Comment 1: "Firstly, the classification results from the WIP Tool appear to be far from perfect, as the overlay on the Esri satellite basemap shows that the wetland area is significantly overestimated, and the commission error seems high. There are many instances of tree shadows and roads being classified as wetlands."

Thank you for your comment. I humbly disagree that the wetland area is significantly overestimated. However, I can understand your skepticism. Forested wetlands that do not have standing water, but rather saturated soils, are difficult to detect in imagery, especially in evergreen forested areas of the Pacific Northwest US that do not lose their leaves, so it may appear that there is a high error of commission and make it difficult for the reviewer. Even when on the ground these cryptic wetlands can be difficult to find (Figure 1).

Here we provide an example of an area that is correctly mapped by the WIP to highlight the difficulty in identifying wetlands by spectral imagery alone (Figure 2). This example demonstrates how the NWI misses large areas of wetlands – some are easily detected because of the yellow stressed-out vegetation from treed bogs (red arrows), but others are impossible to detect in the imagery alone (yellow arrows). In our region these are often referred to as cryptic wetlands. The areas that appear as tree shadows are in fact small hummocky wetlands. We have included a photo from the field to help demonstrate the challenging nature of this landscape (Figure 1).

To test the assertion that there are many instances of tree shadows and roads being classified as wetlands we re-ran the model without any spectral or tree height data in our input datasets. The results (Figure2) demonstrate that spectral imagery was not an important variable in our model as the model results are similar when spectral imagery was not used (Figure 2, lower right) and agree with our hierarchy of variable importance published in our manuscript. The most important input variables are our unique multi-scale terrain indices created as part of the WIP tool and described in detail in our methods.

Figure 1: Field photo of a forested wetland in the PNW. These cryptic wetlands are difficult to detect in aerial imagery. They are statured long enough in the growing season to support wetland vegetation species (skunk cabbage – center), develop hydric soils, and have saturation well into the summer drought months. These

cryptic wetlands provide critical ecosystem services by providing drought refugia, storing large amounts of carbon, and supporting unique species, but they are currently missing from most inventories.

