer 1:

The MS titled, 'Uncertanity estimation for a new exponential filter-based long-term root-zone soil moisture dataset from C3S surface observations' by Pasik et al. describes a methodology to deduce rootzone SM using only satellite derived surface soil moisture using a well-known exponential filter approach with its error characterization. In general, I found this study lacking in novetly factor. The exponential model has been around for a while now and is shown to have limited success in estimating SM at deeper layers (>40cm) as also shown in this study. Furthermore, there are some glaring gaps in background infromation, method descriptions etc. Therefore, despite tackling a critical issue in sub-surface hydrology, I would recommend rejection in its current form. However, I would be more than happy to see it re-submitted with significant revisions.

We regret to learn that the novelty in the uncertainty estimation scheme for the exponential filter (EF) is not clear. Our intention is for the uncertainty estimation method (rather than the dataset or the EF itself) to be the main focus of the paper. To our best knowledge, this study is the first one to provide uncertainty estimates for a global, EF-based soil moisture dataset which also accounts for uncertainties of the EF method itself. Moreover, while the application of the EF approach at greater soil depths yields less accurate results, we believe that there can still be merit found in such estimates provided that the uncertainties are known (which, again, are the main focus of this study). We will revise the manuscript to make this more clear, and address the gaps in background information and method descriptions mentioned in the major and specific comments below.

Major Comments	Reply
1) As mentioned earlier, the exponential filters have been studied extensively (as also acknowledged in the MS) in the past. Also has well docmented issues such as poor performance in the deeper layers, struggles with widely different soil types between surface and lower layers (no mention or discussion around this in the MS), summer season decoupling etc. This study acknowledges and re-affirms most of these issues but presents no path forward in trying to solve them. Its does not seem to be taking the scinece forward either by reducing the model known limitations, or sheding new lights on model performance with its global implementation or discussion. Therefore, in my opition, in the current form, this study lack the novelty factor and may need significant revisions.	Thank you for pointing out these previously unacknowledged limitations of the EF. We will mention them while discussing the limitations of the method: *"Other limitations of the method include generally poorer performance in arid zones and when soil texture is not homogeneous throughout the soil column (Yang et al., 2022; Ford et al. 2014)"*. In this MS, we focus on advancing the understanding of the EF method by contributing our uncertainty estimation scheme to describe, rather than reduce, the model's known limitations. We will make this intention clear by adding the following sentence to conclude section **3.1 Exponential filter**: *"It is precisely such limitations that we attempt to describe with the uncertainty estimation scheme developed in this study, and hence advance the understanding of the EF method's performance"*.
2) I agree with the authors that their product could be the covering the longest period of record for observation-based RZSM, but shoud have at least acknowledged other global products such as SMAP and SMOS L4 products in the MS. In fact, SMOS L4 product is similar to this study, they are also using EF (https://sextant.ifremer.fr/record/316e77af-cb72-4312-96a3-3011cc5068d4/) whereas SMAP L4 uses data assimialtion approach to merge SMAP with the Catchment Land surface Model with detailed uncertainty analysis and published ATBDs.	Thank you for the suggestion. We will explicitly reference SMAP L4 RZSM while discussing various approaches to estimating RZSM in the **Introduction**: *"Satellite-based SSM observations can also be assimilated into a land surface model to produce estimates of RZSM with global coverage, as in the case of the SMAP L4 RZSM product (Reichle et al., 2017)."* We will also edit the concluding sentences of the introduction to acknowledge other EF-based datasets (incl. SMOS L4 RZSM): *"While other EF-based datasets exist (e.g., the SMOS L4 product), they offer limited spatio-temporal coverage and lack quantitative uncertainty information (Al Bitar and Mahmoodi, 2020; Bauer-Marschallinger et al., 2018)."*

