Hydroinformatics Blog Post

Organized by the [CUAHSI Informatics Standing Committee](). Contributions are welcome, please contact Veronica Sosa Gonzalez at vgonzalez@cuahsi.org.

Arc Hydro's WIM: A Machine Learning Framework for Wetland Identification

By: Gina L. O'Neil, Ph.D.

Wetlands are a vital ecosystem that provide ecological habitat, improve water quality, and ease flood and drought severity. As rapid development and climate change continue to threaten wetland health, it is increasingly important to build and share tools for wetland management and conservation. Tools to identify wetlands with greater efficiency than traditional manual processes are particularly valuable. The National Wetland Inventory (NWI) recognizes this need and is working with the wetland science community to develop an updated, semi-automated wetland mapping workflow. Arc Hydro's Wetland Identification Model (WIM) is a candidate toolset to meet this goal.

Method and Infrastructure

While there are many types of wetlands, all can be identified by common features, including the presence of hydrologic conditions that inundate the area, vegetation adapted for life in saturated soil conditions, and hydric soils. Remote sensing data offer new opportunities to observe these patterns at varying scales. LiDAR DEMs are particularly well-suited to modeling hydrologic drivers of wetland formation given their recent collection dates, wide availability, and high spatial resolution. Many studies have shown the ability of DEM derivatives to model flow convergence and act as proxies for near-surface soil moisture (e.g., Lang et al., 2013; Lang & McCarty, 2014; Millard & Richardson, 2013; Millard & Richardson, 2015; O'Neil et al., 2018).

WIM, implemented as Arc Hydro tools for ArcGIS Pro (Figure 1), is an automated framework for identifying likely wetlands using machine learning and user-specified predictor variables. The intended WIM workflow is:

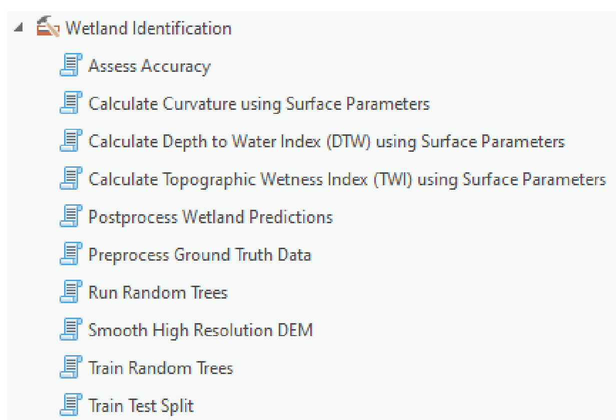

Figure 1 WIM Arc Hydro tools for ArcGIS Pro 3+

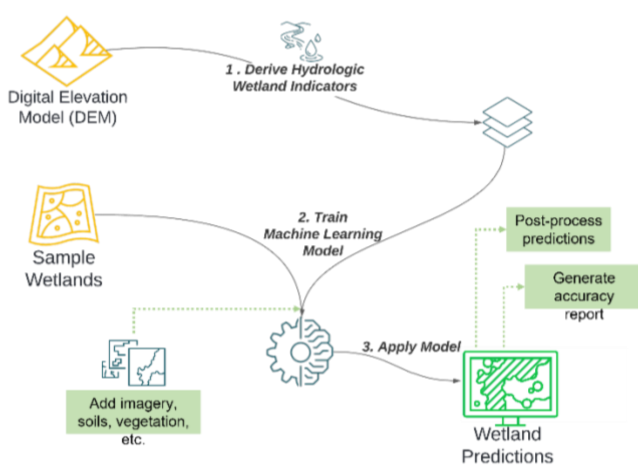

Figure 2 WIM Workflow

generate predictor variables, train and apply a Random Forests model (Breiman, 2001), and assess accuracy (Figure 2). In its baseline implementation, WIM calculates LiDAR-derived hydrologic drivers of wetland formation as predictor variables. Required input data are a high-resolution DEM, surface water data, and ground truth wetland coverage for a subset area to be used to train and assess the model. WIM outputs are wetland predictions and an accuracy report. WIM is designed to be flexible, allowing users to modify predictor variables, classify multiple wetland and nonwetland classes, and configure random forest model parameters. See below for highlights of the WIM component.