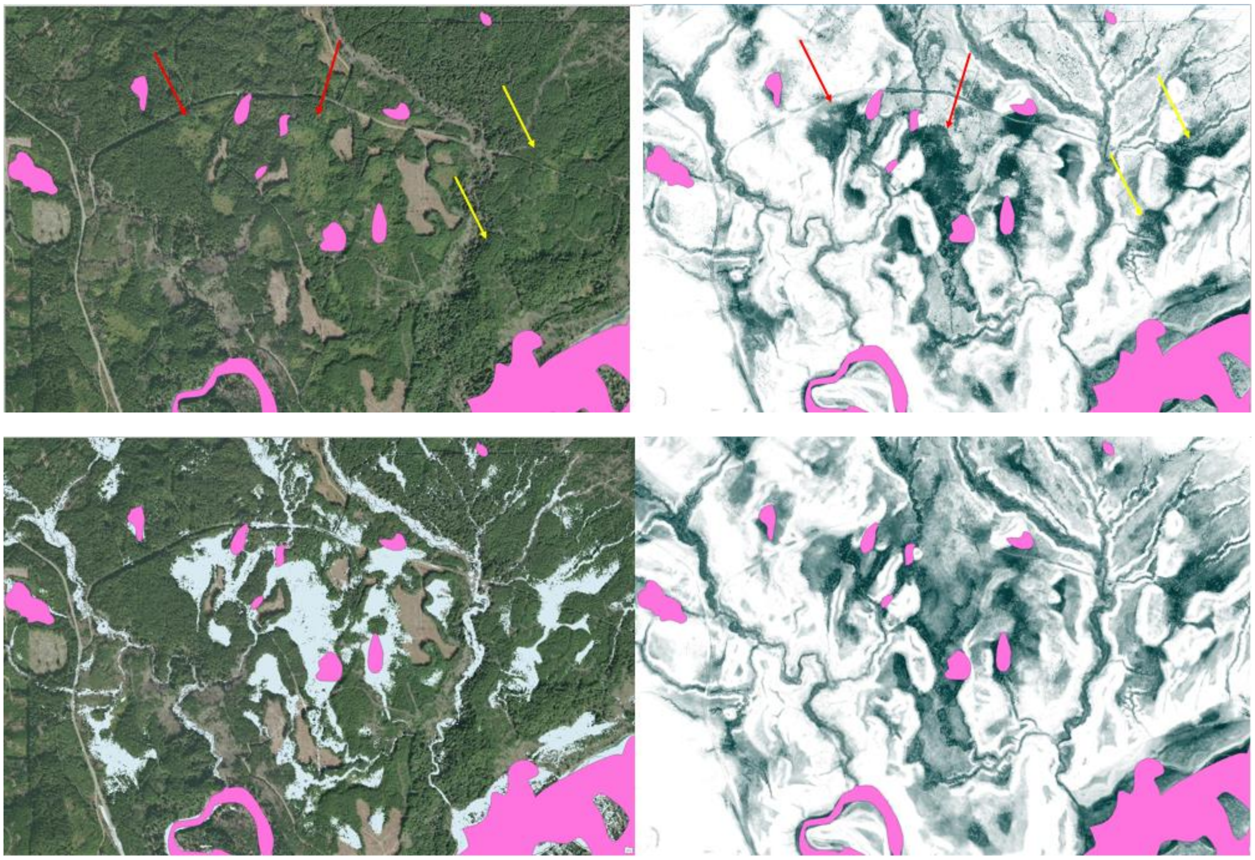


Figure 1: Example of a forested wetland area in the Hoh watershed. The image in the upper left shows areas missed by the NWI (pink) pointed out by red and yellow arrows. Red arrows represent wetlands easily detected in the imagery due to the stressed evergreen vegetation. The yellow arrows represent forested wetlands that are difficult to detect in the imagery due to dense canopy but validated on the ground. The image on the upper right shows the output of the WIP probability. It would be difficult to improve upon this map. The WIP picks up all the forested wetlands that are missed. The bottom left shows the binary classification using a threshold of 0.5, which represents a correct estimate of wetlands in the area. The image on the bottom right is the WIP tool re-run without any spectral or vegetation height data used. While removing the spectral imagery reduces the visible sign on roads, we felt that the error of commission from roads was small and decided to keep all input layers in our model as it did improve overall accuracy. I can find no evidence of shadows in either of these outputs. If the shadow effect had been large, it would have been removed once the spectral imagery had been removed from the model. This provides further proof that these wetlands are small wet areas and not caused by tree shadows.

Comment 2: "Also, the results are quite noisy with numerous tiny and irregular shaped polygons. I think the tool still has a long way to go before it can become a practical dataset complement to NWI."

We can understand how you may be disappointed if expecting a dataset with a similar look and feel to the NWI. We have added a statement to make it explicitly clear that this was not our goal and in no way do we recommend our WIP output as a replacement to the NWI. Rather the WIP output offers a different paradigm to wetland identification by providing a raster-based product. Many end users prefer the WIP probability output for wetland identification especially for areas that do not have clear borders as it highlights the gradient they see on the ground and also provides model uncertainty information. Additionally, a raster based gradient of wetland probability can be used for landscape modelling in ways that a vector based dataset cannot, especially useful for Bayesian probability estimates of wetland ecosystem services such as above and below ground carbon stocks (Hudak et al. 2019[1], Moskal et al. 2023[2]) However, there are many users that prefer the look and feel of the NWI and are using this tool as a screening tool in addition to manual photo interpretation to update the NWI, which is still the standard method within the U.S. However, to your point there are many further steps that could be taken to smooth and present a binary classification such as applying a focal smoothing filter. However, applying such a filter may arbitrarily alter the model results in other ways not related to the model inputs. Because presentation and wetland delineation was not the goal of our research, we did not focus on smoothing or clean up. We simply selected a 0.5 threshold to assess accuracy as one cannot assess accuracy for a probability gradient.

We have observed in the field that many of the small tiny and irregular shaped areas are in fact hummocky flats with small depressions that can cover large areas (Figure 2). Because the model output is a pixelated raster model, and not polygons these areas can look irregular and are difficult to delineate through remote sensing imagery or on the ground. However, small wetlands interspersed throughout a landscape can provide critical ecosystem services, even being termed ‘wetlandscapes’ to reflect their complexity and difficult delineation (Thorslund et al., 2017[3]). However, we believe our continuous wetland probability better reflects potential wetland presence in these landscapes which are currently missing from most inventories.

Above we provide a qualitative example, but our accuracy assessment provides the quantitative analysis to support the strength of the model in this challenging landscape. The error of commission was not substantially high at only 10.24%. That is within the range of other published datasets and if used as a screening tool can be easily dealt with.

We have strengthened our statement in the document to make it clear that we do not intend for the WIP to replace the NWI (Lines 344 – 353). “While we used the NWI as a comparison baseline we want to make it explicitly clear that developing a method to replace the NWI was not our goal here and in no way do we recommend our WIP output as a replacement to the NWI. Rather the WIP output offers a different paradigm to wetland identification by providing a raster-based product that also provides continuous model probability. Our WIP probability output in many cases may be preferable to a vector based binary classification for wetland identification especially for wetlands that do not have clear borders or for use in other landscape models that require continuous raster datasets. The WIP probability output can also be used to detect wetlands that do not meet the jurisdictional or Cowardin

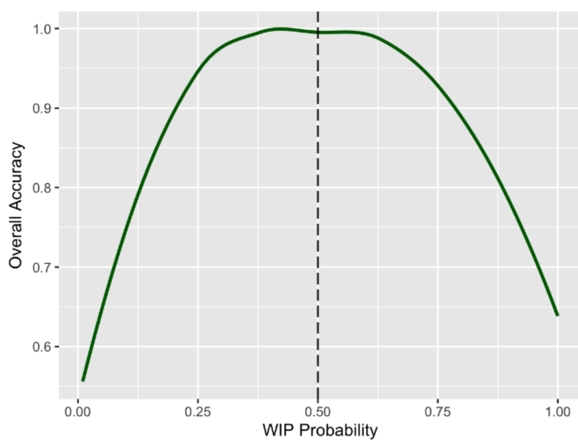
---

[1]NASA CMS https://cce-datasharing.gsfc.nasa.gov/cmsprojects/list/h/0/?projType=project&progID=5&projID=4096

[2] NASA CMS https://cce-datasharing.gsfc.nasa.gov/cmsprojects/list/h/0/?projType=project&progID=5&projID=4869).

definition of wetlands, yet still offer substantial ecosystem services such as carbon storage, habitat, and drought refugia. While not a replacement to the NWI, the WIP tool can be a screening tool to identify omitted wetlands in the NWI (as high as 47.5% in our study area) and to reduce bias for future NWI updates created through traditional manual photo interpretation."