	In addition to the discussion on the root-zone estimation techniques (and the EF method in their context) already present in L44-60 of the submitted MS, we will also include specific references to SMOS L4 and SMAP L4 RZSM products as per the previous comment. Furthermore, in the context of product performance, we will add the following as the concluding sentence of section **4.2 Global RZSM product quality assessment**: *"The performance of our product is similar to that of other satellite-based RZSM products found in other studies, especially when considering the same regions for assessment (Xu et al. 2021, Reichle et al. 2017). While the data set presented here does not outperform other existing RZSM products, it distinguishes itself as the only purely observation-based global product covering such a long time period, and the only EF-based product that has uncertainty estimates provided with it"*.
3) Furthertmore, I think the MS could be improved further with some more background information on other rootzone SM estimation techniques, currently the MS does not talk about other methods and why EF might be better than others, in that context an best average correlation of 0.56 doesnt not inspire too much confidence. The authors simply skipped over some of the detailes of satellite SM estiamtions like difference in bands (X&C for AMSR-E vs L for SMAP and SMAOS etc.), I think it would improve MS.	The input C3S SM dataset is a harmonized product where the biases between bands are mitigated by the inter-calibration of the sensors. Furthermore, each estimate is a weighted combination of the individual sensor observations available on that day. However, the spectral band of the sensor, via its suitability for retrieving soil moisture information (e.g., different sensitivity to precipitation and evaporation), impacts the uncertainty estimates. We will make this more clear when introducing the C3S dataset in **section 2.1** by including the following: *"Note that the distinctive life spans and spectral bands of the used satellite missions (e.g., C and X-bands used by AMSR-E, L-band used by SMOS and SMAP) can potentially also lead to distinctive changes in the data quality of the merged product via the differences in their sensitivity to precipitation or evaporation. These sudden changes in SSM and uncertainty data are hereinafter referred to as systemic breaks (Preimesberger et al., 2021). Although said breaks have a marginal impact on the SSM signal itself due to the inter-calibration of sensors, they are distinct in the uncertainty estimates. As more and newer sensors provide better retrievals, mean uncertainty values typically decrease distinctively with every new satellite launch in more recent periods (Gruber et al., 2017)"*.

	It is correctly pointed out that the reference in situ sites used for optimizing the T-parameter do not equally represent the variety of land cover classes. Similarly, their geographical distribution is largely skewed towards the Global North (as acknowledged in L27-29 of the submitted MS). However, as other studies have shown, optimizing the T-parameter per, e.g., soil type, does not yield an improvement versus using an averaged T-value (e.g., De Lange et al. 2008, Grillakis et al. 2021). This is why we choose to lump all the available in situ stations together in the optimization of the T-parameter. Even though their distribution might be skewed toward particular land cover classes (i.e., grasslands or croplands) the impact of this is probably negligible (e.g., Stefan et al. 2021). This limited sensitivity to variation in T-parameter values is discussed in L135-141 of the submitted MS.
4) If I understand correctly, most of the in-situ sites are in open fields (usually near agricultural land). Therefore, Topt obtained may only be able to represent (presumabely) a particular landcover type. Is there any analysis authors have performed to assess the model perfromance at other locations? While reading the MS, I could not figure out if the model was implemented at gridded scale or only at the ISMN sites. Perhaps, this could be made more clearer.	Indeed we evaluate our product globally at grid cell level against ERA5-Land data. We will include spatial correlation maps for each of the product layers in an additional new Figure (please find it included at the end of this document), where the performance of the RZSM product can be seen at every single location. Performance patterns observed in these maps do not show correlation with the distribution of the ISMN stations used for the T-parameter optimization (i.e., Fig. A1 in the submitted MS), which too suggests that the impact of their uneven distribution and over-representation of certain land cover classes does not have a substantial impact. The EF model was optimized with point-scale ground measurements and implemented globally at the grid scale. We will add the following text to make this clearer throughout the MS: 1) in the opening section of the **Results**: *"In this section, we first show results of the point-scale T-parameter optimization. Next, we compare the gridded RZSM product globally to E5L."* 2) in the opening sentence of section **4.2 Global RZSM product quality assessment**: *"A global SM dataset spanning the 2002–2020 period was computed using the EF method and T-parameters optimized at the point scale with the approach described in section 4.1."*

5) Finally, I think the discussoin could be further improved (especially if study is simply focused in implementing at larger scale with its limitations intact) by talking about if there is any regional pattern in model performance (arid vs humid conditions); tropical vs sub-tropical region? Does soil types play any role in Topt and uncertanity estimations? What is the dominent landcover type and can these different rooting systems (barren soils vs cropland vs deep rooted trees) explain some of the issues beign faced?