After applying smoothing to the input DEM, WIM automates the calculation of Mean Curvature, Topographic Wetness Index, and Depth-to-Water Index rasters (Figure 3). Calculations employ adaptive neighborhoods to capture hydrologic patterns at varying scales. The adaptive neighborhoods method uses varying window sizes to calculate curvature and slope, changing the window size based on local topographic variability.

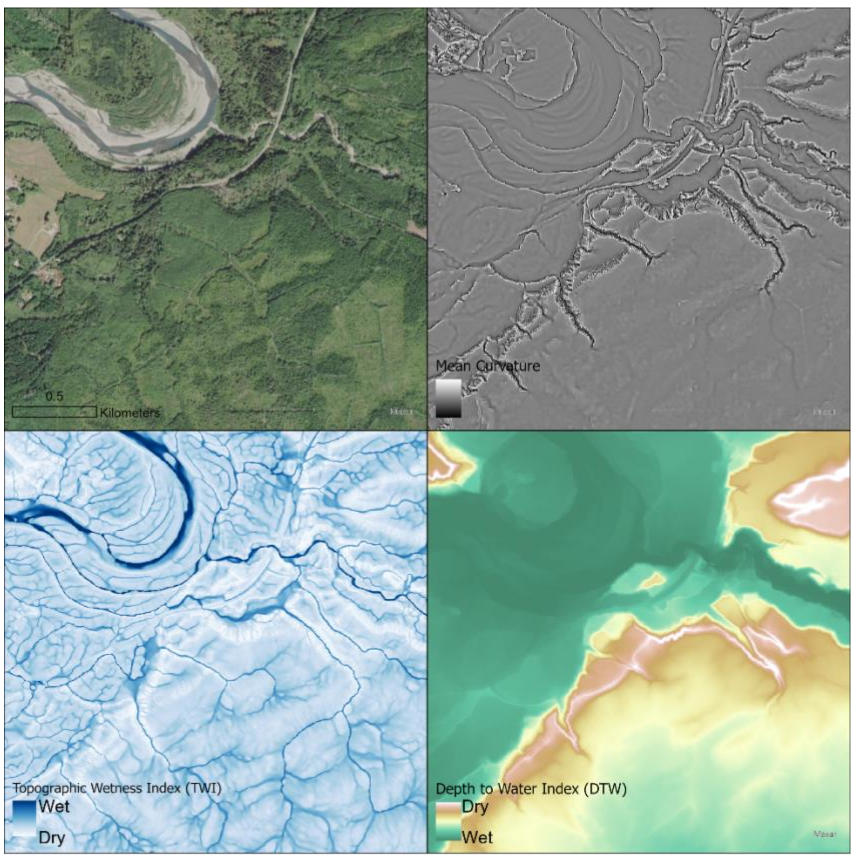

Figure 3 Topographic Predictor Variables

These baseline predictor variables are intended to be a starting point for wetland identification where the model relies primarily on high-resolution and widely available elevation data. However, users may incorporate additional vegetation and landcover information to capture a more robust set of wetland characteristics. WIM accepts additional predictor variables, as categorical or continuous data, in the TIFF raster format.

WIM produces Random Forest predictions in two formats: hard classifications where output raster cells are assigned a target classification, and classification probabilities where raster cells are assigned a probability of belonging to a target classification (Figure 4).

