Hydrology and
Earth System
Sciences

# The Wetland Intrinsic Potential tool: mapping wetland intrinsic potential through machine learning of multi-scale remote sensing proxies of wetland indicators

Meghan Halabisky[1], Dan Miller[2], Anthony J. Stewart[1], Amy Yahnke[3], Daniel Lorigan[2], Tate Brasel[2], and Ludmila Monika Moskal[1]

[1]School of Environmental and Forest Sciences, University of Washington, Seattle, WA, USA
[2]TerrainWorks, Seattle, WA, USA
[3]Washington Department of Ecology, Lacey, WA, USA

**Correspondence:** Meghan Halabisky (halabisk@uw.edu)

**Abstract.** Accurate, unbiased wetland inventories are critical to monitor and protect wetlands from future harm or land conversion. However, most wetland inventories are constructed through manual image interpretation or automated classification of multi-band imagery and are biased towards wetlands that are easy to directly detect in aerial and satellite imagery. Wetlands that are obscured by forest canopy, that occur ephemerally, and that have no visible standing water are, therefore, often missing from wetland maps. To aid in the detection of these cryptic wetlands, we developed the Wetland Intrinsic Potential (WIP) tool, based on a wetland-indicator framework commonly used on the ground to detect wetlands through the presence of hydrophytic vegetation, hydrology, and hydric soils. Our tool uses a random forest model with spatially explicit input variables that represent all three wetland indicators, including novel multi-scale topographic indicators that represent the processes that drive wetland formation, to derive a map of wetland probability. With the ability to include multi-scale topographic indicators that help identify cryptic wetlands, the WIP tool can identify areas conducive to wetland formation while providing a flexible approach that can be adapted to diverse landscapes. For a study area in the Hoh River watershed in western Washington, USA, classification of the output probability with a threshold of 0.5 provided an overall accuracy of 91.97 %. Compared to the National Wetlands Inventory, the classified WIP tool output identified over 2 times the wetland area and reduced errors of omission from 47.5 % to 14.1 % but increased errors of commission from 1.9 % to 10.5 %. The WIP tool is implemented as an ArcGIS toolbox using a combination of R and Python scripts.

## 1 Introduction

Wetlands provide a vast array of ecosystem services, including water storage, carbon sequestration, sediment removal, and wildlife habitat (Davidson et al., 2019). Despite their value, over 50 % of wetlands worldwide have been lost through draining and filling (Davidson, 2014; Davidson and Finlayson, 2018). The remaining wetlands are surrounded by an increasingly modified landscape that can adversely affect both the condition and function of wetlands (Calhoun et al., 2017; Tiner, 2009). An accurate inventory of wetland locations is necessary to protect wetlands from further land cover changes and degradation. However, in many regions, wetland inventories do not exist or are inaccurate with high errors of omission (Halabisky, 2018). Wetlands under partial or closed canopy, ephemeral wetlands that are flooded for only a portion of the year, and wetlands with no visible standing water (i.e., saturated soils) are often missing from wetland inventories (Halabisky, 2018).

On the ground, wetlands are identified by the presence of three wetland indicators: hydrophytic vegetation, surface hydrology (e.g., inundation or signs of inundation), and hydric soils (Cowardin, 1979). At the landscape scale, how-

ever, wetlands are primarily identified using remotely sensed data. Hence, wetland inventories have been commonly created through manual image interpretation by directly identifying wetland characteristics (e.g., presence of water) or proxies that represent wetland characteristics (e.g., areas of low slope represent areas more likely to be flooded) in imagery (Brinson, 1993; Tiner, 1990). In the last decade, there have been great strides in mapping wetlands through automated or semi-automated processes using remotely sensed multispectral data that provide indicators of hydric soil and hydrophytic vegetation (Dronova, 2015; Halabisky, 2019; Kloiber et al., 2015; Lang and McCarty, 2009).

However, small, ephemeral wetlands with dense canopy cover are virtually undetectable in aerial imagery (Fig. 1). Even in areas without dense canopy, trees and topography can create shadows that can resemble flooded wetlands in the imagery and confuse automated methods based on spectral features alone. Wetlands with fluctuating water levels, or wetlands without visible surface-water expression, may not be easily detected in the imagery due to a mismatch in the image acquisition timing or poor spectral or spatial image resolution. These cryptic, undetected wetlands can comprise a substantial portion of total wetland area in certain landscapes (Creed et al., 2003; Janisch et al., 2011).

With the widespread availability of lidar-derived elevation data, topographic information has increasingly been included as an indicator of wetland potential in analysis of remotely sensed data. Coincident with lidar availability, the development of machine learning techniques has enabled the analysis of large multivariable datasets both at very high spatial resolution (e.g., Ågren et al., 2021; O'Neil et al., 2020; Montgomery et al., 2021; Du et al., 2020) and over very large areas (e.g., Zhang et al., 2023; Woznicki et al., 2019). This work seeks to build on those efforts, with a focus on the development of new methods to identify difficult-to-detect, cryptic wetlands. We seek to incorporate the same suite of physical indicators used for ground-based wetland mapping, although using remotely sensed data, so that these methods can be applied at regional extents. Because those cryptic wetlands are both difficult to detect with optical or multispectral imagery and typically small, though potentially numerous, we rely on high-resolution lidar elevation data to resolve intrinsic topographic controls on water flux. Recognizing that topographic features that affect water fluxes through a landscape span a large range of spatial extents, we include tools developed to measure topographic attributes over multiple length scales.