Thank you for this suggestion, we will include spatial correlation maps in a new figure (mentioned in our response to comment #4) and discuss the observed patterns in section **4.2 Global RZSM product quality assessment**: "*A global SM dataset spanning the 2002–2020 period was computed using the EF method and T-parameters optimized at the point scale with the approach described in section 4.1. Figures Xa–e) show correlation maps of each of the RZSM product layers as well as the input C3S SSM dataset with E5L. The spatial patterns observed in the C3S SSM data (Figure Xa) are strikingly similar to those in RZSM layer 1 (Figure Xb) with slight to moderate deterioration in performance over the high latitudes (*$> 60°N$*). This is not surprising given that both products differ only by a small degree of smoothing applied to RZSM layer 1 and are compared to the same E5L layer (0-7 cm). RZSM layers 2 and 3 (Figure Xc-d) are compared to E5L layers 7-28 and 28-100 cm, respectively, and largely preserve good performance in regions where the input C3S SSM product also performs well, i.e., in Europe (bar Scandinavia), the Caspian and Aral Sea basins, the Eastern United States, India, Southeast Asia, South America, Sub-Saharan Africa, and Australia. At the same time, deterioration of performance is observed in high latitudes and in arid environments such as the Sahara desert and the Arabian Peninsula where the reduced strength of coupling between the surface and root-zone dynamics may hinder the EF performance (Yang et al., 2022). The patterns of good and poor performance visible in RZSM layers 1-3, are not replicated in RZSM layer 4 (Figure Xe) where the agreement with the reference E5L 100-289 cm layer is spatially very heterogeneous and worse overall. The few regions where the good performance observed in shallower layers is preserved include India, Southeast Asia, and the Eastern United States*".

The impact of soil type or land cover on the performance of the EF remains ambiguous (e,g., Stefan et al. 2021). External variables could potentially influence the uncertainty estimates, but such detailed analysis is outside of the scope of this study. We will reiterate the focus of this study throughout the MS, e.g., as per major comment #1 and also by adding the following concluding sentences to the MS' Introduction: "*The focus and novelty of this paper lie in quantifying, rather than reducing, the EF model's known limitations by providing a methodology for comprehensive uncertainty estimation for the EF method. Additionally, to our best knowledge, this dataset is, as yet, the longest available solely observation-based, error-*

Specific Comments	Reply
1) Figure 1, there seems to be huge overlaps between Topt for different layers (25th-75th percentile box), how would this impact the results? Have the authors consider perhaps running the model woth those as upper and lowes limits on Topt to see the impact on performance?	The variability in T_{opt} likely reflects the differences in environmental conditions and sensor depths between calibration sites, resulting in large overlaps in T_{opt} IQR between product layers. Product comparison results presented in Figure 3 reaffirm the limited sensitivity of the EF to variations in T-parameter observed by others (Ford et al. 2014; Grillakis et al. 2021). This is especially apparent in case of E5L layer 7–28 cm, where the performance of all four RZSM product layers is very similar. Nonetheless, each of the RZSM product layers correlates best with its most-approximate E5L counterpart in all but one case.
2) Typically EF is implemented in a normaliezed SM scale (SWI either by scaling from 0-1 using min/max or using soil characteristic properties). In the MS, it's not mentioend which expecific method was used (if at all).	The scaling of the input SM data between 0–1 and subsequent rescaling between the wilting point and field capacity values is usually done when dealing with datasets that express SM as the percentage of saturation rather than in volumetric units (m^3/m^3), such as in the EUMETSAT H SAF soil moisture data records. Given that the C3S SM data is already in volumetric units, there is no need for rescaling.
3) Lines 45-48, could use more detailes on existing methods for rootzzone SM estiamtions and their challenges.	We will revise these lines to include more detail: *"The existing link between SM dynamics in the surface layer and the root zone (Albergel et al., 2008; Wang et al., 2017; Ford et al., 2014; Sure and Dikshit, 2019) allows for estimating RZSM from surface SM (SSM) observations via a variety of hydrological models. These include relatively simple two-layer approaches approximating RZSM as a function of SSM (Manfreda et al., 2014), compound process-based models requiring sophisticated parameter calibration (Bouaziz et al., 2020), as well as complex and computationally expensive land surface models requiring many auxiliary inputs (Muñoz Sabater et al., 2021; Rodell et al., 2004). Satellite-based SSM observations can also be assimilated into land surface models to improve the model simulations of RZSM with global and temporally-complete coverage, as in the case of the SMAP L4 RZSM product (Reichle et al., 2017)."*