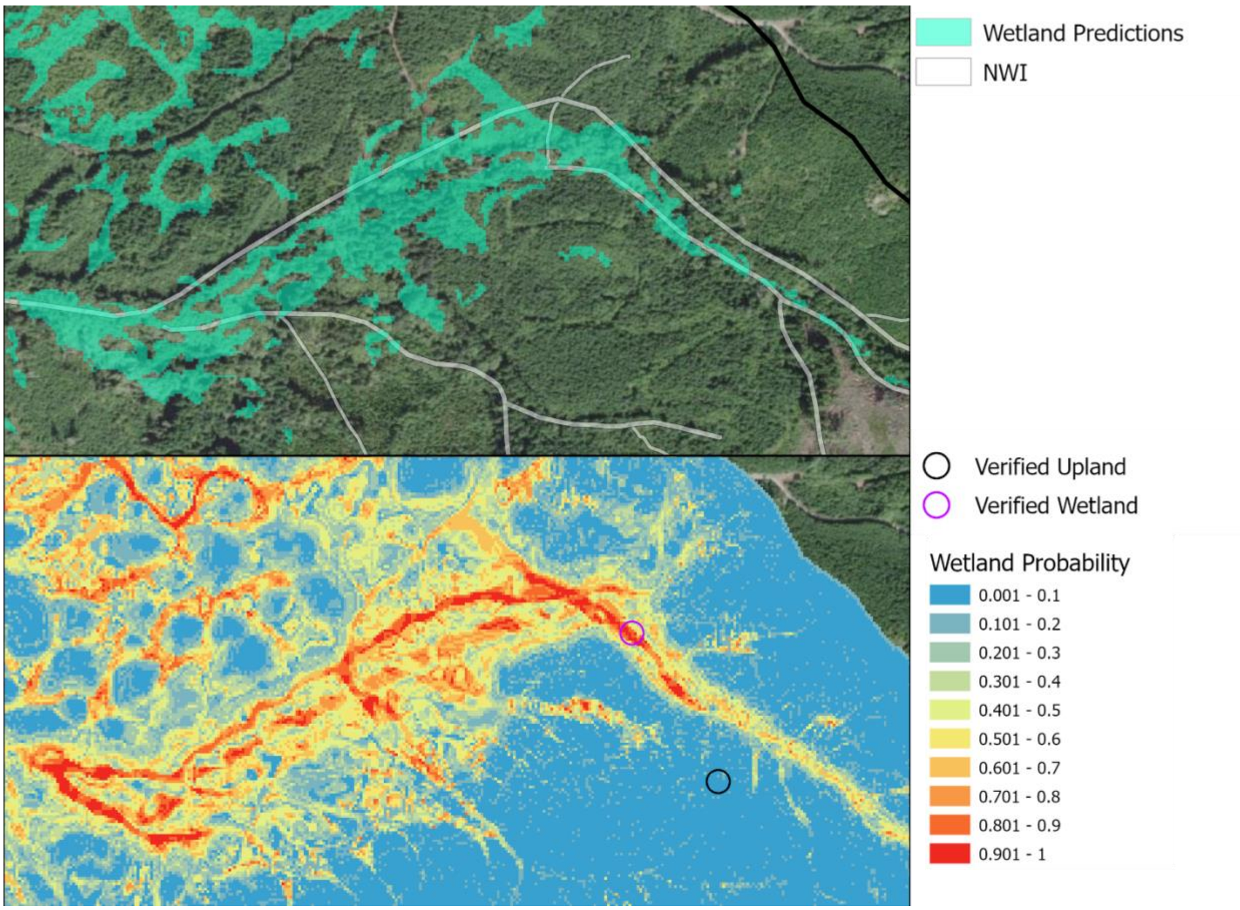

Figure 4 WIM Results formatted as hard classifications and classification probabilities

Results and Conclusion

The original WIM workflow, including the effectiveness of the default topographic predictor variables, was developed through three peer-reviewed publications (O'Neil et al., 2018, O'Neil et al., 2019, O'Neil et al., 2020). After calibration for four geographic regions in Virginia using a rich ground truth dataset of jurisdictionally confirmed wetlands, WIM was able to identify 80-90% of true wetlands across the sites. Wetland prediction precision varied from 22 to 69%, revealing a tendency to overpredict wetlands when using topographic inputs alone. Overall, the results suggest strong potential for WIM to support wetland delineation, but users must leverage the configurable components to develop best-performing models in their areas of interest. Like most modeling outputs, WIM outputs, cannot replace on-the-ground surveying and manual analyses. Instead, WIM outputs are intended to streamline and guide manual wetland mapping.