## 1.1 Topographic and hydrologic indices

Cryptic wetlands can be indirectly identified by mapping the hydrologic processes driving wetland inundation patterns (Lang et al., 2013; Wu and Lane, 2017). Many studies have shown that the delineation of terrain attributes indicative of these processes is effective at predicting wetland locations (Lang et al., 2013; Maxwell et al., 2016; O'Neil et al., 2018,

2020), particularly when these attributes are calculated using high-resolution lidar elevation data. The primary attributes explored in the literature include local topographic position, slope gradient and curvature, the topographic wetness index (TWI), and the cartographic depth to water (DTW) (Maxwell et al., 2018). These attributes are calculated using digital elevation models (DEMs), which provide point measures of elevation over a regular grid.

Local topographic position provides a measure of vertical position in the landscape and can differentiate between the higher and low-lying terrain where wetlands tend to occur. There are a variety of methods to calculate local topographic position (Newman et al., 2018), all of which involve comparison of the elevation of a DEM grid point to the elevations of all the other grid points within a neighborhood of a specified radius. The variety of methods for local topographic position differs in how these comparisons are made. The center-cell elevation can be compared to the minimum and maximum elevations or to the mean elevation. That elevation difference can then be used directly or normalized by the range of elevations, the mean, or the standard deviation. For mapping wetland potential, measures of local topographic position are used for identifying landforms where water may tend to accumulate (Branton and Robinson, 2020; Riley et al., 2017).

Slope gradient and curvature are related to the direction and rate of surface and shallow subsurface-water flow across the terrain. Water tends to drain quickly from steep slopes and less quickly from lower-gradient slopes. Curvature can indicate areas where flow directions converge and where rates of flow decrease, both of which are associated with zones of increased soil moisture (Fink and Drohan, 2016).

The topographic wetness index is based on a simple conceptual model of shallow subsurface flow (Beven and Kirkby, 1979). The depth of soil saturation at a point, or at a DEM cell, is determined by the amount of water flowing to that cell, the degree of convergence or divergence of the topography there, and the effective velocity (the Darcy velocity) of saturated flow through the soil. Under steady-state rainfall, the amount of water is proportional to the area of the flow tube draining to that DEM cell. The effect of topographic convergence is accounted for by dividing that contributing area by the width of a contour line crossed by water flowing through the cell, giving the specific contributing area $A_s$. The flow velocity is proportional to the tangent of the slope $\theta$. With these definitions, saturation depth varies with $A_s / \tan\theta$. The topographic wetness index is defined as $\text{TWI} = \ln(A_s / \tan\theta)$. TWI, also called the compound topographic index (CTI), is used as a topographic indicator of relative soil moisture (Kopecký et al., 2021).

The cartographic depth to water (DTW) provides an estimated depth from the ground surface to the saturated zone in the soil column (Murphy et al., 2007). DTW calculates the elevation difference between a DEM grid point and a nearby location of water at the ground surface, such as a river or lake, which is included as inputs in the model. The location of the

associated surface-water point is found by repeatedly finding the adjacent DEM cell with the smallest downslope elevation difference, jumping to that point, and repeating that procedure until surface-water is encountered; that is, the least-cost path using slope as the measure of cost. Small DTW values can be good indicators of wetland occurrence (White et al., 2012). Height above the nearest drainage (Nobre et al., 2011) offers an alternative method for estimating depth to the saturated zone. This method finds the elevation difference between a DEM cell and the surface-water point it drains to, based on the downslope flow path traced from each DEM cell (Rennó et al., 2008).

Various combinations of these terrain attributes have all been used for wetland identification. The degree of success and the attributes of primary importance vary across studies. This variability reflects not only the intrinsic differences across landscapes (Branton and Robinson, 2020), but also differences in the spatial resolution of the data used (Fink and Drohan, 2016), preconditioning of those data (O'Neil et al., 2018), and the specific topographic attributes examined. Another source of variability is differences in the spatial scale of the terrain attributes examined.

## 1.2 Multi-scale indices for complex topographic features

All of the terrain measures outlined above are dependent on the length scales over which measurements are calculated. For example, the local topographic position will vary depending on the neighborhood radius used (De Reu et al., 2013). A neighborhood spanning 20 m will differentiate tree-fall pits and mounds (if resolved by the DEM), while a neighborhood radius spanning kilometers will differentiate valley floors and ridge tops. Gradient and curvature measured over a 5 m length might also detect pits and mounds; gradient and curvature measured over 50 m will miss those pits and mounds but will detect a broad swale. With measurement of any topographic attribute, it is important to match the scales of the landforms we wish to detect and of the processes we wish to characterize.

In regions with complex topography, wetlands are found in topographic features that occur at multiple, interconnected scales (Bertasello et al., 2018; Wu and Lane, 2017). These scales and the degrees of interconnectedness vary across and within landscapes, depending on the landforms and hydrologic processes involved with wetland formation. This variability challenges our ability to use topographic attributes as general indicators of wetland potential. Is a 50 m wide depression as important as a 300 m wide depression, or a 1000 m wide depression? Does it matter if that 50 m depression is inside of a 1000 m depression? Likewise, does a depression on a valley floor have the same importance as a depression on a ridge top? Do the relevant scales differ across landscapes? To answer such questions, we must examine topographic attributes over multiple spatial scales.

## 1.3 Random forest

A large range of factors can be considered for wetland detection: multispectral imagery, multiple interacting topographic attributes over a range of spatial scales, or variations in substrate and land use. Analysis of such large and diverse datasets has benefited from the development of machine learning algorithms that do not require assumptions about the statistical distribution of input data (Maxwell et al., 2018). Non-parametric supervised classification approaches to land cover mapping produce more efficient and accurate results than earlier supervised parametric classification methods (e.g., maximum likelihood), primarily because satellite image data values are not normally distributed (Wulder et al., 2019). Random forest modeling is a commonly used non-parametric classification method (Breiman, 2001) that allows for the use of multiple, correlated input variables that are not normally distributed. These methods are increasingly being used for remote detection of wetlands (Halabisky et al., 2018; Kloiber et al., 2015; Maxwell et al., 2016; O'Neil et al., 2018; Zhang et al., 2023).