4)Line 69, EF typically is used to estimate SM at specific layer depth not a composites like 0-10 or 0-40. It's either 0 or 10 or 40 cm ($\pm$ few cms)	Both approaches are common practice. For example, in the aforementioned SMOS L4 product, the EF is used to represent a 5–40 cm soil layer (Al Bitar and Mahmoodi, 2020). Other studies have used 0–100 cm (Wagner et al. 1999, De Lange et al. 2008); 0–25, 50–100, and 0–100 cm (Ceballos et al. 2005); or 25–60 cm (Ford et al. 2014). Moreover, we believe that, given the large variations in land cover and soil texture within single satellite grid cells, it is impossible to assign a single specific depth to the EF-based RZSM estimates. Other practical considerations were also taken into account here, e.g., binning of the very limited number of in situ sensors operating at greater depths is necessary to obtain a reasonable sample against which to calibrate the T-parameter. Therefore, we think that a depth range accompanied by rigorous uncertainty estimates (which are the main goal of the study) provide a more realistic description of the product.
5) Section 2.1, I would suggest to include a table showsing the timeline of various satellite SM products being part of C3S with band information that would help understanding the dataset better. Also, I would liked to see some examples of mentioned structural breaks either as timeseries.	Individual sensor data are intercalibrated within the input C3S SM dataset and the breaks in the *SSM* signal occurring at sensor changes were demonstrated by Preimesberger et al. (2021) to be marginal. However, the structural breaks—in the new manuscript referred to as systemic breaks, see above—are distinct in the C3S product *uncertainties*. We will include a sensor timeline table in section **2.1 C3S surface soil moisture**. We will also clarify this in the MS (see our response to major comment #3). A time series example of a systemic break in C3S is shown in Figure 6. We will redesign Figure 6 to accommodate specific comment #9 and also to make the systemic breaks in C3S uncertainty time series clearly visible (revised Figure 6 is included at the end of this document).

6) Figure 3, I dont think there is any need to compare all the layers at each depth. For instance, deeper layers should not be compared with 0-7 cm modeled SM. They are not the same thing to be comapared and does not add any value to the MS. Similarly, at 100-200 cm depth, surface SM correlations and their discussions (Lines 241-250) could be avoided.	Regarding Figure 3, the reason for comparing all of the RZSM product layers as well as the input C3S SSM layer against E5L is to demonstrate that our approach to T-optimization works and that the sensitivity of the EF method to variations in T_{opt} is limited, as discussed in the MS earlier. The former is reflected in the best performance of each RZSM layer being achieved again at the depth of the E5L product we intended it to; the latter is apparent in the very similar performance of all RZSM product layers against E5L reference (regardless of significant differences in T_{opt} between them). For these reasons we believe this figure provides more information in its current form and we would like to keep it. We will also annotate the median lines in every box for an easier comparison of the results (revised Figure 3 is included at the end of this document).
7) Line 259, the structural uncertainty mentioned here is it same as the structural breaks (Line 87)? if not, perhaps use some other terminology to differentiate is further.	Model structural uncertainty refers to an error inherent in the EF method and is introduced and explained much earlier in L64-66 of the submitted MS (although in first instance referred to as model structural *error*, which we will change also to model structural *uncertainty* for consistency). Subsections 3.2.3 (in methods) and 4.3 (in results) discussing the model structural uncertainty were both given the same name to make it clear what is being referred to. Structural breaks are sudden changes in the input uncertainty data and are described when introducing C3S dataset (L87 of the submitted MS). We will use the term *systemic breaks* instead to better distinguish the two terms. Further clarification regarding the difference between the systemic breaks in the SSM signal and those in the uncertainty data will be provided as per our response to major comment #3.
8) Line 268, please add exact location of the data.	Thank you for spotting the missing location. We will revise the caption to: *Figure 5a shows a time series of RZSM uncertainties from the baseline method at an arbitrary example location in Benin (9.875N, 1,625E).*
9) Figure 6, very busy plot. Hard to read.	Thank you for the suggestion. We will redesign Figure 6 for more clarity by removing the not relevant in situ signal and placing uncertainties in separate panes with independent scales to make their temporal dynamics better visible (find the revised version at the end of this document).