To reiterate, while it is fine to convert the WIP tool continuous probability index to a binary classification that was not our goal here but understand that land managers may set thresholds for determining wetland presence to create output products like the NWI. To further address this comment, we have added an additional WIP tool output that provides information to help select a threshold to reduce errors of omission or commission and optimized overall accuracy (Figure 3, left). We used 0.5 to create our binary classification, but if users wanted to remove the commission error they could raise the threshold value without a huge loss in overall accuracy.



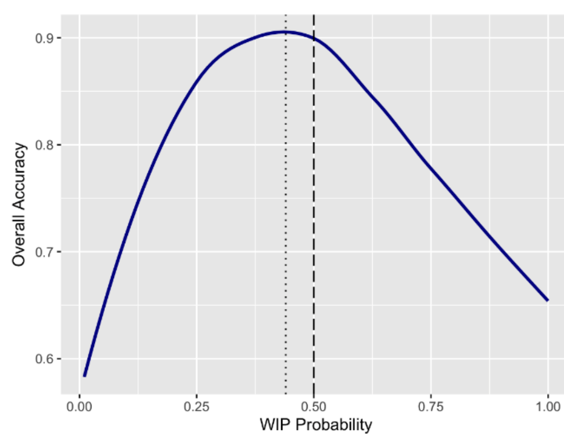


Figure 3: The left-hand graph shows how the overall accuracy of the WIP model varies across the range of WIP probability values taken from the training reference data. A threshold of 0.50 was chosen in order to validate the results for wetland and upland classification. The right-hand graph shows how the overall accuracy of the classification for the validation data used in the WIP model varies across the range of WIP probability values. Users who are interested in optimizing the overall accuracy may want to reduce the threshold to 0.44.

Comment 3: "Secondly, it is unclear how the proposed framework would function on other areas, as the authors did not develop a generalized model that could potentially be applied elsewhere. Without addressing the transferability of the method, it would remain a case study."

The model was initially developed as a tool to identify wetlands that are difficult to detect in the spectral imagery as requested by the Washington State Department of Natural Resources. It was initially tested on several watersheds across the PNW and published in a report the WA State DNR[3], however the results were not peer-reviewed outside of internal agency peer-review. In order to build statistical confidence in the method we focused analysis on the Hoh watershed, which is considered one of the most difficult areas to map because of the tall evergreen trees and did an intensive effort to create labelled training data and follow up on the ground. This was done to provide confidence in the model and is now being rolled out across the Pacific Northwest, in other parts of the U.S. and now in 3 countries in Africa. The WIP tool was absolutely developed with flexibility in mind with the creation of the wetland indicator framework and draws upon existing literature of proven input datasets for other areas. In some areas other input datasets may be more important model variables.

All machine learning models require training data, and this is true with our model, yet others have shown that machine learning can be a suitable framework and transferred elsewhere when trained or calibrated with local data on new areas (Rußwurm et al 2023). Indeed non-profit organizations exist to make repeatable models available as well as different training datasets so that these models can be trained on new data and calibrated to new locations (https://mlhub.earth/models). We addressed transferability in Section 5.2 (lines 362 – 385). We clearly state that the model itself is not transferable, but the method and the tool can be transferred to other areas and scaled to larger extents with collection of new training data and provide several examples of where it has been successfully run. The method here can be used as an example of how to more efficiently create training data through prelim models run, which can substantially reduce the necessary training data needed to run machine learning model.

Thorslund, Josefin, Jerker Jarsjo, Fernando Jaramillo, James W. Jawitz, Stefano Manzoni, Nandita B. Basu, Sergey R. Chalov, et al. “Wetlands as Large-Scale Nature-Based Solutions: Status and Challenges for Research, Engineering and Management.” *Ecological Engineering*, Ecological Engineering of Sustainable Landscapes, 108 (November 1, 2017): 489–97. https://doi.org/10.1016/j.ecoleng.2017.07.012.

Rußwurm, M., Courty, N., Emonet, R., Lefèvre, S., Tuia, D., Tavenard, R. (2023). End-to-end learned early classification of time series for in-season crop type mapping. ISPRS Journal of Photogrammetry and Remote Sensing, 196, 445-456. ISSN 0924-2716. https://doi.org/10.1016/j.isprsjprs.2022.12.016.

[3] https://www.dnr.wa.gov/publications/bc_fpb_wip_final_report_20210721.pdf