## 1.4 Research goal

Our goal was to develop a methodology to map intrinsic wetland potential using spatially explicit proxies for three wetland indicators: hydrophytic vegetation, hydrology, and hydric soils. A key objective was to test the inclusion of novel multi-scale terrain indices as a proxy for hydrologic features of variable shapes and sizes. We used a wetland-indicator framework to ensure comprehensive inclusion of wetland characteristics in development of a model, reflecting common wetland identification practices used by wetland ecologists. Framing model development using this framework (and developing a tool for model building with this in mind) helped us ground our approach in wetland ecology, enabling us to more easily bring our domain knowledge into a remote sensing solution. We applied and tested this approach in the Hoh River watershed of northwest Washington State, a particularly challenging area to map due to its complex topography; its tall and structurally complex forests; and the high variability of wetland types, including many ephemeral wetlands under dense canopy. We have incorporated the methods outlined here into a flexible ArcGIS toolbox called the Wetland Intrinsic Potential (WIP) tool to provide an end-to-end workflow that enables users to develop proxies of wetland indicators for their area of interest, including a wide range of topographic indices at multiple scales; to evaluate those indicators using a random forest model; and to use that model to create maps of wetland potential.

## 2 Study area and datasets

Here we define wetlands broadly as wet areas that have one of three wetland indicators: hydric vegetation, hydric soils, or

signs of inundation for at least 2 weeks during the growing season. We included both ephemeral and permanent water-bodies such as rivers and streams in our wetland definition. This decision was driven by the National Wetlands Inventory, which includes open water features such as lakes and rivers (Cowardin, 1979).

## 2.1 Study area

Data collection and analyses were performed in the middle and lower Hoh River watershed on the Pacific Northwest coast of Washington State, USA (Fig. 1). The Hoh River watershed contains a broad valley filled with alluvial and alpine glacial deposits, with steep alpine zones in predominately marine sedimentary rocks. The main river channel is active and unconfined and has formed terraces from previous higher flows. The Hoh River watershed is part of the Olympic temperate rainforest, receiving between 2.8 and 4.3 m of precipitation a year, based on PRISM 30-year normals (https://prism.oregonstate.edu/normals/, last access: 1 July 2022). While the majority of the lower watershed has undergone significant impacts from forest harvest, the upper watershed and area along the coast are within the Olympic National Park (ONP), where forest harvest is prohibited. The trees of the old-growth forest in the ONP can be up to 80 m in height (Harmon and Franklin, 1989), while the lower watershed is dominated by plantation forests managed for timber harvest (Pelt, 2001). The wetlands within the Hoh River watershed are diverse, from precipitation-driven peat bogs to riparian wetlands driven by streamflow inputs, as well as upland wetlands driven by surface-water flows and groundwater inputs. The National Wetlands Inventory identifies 3084 ha of wetlands (not including buffered National Hydrography Dataset streamlines), which comprises 4.4 % of the study area (69 558 ha). Many of the wetlands are under completely closed canopy cover; however, trees in areas of high levels of inundation can display stunted growth.

## 2.2 Data sources

We used multiple raster and vector datasets as inputs and training data into our random forest model: four-band aerial imagery acquired by the National Aerial Imagery Program (NAIP) in 2017 at 1 m resolution; a DEM and digital surface model (DSM) derived from lidar acquired in 2012 and 2013 by Watershed Sciences at 3 foot (= 0.914 m) pixel resolution and downloaded from the Washington State Department of Natural Resources Lidar data portal (https://lidarportal.dnr.wa.gov/, last access: 1 September 2020) (a DSM is a surface model created from the highest hit object in the lidar point cloud, subtracting the DEM from the DSM provides estimates of canopy height); and two data layers from the United States Department of Agriculture SSURGO soils data for the Hoh River watershed: the depth to any restricted layer and the hydraulic conductivity.

We also used the National Wetlands Inventory to create an initial sample training dataset for our preliminary model and for model output comparison. We removed buffered streamlines added into the NWI from the National Hydrography Dataset (NHD) because of the high positional error and the use of a default uniform buffer. Before processing, we rescaled all of the raster input datasets to match a 4 m pixel resolution. The reason for rescaling to a coarser pixel resolution was to reduce processing time while still preserving the resolution needed for wetland identification.

## 3 Methods

### 3.1 Developing proxies for wetland indicators

As a first step, we identified spatially explicit proxies that represent wetland indicators for hydrophytic vegetation, hydrology, and hydric soils that we could derive from freely available data sources (i.e., SSURGO soils), and aerial imagery or lidar data (Fig. 2). This framework provided us with a systematic way to consider the characteristics used to identify wetlands in the field and in imagery and determine the ideal proxy that could represent these characteristics as inputs in a random forest model. This framework also allowed us to test which group of indicators was most useful in identifying wetlands. We identified datasets that represented proxies based on our own experience and from a thorough literature review.

### 3.1.1 Hydrologic indicators

We identified surface-water directly in the imagery using the normalized difference water index (NDWI) created from the 2017 NAIP imagery. The NDWI is a normalized band ratio between the near-infrared and green bands and is useful to identify open water (McFeeters, 1996). We generated the TWI and the DTW indices using the Arc Hydro toolbox in ArcGIS Pro using permanent riverine features and waterbodies from the National Hydrography Dataset as input water features for DTW (O'Neil et al., 2018).

In addition to the TWI and the DTW, we explored the use of topographic indices calculated at different length scales.