10) Figure 7, how is the sample size in the subset bigger than the whole dataset (Fig d vs h).	Thank you for pointing this out. The time series selected for this analysis were filtered for a minimum Pearson's r ($r \geq 0.5$) to mitigate the impact of the spatial mismatch between point-scale in situ measurements and the large footprint of the satellite observations (as described and referenced in L162-164 of the submitted MS). In this experiment, several time series that did not satisfy this minimum correlation criteria in the 2002-2020 period (7d) reached the $r \geq 0.5$ threshold in 2015-2020 (7h). This is due to the latter part of the time series being based on more modern satellite sensors providing more accurate SM retrievals. We will reprocess Figure 7 using the exact same sample, i.e., only sites where $r \geq 0.5$ in both periods. We do not expect this to impact the overall result in a significant way, though.

References:

Al Bitar, A. and Mahmoodi, A.: Algorithm Theoretical Basis Document (ATBD) for the SMOS Level 4 Root Zone Soil Moisture (Version v30_01), Tech. rep., https://doi.org/doi.org/10.5281/zenodo.4298572, 2020

Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Piguet, B., and Martin, E.: From nearsurface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations, Hydrology and Earth System Sciences, 12, 1323–1337, https://doi.org/10.5194/hess-12-1323-2008, 2008.

Bauer-Marschallinger, B., Paulik, C., Hochstöger, S., Mistelbauer, T., Modanesi, S., Ciabatta, L., Massari, C., Brocca, L., and Wagner, W.: Soil Moisture from Fusion of Scatterometer and SAR: Closing the Scale Gap with Temporal Filtering, Remote Sensing, 10, https://doi.org/10.3390/rs10071030, 2018.

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, e2019WR026 365, https://doi.org/10.1029/2019WR026365, 2020.

Ceballos, A., Scipal, K., Wagner, W., and Martínez-Fernández, J.: Validation of ERS scatterometer-derived soil moisture data in the central part of the Duero Basin, Spain, Hydrological Processes, 19, 1549–1566, https://doi.org/10.1002/hyp.5585, 2005.

de Lange, R., Beck, R., van de Giesen, N., Friesen, J., de Wit, A., and Wagner, W.: Scatterometer-Derived Soil Moisture Calibrated for Soil Texture With a One-Dimensional Water-Flow Model, IEEE Transactions on Geoscience and Remote Sensing, 46, 4041–4049, https://doi.org/10.1109/TGRS.2008.2000796, 2008.

Ford, T. W., Harris, E., and Quiring, S. M.: Estimating root zone soil moisture using near-surface observations from SMOS, Hydrology and Earth System Sciences, 18, 139–154, https://doi.org/10.5194/hess-18-139-2014, 2014.

Grillakis, M. G., Koutroulis, A. G., Alexakis, D. D., Polykretis, C., and Daliakopoulos, I. N.: Regionalizing Root-Zone Soil Moisture Estimates From ESA CCI Soil Water Index Using Machine Learning and Information on Soil, Vegetation, and Climate, Water Resources Research, 57, e2020WR029 249, https://doi.org/10.1029/2020WR029249, e2020WR029249 2020WR029249, 2021.

Gruber, A., Dorigo, W., Crow, W., and Wagner, W.: Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals, IEEE Transactions on Geoscience and Remote Sensing, PP, 1–13, https://doi.org/10.1109/TGRS.2017.2734070, 2017.

Manfreda, S., Brocca, L., Moramarco, T., Melone, F., and Sheffield, J.: A physically based approach for the estimation of root-zone soil moisture from surface measurements, Hydrology and Earth System Sciences, 18, 1199–1212, https://doi.org/10.5194/hess-18-1199-2014, 2014.

Muñoz Sabater, J.: ERA5-Land hourly data from 1950 to 1980. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.e2161bac, 2021.

Preimesberger, W., Scanlon, T., Su, C.-H., Gruber, A., and Dorigo, W.: Homogenization of Structural Breaks in the Global ESA CCI Soil Moisture Multisatellite Climate Data Record, IEEE Transactions on Geoscience and Remote Sensing, 59, 2845–2862, https://doi.org/10.1109/TGRS.2020.3012896, 2021.

