We wish to thank the reviewer for the helpful comments. Please find our replies below in red.

The authors mapped the open water body in the PPR using SAR images and high-resolution DTM data. Time series of water body maps were obtained and the dynamics of open water bodies were analyzed. Generally, this manuscript is well prepared with good structures and figures. However, I have some major concerns about the current work, as:

1) The validation (or accuracy evaluation) of the method. Currently, the authors evaluate the accuracy from a viewpoint of classification, using user accuracy and produce accuracy for the pixels selected. However, as the authors analyzed in section 3.2, what the hydrologist cares about is the number and area of the water body. So, the authors are suggested to validate (evaluate) their method by comparing the area and number estimated from SAR with those from NAIP.

We agree that this would be an interesting information. However, a classification of NAIP mosaics, such as the ones shown in Fig. A1, using either a supervised or unsupervised approach into water bodies and non-water areas is also prone to errors and would require an accuracy assessment on its own. Manual delineation would likely not find all the different water bodies in the mosaic. Hence, we opted for a "traditional" point sample-based accuracy assessment for the manuscript. In response to your comment, we classified one of the NAIP mosaics using a supervised approach. You can find the methods and results below illustrating an underestimation of the number and total area by Sentinel-1. However, it also shows the challenges in connection with creating such a reference dataset from optical airborne data, e.g. false positives in areas with low vegetation cover or tree shadows. Such a small analysis cannot properly address these issues and would warrant a study of its own. We therefore think that this validation is beyond the scope of our study.

Classification of NAIP mosaic and comparison with Sentinel-1 water bodies:

We classified the 2019 NAIP mosaic (shown in Fig. A1c) using a Random Forest (RF) classifier that was trained using 350 of the 400 reference points used in the original accuracy assessment. The remaining 50 points were set aside for model testing. In addition to the four NAIP bands, also NDVI was used as a feature. The classified raster contained many false positives in areas with little vegetation, such as recently ploughed fields and tarmac, as well as areas shadowed by tall vegetation and buildings (Figure R1). For this reason, we intersected the RF-derived areas with high water probability ($p > 0.8$) with the original, LiDAR-derived, potholes and retained those objects as seeds. These seeds were used for region growing with a relaxed probability threshold ($p > 0.5$). For comparison with the coarser-scale Sentinel-1 product, the NAIP product was upsampled to 10 m, whereby only pixels with at least 50% water coverage were retained.

The RF model overall accuracy was estimated at 96% using the test dataset of 50 points. The classification based on NAIP yielded a number of 449 water bodies covering an area of 1967 ha. In contrast, the Sentinel-1 wetland product over the same sub-region contained 243 water bodies with a total area of 1634 ha. However, in the size class > 1 ha, Sentinel-1 had 183 water bodies, while the NAIP-based result contained 185. This demonstrates that especially smaller water-filled potholes remained undetected in the Sentinel-1 product. This seems plausible due to the higher proportion of mixed pixels along the edges as well as shallower water depths which may lead to vegetation protruding through the water surface. The large amount of small erroneous water pixels in Figure R1 suggests, however, that at least some of the NAIP-based small water bodies are false positives, even after the post-processing steps described above.

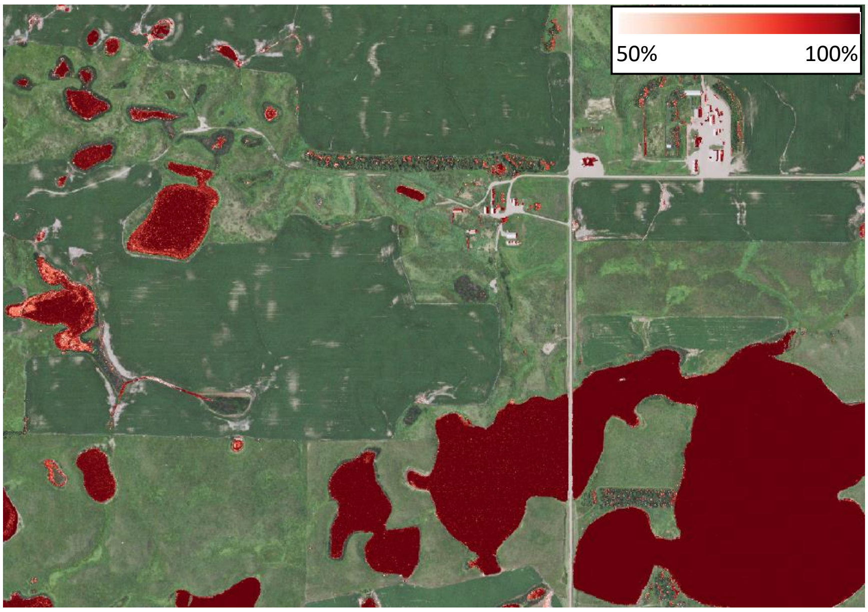

Figure R1: Subset of random forest water probabilities > 0.5 for the NAIP mosaic of 2019.