Gradient and curvature were calculated using the methodology described by Zevenbergen and Thorne (1987) in which the shape of the ground surface at a DEM grid point is interpolated as a smooth polynomial surface that matches elevations of the grid point and its eight adjacent points. This methodology was modified to use a circular neighborhood (Shi et al., 2007) of arbitrary radius, with elevations along the circle interpolated from adjacent DEM grid points. This procedure allows estimates of gradient and curvature for each DEM point measured over any length scale, down to the DEM grid size. This is similar to the "local quadratic regression" described by Newman et al. (2022) but uses a slightly higher-order polynomial with an exact fit to only nine points,
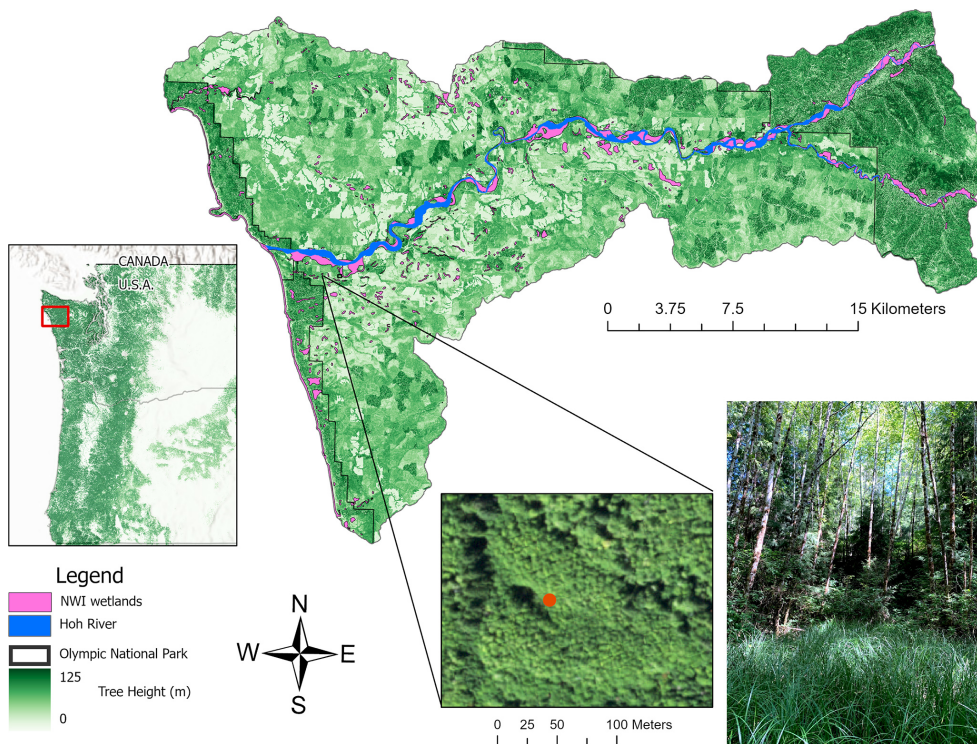
**Figure 1.** Study area in the Hoh rainforest located in the Pacific Northwest of the United States. Study area shows the variability in tree height, largely determined by a legacy of forestry. National Wetlands Inventory (NWI) wetlands are represented as pink polygons. Areas within Olympic National Park have not been logged. The photo on the right is a picture taken on the ground of the forested wetland shown in the aerial image from the 2017 National Aerial Imagery Program (NAIP) (orange dot). This wetland was missed in the NWI and is hard to detect in the aerial imagery. The dark areas in the aerial image are created by shadows from trees and are not standing water.
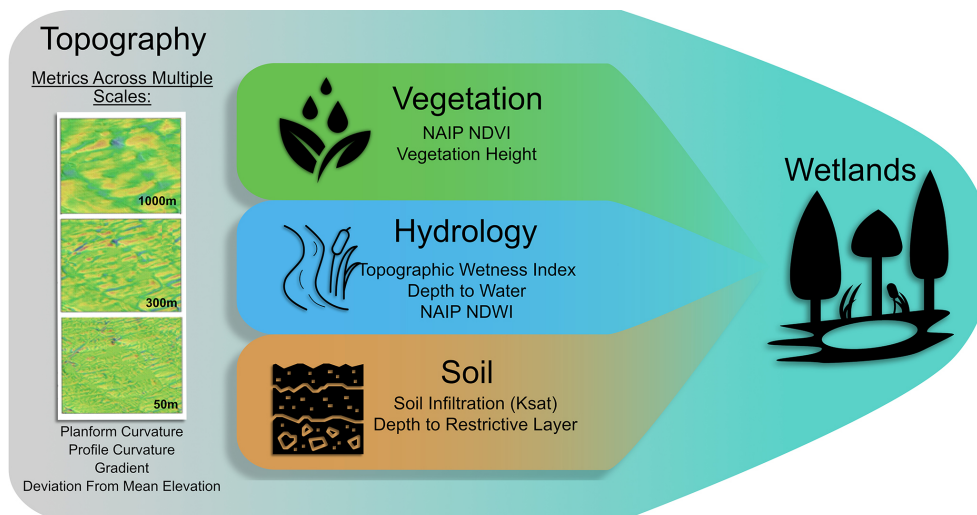


**Figure 2.** Input variables used in the random forest model represent proxies of wetland indicators used for wetland identification. Topographic indices, calculated at multiple scales, represent areas where water flows and collects. Profile curvature calculated at three different scales; 50, 300, 1000 m scales are shown as an example.

elevation at the current DEM grid point and elevations at eight equally spaced points on the circumference of a circle of a specified radius. This effectively smooths the DEM over the diameter of the circle with no increase in processing time with an increasing spatial scale; i.e., with larger circle diameters.