Reichle, R., DeLannoy, G., Koster, R. D., Crow, W. T., and Kimball, J.: SMAP L4 9 km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data, Version 3, Tech. rep., TU Vienna, https://doi.org/10.5067/B59DT1D5UMB4, 2017.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, Bulletin of the American Meteorological Society, 85, 381 – 394, https://doi.org/10.1175/BAMS-85-3-381, 2004

Stefan, V.-G., Indrio, G., Escorihuela, M.-J., Quintana-Seguí, P., and Villar, J. M.: High-Resolution SMAP-Derived Root-Zone Soil Moisture Using an Exponential Filter Model Calibrated per Land Cover Type, Remote Sensing, 13, https://doi.org/10.3390/rs13061112, 2021.

Sure, A. and Dikshit, O.: Estimation of root zone soil moisture using passive microwave remote sensing: A case study for rice and wheat crops for three states in the Indo-Gangetic basin, Journal of Environmental Management, 234, 75–89, https://doi.org/10.1016/j.jenvman.2018.12.109, 2019.

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, https://doi.org/10.1016/j.jhydrol.2017.03.055, 2017.

Lei Xu, Nengcheng Chen, Xiang Zhang, Hamid Moradkhani, Chong Zhang, Chuli Hu: In-situ and triple-collocation based evaluations of eight global root zone soil moisture products, Remote Sensing of Environment, 254, 0034-4257, https://doi.org/10.1016/j.rse.2020.112248, 2021.

Yang, Y., Bao, Z., Wu, H., Wang, G., Liu, C., Wang, J., Zhang, J. An Exponential Filter Model-Based Root-Zone Soil Moisture Estimation Methodology from Multiple Datasets. Remote Sens. 2022, 14, 1785. https://doi.org/10.3390/rs14081785