Since its implementation in the Arc Hydro software, the applicability of WIM to various use cases has been evaluated by Esri users and through collaborative research efforts. We are grateful to our collaborators within the NWI and its New Mapping Technologies working group for helping us to improve WIM and make it an impactful tool for wetland scientists. We hope that WIM can act as a user-friendly and reliable computational engine for ongoing wetland identification research.

Additional Resources

Install Arc Hydro for Pro 3+ to use the WIM tools and read more about the recent updates to WIM here.

Additional details on WIM functionality and demo are detailed in this recent Esri webinar.

Contact the Arc Hydro team through Esri Community or email Gina directly at goneil@esri.com.

Acknowledgments

Dr. Jon Goodall and the UVa Hydroinformatics Lab

Dr. Dean Djokic and the Arc Hydro team

Dr. Megan Lang, Chief Scientist, National Wetland Inventory

About the author:

Dr. O'Neil joined Esri Professional Services in 2019, following the completion of her Ph.D. in Civil and Environmental Engineering from the University of Virginia. Dr. O'Neil's dissertation focused on the development of a wetland identification model, involving work in the areas of hydrologic and geomorphologic modeling, GIS, python development, and machine learning. She continues to research hydrographic feature extraction, including wetland identification, as a member of Esri's Arc Hydro team. As a Data Scientist with Esri, Dr. O'Neil works on a range of geospatial projects in the natural resources and water resources sector.

Author's website: *https://www.linkedin.com/in/gina-o-neil/*

References

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Lang, M., & McCarty, G. (2014). Light detection and ranging (LiDAR) for improved mapping of wetland resources and assessment of wetland conservation projects, (September), 7. Retrieved from http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1260970.pdf

Lang, M., McCarty, G., Oesterling, R., & Yeo, I. Y. (2013). Topographic metrics for improved mapping of forested wetlands. Wetlands, 33(1), 141–155. https://doi.org/10.1007/s13157-012-0359-8

Millard, K., & Richardson,M. (2013). Wetland mapping with LiDAR derivatives, SAR polarimetric decompositions, and LiDAR-SAR fusion using a random forest classifier. Canadian Journal of Remote Sensing, 39(4), 290–307. https://doi.org/10.5589/m13-038

Millard, K., & Richardson, M. (2015). On the importance of training data sample selection in Random Forest image classification: A case study in peatland ecosystem mapping. Remote Sensing, 7(7), 8489–8515. https://doi.org/10.3390/rs70708489

O'Neil, G. L., Goodall, J. L., Behl, M., Saby, L. (2020). Deep Learning using Physically-Informed Input Data for Wetland Identification. *Environmental Modelling and Software*. 104665. https://doi.org/10.1016/j.envsoft.2020.104665.

O'Neil, G. L., Saby, L., Band, L. E., Goodall, J. L. (2019). Effects of LiDAR DEM Smoothing and Conditioning Techniques on a Topography-Based Wetland Identification Model. *Water Resources Research*, 55. https://doi.org/10.1029/2019WR024784.

O'Neil, G. L., Goodall, J. L., Watson, L. T. (2018). Evaluating the potential for site-specific modification of LiDAR DEM derivatives to improve environmental planning-scale wetland identification using random forest classification. *Journal of Hydrology*, 559, 192-208. https://doi.org/10.1016/j.jhydrol.2018.02.009.