Several topographic-position indices have been developed to provide different measures of local relief (Newman et al., 2018). Of these, deviation from mean elevation (DEV) proved most appropriate for delineating low-lying areas across topographically diverse terrain: $DEV = (z - z_{mean})/\sigma$, where $z$ is elevation at the point of measurement, $z_{mean}$ is the mean elevation within a neighborhood of a specified radius, and $\sigma$ is the standard deviation of elevation within that neighborhood (Newman et al., 2018). Positive values of DEV indicate the point is higher than the mean of neighboring points (within the specified radius); negative values indicate the point is lower. Dividing by the standard deviation – a measure of how variable elevations are within the neighborhood – acts to normalize DEV values so that depressions in gentle, low-relief terrain, like broad river valleys, are recognized just as well as depressions in high-relief terrain, like alpine glacial cirques.

We calculated the topographic indices at five different length scales, 50, 150, 300, 500, and 1000 m, to approximate the variability of topographic features across the landscape. We visually assessed each topographic index at these scales and decided to only use the 50, 300, and 1000 m scales as they captured the most variability across the landscape and to reduce the number of input datasets to improve processing time.

Topographic indices were calculated using compiled Fortran programs from the NetStream program suite (Miller, 2003). These programs implement the procedures described above for calculating gradient, curvature, and local relief over any length scale. We developed an ArcGIS Pro toolbox called DEM Utilities for users to create topographic indices at multiple scales.

### 3.1.2 Hydrophytic vegetation and hydric soil indicators

To detect hydric vegetation, we created a normalized difference vegetation index (NDVI) from the 2017 NAIP imagery, rescaled to 4 m. NDVI is a normalized band ratio between the near-infrared and red bands that is useful for distinguishing wetland from non-wetland vegetation, as well as vegetation that may be stressed from inundation (Halabisky et al., 2011). We created two raster datasets from the SSURGO soil database, depth to any restricted layer and the hydraulic conductivity, to differentiate the soil properties that influence soil saturation and drainage.

### 3.2 Training data

Without knowing the location of forested wetlands a priori, it was difficult to develop an efficient and unbiased sampling design. Therefore, to aid in placement of points for a training dataset, we used a stratified random sample from a preliminary wetland model developed from the National Wetlands Inventory (NWI, https://www.fws.gov/program/national-wetlands-inventory, last access: 1 September 2020) for the Hoh River watershed. The preliminary model was based on a random forest model using the topographic indices and trained on 1000 wetland and 2000 non-wetland locations sampled from the NWI. The preliminary model then consisted of a raster of wetland probability with values from 0 to 1. To generate point locations for training the final model, we randomly placed 600 sample points equally into four strata based on the preliminary wetland probability raster: 0–0.25, 0.25–0.5, 0.5–0.75, 0.75–1.0. This provided an efficient way to identify potential wetland (high probability) and non-wetland (low probability) areas for a balanced point placement, as well as areas where there is high model uncertainty (i.e., probability near 0.5). We felt that stratifying the sample points using the preliminary model would reduce potential bias introduced by referencing the NWI better than if we had solely used the NWI to create our sample stratification.

Each sample point was evaluated by two analysts and labeled as wetland or upland using available datasets, including a hillshade and slope index from the lidar DEM; pre-existing wetland inventories including the NWI, NAIP imagery; and Forest Practices permits issued by the Washington State Department of Natural Resources, which indicate the presence of wetlands in areas where timber harvest occurs. If a point could not be determined as a wetland or non-wetland in aerial imagery or any other available datasets, it was marked as unknown. The challenge with this approach is that many of the areas with model certainty close to 0.5 are hard to assess using image interpretation. We made several site visits to ensure that assumptions made in manual image interpretation aligned with the ground truth. Out of the total points, 10 % were visited in the field. In 25 cases where the edge of the wetland was difficult to determine, the point was moved to an area clearly inside or outside the wetland. We removed two points because we could not agree on the label. We were unable to identify any wetlands formed by groundwater expression on slopes with no channel formation to include in the training or validation dataset. Therefore, we expected that the model could not predict or validate the presence of these types of slope wetlands.

### 3.3 Random forest model

We used the `randomForest` package in R (Breiman, 2001) with 598 sample points and 200 trees to train random forest models using 19 wetland indicators (Fig. 2). We decided to

use the most complete model with all 19 input data layers based on comparison of the out-of-bag error, a bootstrapped validation approach using sub-selections of the training data. The final model provided a raster showing the probability that a wetland will be found at each DEM grid cell (Fig. 3). The Gini coefficient provides a measure of the relative importance of each input indicator in the final model (Fig. 4). We classified the wetland probability values into a binary classification of upland and wetland classes using a probability threshold of 0.5.

## 3.4 Model validation

The WIP tool outputs the probability that a pixel is a wetland. It does not return a binary classification of wetland or upland that could be compared directly to a wetland inventory. To directly compare WIP modeled probabilities to the National Wetlands Inventory, we classified all pixels with a modeled probability of 0.5 or above as wetland and all others as not a wetland. The choice of 0.5 for the classification threshold simply reflects that, based on this model output, these pixels are more likely within a wetland than not. A lower threshold would increase the area classified as wetland; a higher threshold would reduce the area. We used this classification to randomly distribute 100 points within the wetland area and 200 points in the area outside the wetland classification (i.e., upland). We used the same two-person image interpretation process used for the training sample to label the 300 points. We moved five points because we could not detect the wetland edge and removed one point because the analysts could not agree on a label. We used this validation dataset to assess the accuracy of the random forest output and to identify errors of omission and commission.