2)L332-345, the authors explained the different performances of different polarization combinations. However, this work can be done more physically by introducing the Radar functions. Actually, the authors can start from the radar functions and then build their algorithm on the basis of microwave radiative transfer theory.

Parameters can be retrieved from Earth Observation (EO) data using inversion of radiative transfer models or, as in this case, using statistical approaches. In this case, the use of radiative transfer (RT) modelling would have been very difficult to realise given the large diversity of the background (i.e. non-water) classes along with a high temporal variability of backscatter within each class (after all, we analyse six years of data). This includes different vegetation types and stages, different crop types on agricultural areas, as well as vastly different soil moisture conditions, which all influence backscatter. Furthermore, as mentioned in the manuscript, wind is an important factor on the temporal variability within the water class itself. The parameterisation of RT models would have been extremely challenging and would likely not have yielded robust results as were obtained here. Moreover, a large amount of ancillary data would have been required (e.g., accurate land cover, crop type, vegetation growth stage, soil moisture status, roughness) which are either not available or have their own, often considerable, uncertainties.

3) the abstract is too long, please shorten it.

We agree. We propose to shorten the abstract from ca. 480 to 380 words:

"The North American Prairie Pothole Region (PPR) represents a large system of wetlands with great importance for biodiversity, water storage and flood management. Knowledge of seasonal and inter-annual surface water dynamics in the PPR is important for understanding the functionality of these wetland ecosystems and the changing degree of hydrologic connectivity between them. Optical sensors widely used for retrieving such information are often limited by their temporal resolution and cloud cover, especially in the case of flood events. Synthetic aperture radar (SAR) sensors can potentially overcome such limitations. However, water extent retrieval from SAR data is often impacted by environmental factors, such as wind on water surfaces. Hence, for reliably monitoring water extent over longer time periods robust retrieval methods are required.

The aim of this study was to develop a robust approach for classifying open water extent in the PPR and to analyse the obtained time series covering the entire available Sentinel-1 observation period from 2015 to 2020 in the hydrometeorological context. Open water in prairie potholes was classified by fusing dual-polarised Sentinel-1 data and high-resolution topographical information using a Bayesian framework. The approach was tested for a study area in North Dakota. The resulting surface water maps were validated using high-resolution airborne optical imagery. For the observation period, the total water area, the number of water bodies and the median area per water body were computed. The validation of the retrieved water maps yielded producer's accuracies between 84 % and 95 % for calm days and between 74 % and 88 % on windy days. User's accuracies were above 98 % in all cases, indicating a very low occurrence of false positives due to the constraints introduced by topographical information.

The observed dynamics of total water area displayed both intra-annual and inter-annual patterns. In addition to differences in seasonality between small (< 1 ha) and large (> 1 ha) water bodies due to the effect of evaporation during summer, these size classes also responded differently to an extremely wet period from 2019 to 2020 in terms of increase in number and total area covered. The results demonstrate the potential of Sentinel-1 data for high-resolution monitoring of prairie wetlands. Limitations exist related to wind inhibiting correct water extent retrieval and due to the rather long acquisition interval of 12 days over the PPR."

4) some figures, like figure 5, are suggested to add scales and the compass. Since not all readers are familiar with UTM(zone 14)

We will add scale bar and a north arrow. We will also spell out UTM in the caption (requested by another reviewer).

5) The authors are suggested to pay attention to the usage of abbreviations, for example, please define GRASS and dual-pol.

Thank you for pointing this out. Since we only use the term "dual-pol" twice, we will change it to "dual-polarised". GRASS GIS is a widely used GIS package and commonly referred to as GRASS GIS whereas the full name "Geographic Resources Analysis Support System" is much less known. The full name is given in the Reference with a link with further information (Line 534).