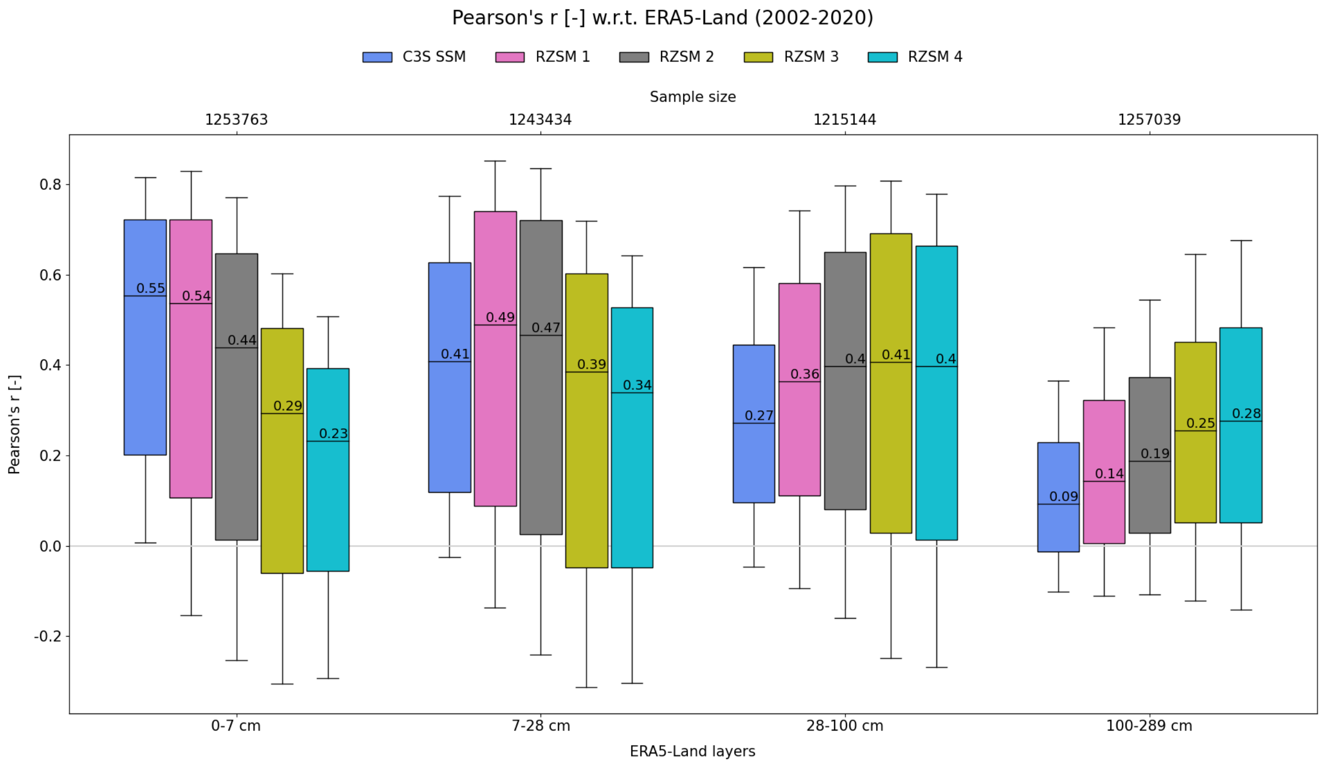
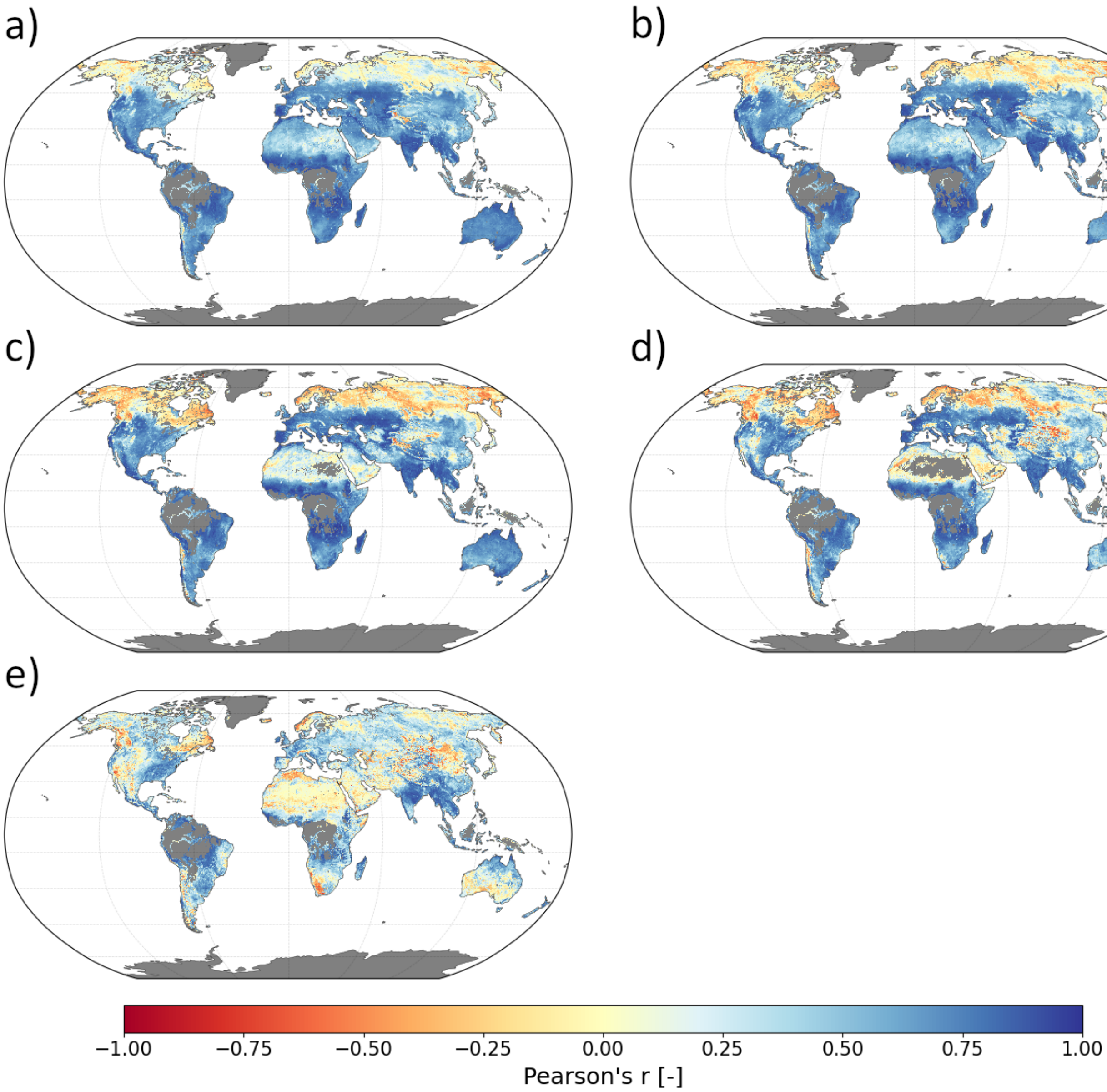