## 4 Results

Our WIP model classification for the Hoh River watershed identified 6995 ha of wetlands using a threshold of 0.5, 2.25 times the area of wetlands mapped by the NWI (3084 ha). Model results for the Hoh River watershed can be viewed in detail on an online map available at https://arcg.is/0HXjbq0 (last access: 18 October 2023). The areas identified as wetland had an overall accuracy of 91.97 % (Table 1). The wetland error of commission (false positives) was 10.53 % and the error of omission (false negatives, missed wetlands) was 14.14 %. In contrast, using the same validation points, the current NWI for the Hoh River watershed had an overall accuracy of 83.95 %, with an error of commission of 1.89 % and an error of omission of 47.47 %.

Gradient calculated at a scale of 50 m, tree height (derived from lidar), and local elevation with a scale of 300 m were identified as the three variables that contributed the most importance to the model as measured by the Gini importance (Fig. 4). Amongst the categories shown in Fig. 2, there were

slightly more topographic indices loading strongly as predictors. Less significant metrics included the coarser 1000 m length scales of topography indices, with the exception of the 1000 m gradient metric. Other metrics of lower importance included the depth to the restrictive layer and the TWI.

Of the 14 wetland points misclassified in the WIP model as upland (errors of omission), 9 of them were within 5 m of the WIP wetland classification. Conversely, none of the 50 wetland points misclassified as upland (errors of omission) in the NWI model were within 5 m of the NWI.

## 5 Discussion

The Wetland Intrinsic Potential tool was designed to improve the detection of wetlands with a specific focus on increasing detection of cryptic wetlands obstructed by vegetation canopy, influenced by shadows from nearby objects and steep topography, and wetlands that do not have visible standing water for some part of the year. Our multi-scale machine learning approach improved the identification of wetlands that were missed in the existing NWI because of these challenging, yet common, remote-sensing issues. Using a WIP probability threshold of 0.5 for our validation, our model reduced the error of omission by over 33 % and the overall accuracy increased by 8 % when compared to the NWI. The increase in overall accuracy from the NWI was driven by the large reduction in errors of omission. It is important to keep in mind that the NWI has a minimum mapping unit of 0.5 ha, while our WIP tool did not set a minimum mapping and is only limited by the resolution of the input DEM.

For our study area, we found that a combination of proxies representing all three wetland indicators contributed to the overall model importance. However, indicators for hydrologic features and hydrophytic vegetation contributed the most. Specifically, three topographic indices that represent hydrologic features were among the top five input variables. It is unsurprising that measurements of gradient contributed the most, as wetlands are found primarily in areas of low slope. Tree height was the second contributing data layer, which may be driven by both the preference for timber companies to harvest outside of wetlands and the stunted height of trees in wetlands. We did notice that while including tree height improved our model, it also led to an increase in errors of commissions in harvested areas. Users who are interested in identifying wetlands in areas with timber harvest may choose not to include tree height to remove this bias. Measurements of DEV at a scale of 300 m were also a top contributing factor, which is useful in identifying medium-sized depressions. Proxies for hydric soils did not contribute as much to the model as other wetland-indicator proxies. The hydrologic indicator DTW contributed more than the TWI; however both were lower in importance than seven of the multi-scale terrain indices.
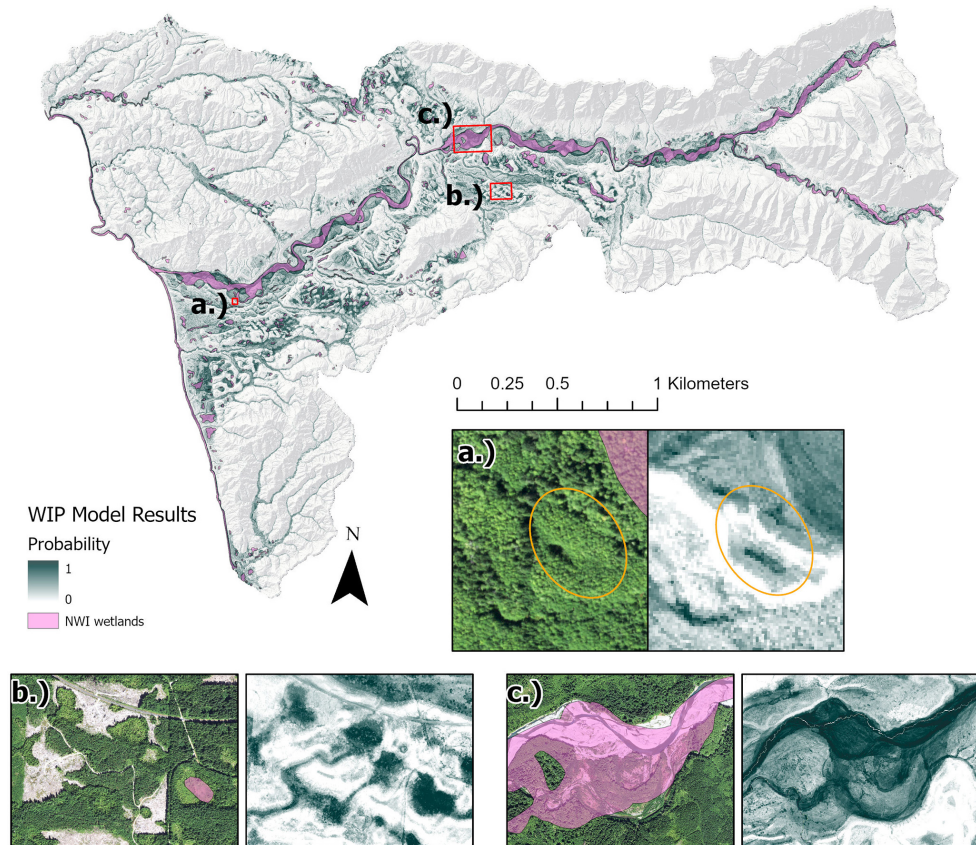
**Figure 3.** Wetland probability map of the entire study area with three examples: depressional wetland **(a)**, peatland **(b)**, and riverine wetland **(c)**.

**Table 1.** Accuracy assessment for WIP model based on 299 reference points (wetland = 99, upland = 200). A total of 275 of the 299 reference points were classified correctly. Wetland commission error was 10.53 % and omission error was 14.14 %.

|  |  | Reference data |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | Wetland | Upland | Total | Commission error |
| Model results | Wetland | 85 | 10 | 95 | 10.53 % |
|  | Upland | 14 | 190 | 204 | 6.86 % |
|  | Total | 99 | 200 |  |  |
|  | Omission error | 14.14 % | 5.00 % | Overall accuracy = 275 / 299 (91.97 %) |  |

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 for 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 definition of wetlands yet still offer substantial ecosystem services such as carbon storage, habitat, and drought refugia. While not a replacement for 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.

## 5.1 Model error

Systematic errors can provide clues for improving model performance. An exploration of the misclassified points shows that, for this study area, zones with the highest commission
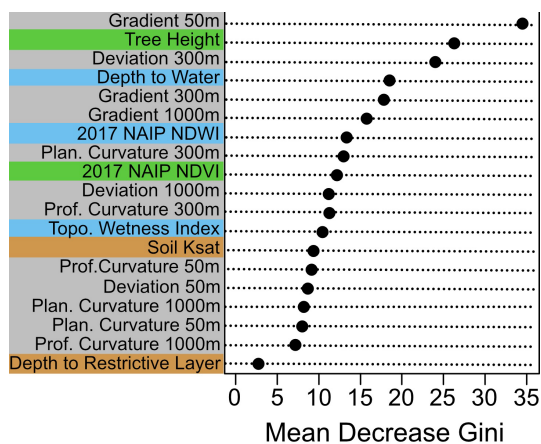
**Figure 4.** Gini coefficient output from the WIP tool random forest model, which is a measure of how each variable contributes to the homogeneity of the nodes and leaves in the resulting random forest. The variables at the top of the chart contributed the most to the model results.

error are located around the main river channel. This suggests that large floodplain and terrace zones should be delineated as categorical variables for input to the random forest model. Several old river terraces in old-growth forest stands were also misclassified as wetlands. Further investigation of these points on the ground suggests that these areas are right on the edge of meeting the definition of a wetland.

The majority of the errors of omission were within 5 of a mapped wetland, suggesting that the model can identify wetland areas but struggles to accurately delineate the wetland boundary. Some wetlands have clear boundaries, while others have a subtle wet–dry gradient. In these locations, the edges of wetlands can be hard to delineate on the ground. For wetland types without hard boundaries, the wetland probability output may provide more realistic information as it picks up the wet–dry gradient. Object-based approaches may help identify wetland boundaries in areas with more distinct wetland boundaries but would require an additional step of segmentation (Halabisky et al., 2011). Regardless, while the WIP tool can be useful to aid wetland delineations, standard field techniques on the ground are required for precise wetland delineation. We did not include slope wetlands in our study because of the difficulty of finding enough samples to train our model for this class of wetlands. We used a threshold of 0.5 and above to classify wetlands for our accuracy assessment. If users want to lower errors of omission, a lower threshold is recommended. Conversely, if users prefer to avoid over-mapping, a higher threshold should be selected.

## 5.2 Extension of model to new locations

The WIP tool is currently available as an ArcGIS toolbox and provides the ability to calculate multi-scale terrain indices. Our wetland-indicator framework allowed us to comprehen-

sively assess a full suite of variables for wetland identification while providing a flexible approach that can be adapted to other areas with different topographic features and wetland types. Extension of the random forest model to new areas requires new training data, which may limit its applicability. The ability of a model to predict wetland occurrence depends on how well the data used to train the model represent the range of wetland types and locations that exist on the ground. Our intention was not to develop a model that could be extended to new areas without the collection of new training data. A model trained on one study area but run on a different study area will not produce accurate results if the two study areas are dissimilar. The importance of different wetland indicators can vary for different study areas, but often the variables themselves will vary in importance as well. For example, in one watershed that contains many surface-water-driven wetlands, the topographic wetness index may be the most important variable that describes the variability between wetlands and uplands, but in another study area, DTW may be ranked as a more important contributing variable.

For application of the WIP tool in a new area, we recommend re-visiting the wetland-indicator framework and considering the wetland types in the area of interest and if new remote sensing proxies that we have not considered in our tool should be added. We have found that local knowledge is a critical component of developing solutions that improve model accuracy by identifying data proxies for local conditions. The WIP tool and the wetland-indicator framework are designed to be a workflow that can be updated and iteratively improved as new applications and datasets are identified. Indeed, for this project the wetland-indicator framework provided our team with a useful framework for testing out existing methods and ultimately led us to identify multi-scale terrain indices that helped identify cryptic forested wetlands and improve our model results. For this study, we tested the WIP tool out in one study area that is considered especially difficult to map. However, the WIP tool has also been applied to several new and distinct geographies. Seattle city government used the WIP tool to aid in wetland delineation in the Skagit basin of Washington (Environmental Science Associates, 2022), and the WIP tool was used to map wetlands on the island of Hawai'i (Tanh et al., 2022), where geology was a key predictor due to the influence of the volcanos.

The WIP tool is designed to be flexible and to allow for iterative improvements from inclusion of additional datasets (e.g., Sentinel-1 data). New datasets can easily be added into the raster stack of input variables in the ArcGIS toolbox. While our goal here was to develop a model with high accuracy and assess multiple wetland-indicator proxies, we also realize that our comprehensive approach may present hurdles to those in areas where some of the data inputs are unavailable. Here we developed a model to optimize for overall accuracy. However, a modified version of the WIP tool with fewer inputs can provide useful results, especially if the

https://doi.org/10.5194/hess-27-3687-2023

Hydrol. Earth Syst. Sci., 27, 3687–3699, 2023

probability gradient does not need to be converted into a hard classification. In cases where a hard wetland classification is not the goal, it may be justifiable to focus only on lidar-derived data inputs as a starting point and include spectral or soils data only if out-of-bag error is not adequate.