Figure 3: Product intercomparison of the C3S SSM and RZSM products against E5L SM.



Spatial correlation maps of the C3S SSM (a) and RZSM products (b-e) with E5L layer 0–7 cm (a-b), 7–28 cm (c), 28–100 cm (d) and 100–289 cm (e).

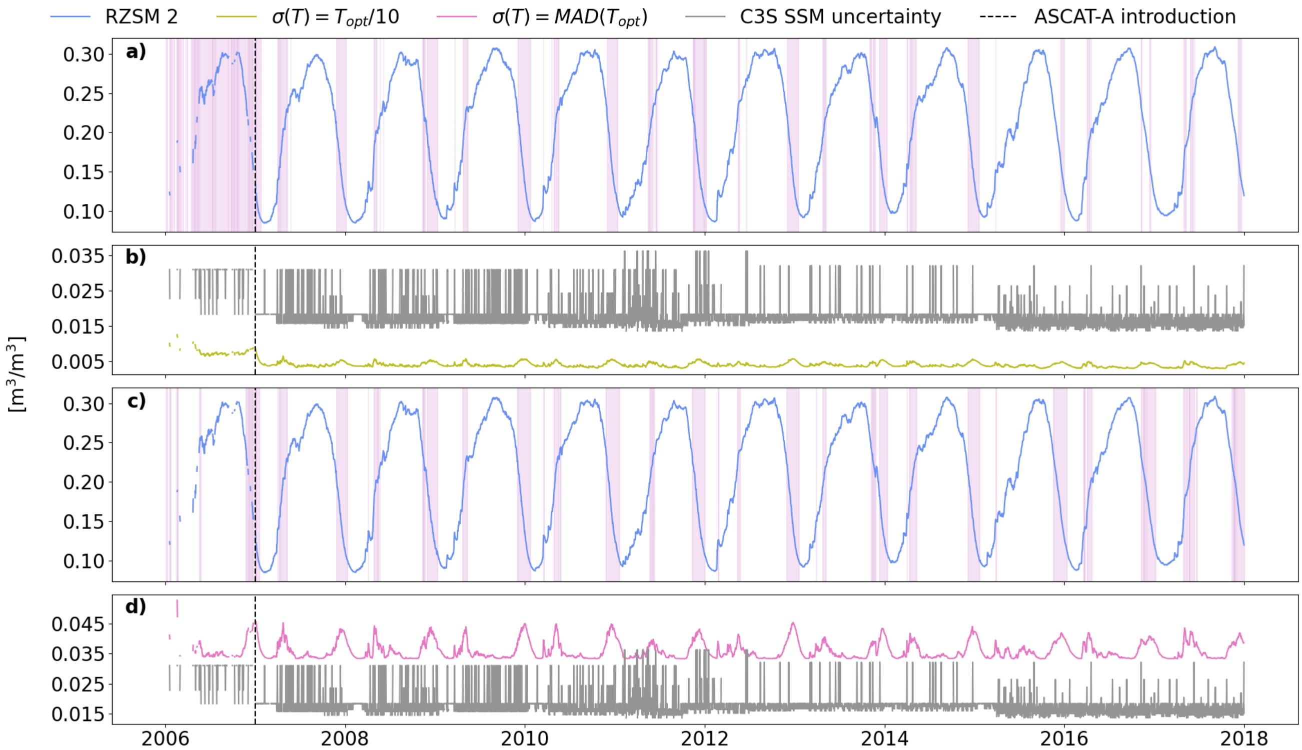


Figure 6: Differences in uncertainty variations of the baseline (a-b) and our proposed uncertainty estimation approach (c-d). Illustrated on the example of RZSM layer 2 at an arbitrary location in Benin (9.875N, 1,625E).