In this study area, we used a lidar-derived DEM to create our random forest model input datasets. While lidar data are becoming increasingly more widespread, they are not available everywhere. For areas without lidar coverage, the WIP tool can still be run with a DEM that was not created from a lidar acquisition. We have qualitatively tested out models using a 1/3 arcsec National Elevation Dataset DEM and an IfSAR-derived DEM created from Synthetic Aperture Radar and found them both to provide potentially adequate results, although at a coarser spatial resolution.

While we tested this model in a heavily forested area, we believe the WIP tool could also be applied to identify wetlands in other landscapes, such as agricultural areas, rangelands, and non-forested areas. However, none of the variables we included in our testing captured water movement influenced by human activity, such as water infrastructure, draining, ditching, or damming. Therefore, we expect that in areas with high levels of human modification of the hydrology the WIP model may identify areas of intrinsic potential and not necessarily areas that meet current definitions of a wetland.

Finally, our approach was a pixel-based probability and identifies areas of wetland intrinsic potential. However, others may prefer an object-based output (polygons). Object-based segmentation can be run on the WIP tool output to produce polygons and may improve results for areas where wetlands have more distinct boundaries.

## 5.3 Future directions

While the WIP tool is currently available as an ArcGIS toolbox, we are currently working to integrate components of the WIP into Esri's Wetland Identification Model (WIM), a random forest approach for wetland identification similar to the WIP. Like the WIP the WIM uses elevation-derived wetland indicators for its baseline implementation (O'Neil et al., 2018, 2019, 2020) and also accepts other raster-based predictors. WIM is available as part of the Arc Hydro toolset for ArcGIS Pro 2.5 and higher. Specifically, we are working to integrate multi-scale terrain indices and inclusion of point-based training data (https://community.esri.com/t5/water-resources-blog/wim-updates-for-arcgis-pro-3/ba-p/1233973, last access: 18 October 2023). Despite our enthusiasm for integrating the WIP into the WIM, we still see value in a stand-alone open-source tool for those without access to Esri products. We are currently working with Digital Earth Africa to develop an open-source Python-based tool to map wetland intrinsic potential using the Open Data Cube (https://www.digitalearthafrica.org, last access: 18 October 2023).

## 5.4 Model availability

We designed the WIP tools for this project expecting that they will evolve over time. The scripts and software are licensed as open source and are publicly available. The Python and R script and any new updates for the DEM utilities and Wetland tool ArcGIS Pro toolboxes are posted to a public GitHub repository at https://doi.org/10.5281/zenodo.10019936 (Miller et al., 2023). Bug reports, comments, and feature requests for these toolboxes can be submitted by posting an issue on GitHub. The random forest model can readily accommodate new terrain attributes as explanatory variables, and the scripts in the Wetland Tools toolbox can accommodate any input grid that can be imported to ArcGIS. We used the R-ArcGIS Bridge to build the Wetland Tools ArcGIS Pro toolbox that implements scripts that call R functions to build and apply random forest models.

## 6 Conclusions

Wetland inventories are critical sources of data to support wetland conservation prioritization, land use permits and regulations, monitoring, and wetland research. While wetland features may individually be small, collectively they cover vast areas and contribute to critical ecosystem services. The omission of a large percentage of wetlands within a region impedes our understanding of the total ecosystem services provided by wetlands and how specific land use regulations and policies may impact these services.

Accurate, unbiased wetland inventories are necessary to avoid further degradation and losses of wetlands. The WIP tool was specifically developed to identify cryptic wetlands that are missing from existing wetland inventories but can also be applied to areas where wetlands have not been mapped well. Our wetland-indicator framework, which includes spatial variables representing hydrophytic vegetation, hydrology, and hydric soils, can be used to quantify probability of wetland occurrence, including cryptic wetlands, with high confidence. The inclusion of novel multi-scale topographic attributes greatly improved model results as they were able to capture the variability of topographic features conducive to wetland formation. Our wetland-indicator framework provides a flexible approach that can be adapted to identify diverse wetland types across varied landscapes. We expect that the capabilities of the WIP tool will expand over time as users determine the most effective wetland indicators used for identifying wetlands in other regions.

lease v1.0.0. TerrainWorks maintains all software developed during collaborative projects. A comparison between the WIP outputs and the NWI for our study area can be viewed through ArcGIS online map at https://arcg.is/0HXjbq0 (Halabisky, 2023).

## References

Ågren, A. M., Larson, J., Paul, S. S., Laudon, H., and Lidberg, W.: Use of multiple LIDAR-derived digital terrain indices and machine learning for high-resolution national-scale soil moisture mapping of the Swedish forest landscape, Geoderma, 404, 115280, https://doi.org/10.1016/j.geoderma.2021.115280, 2021.

Bertassello, L. E., Rao, P. S. C., Jawitz, J. W., Botter, G., Le, P. V. V., Kumar, P., and Aubeneau, A. F.: Wetland-scape Fractal Topography, Geophys. Res. Lett., 45, 6983–6991, https://doi.org/10.1029/2018GL079094, 2018.

Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, Hydrolog. Sci. J., 24, 43–69, https://doi.org/10.1080/02626667909491834, 1979.

Branton, C. and Robinson, D. T.: Quantifying Topographic Characteristics of Wetlandscapes, Wetlands, 40, 433–449, https://doi.org/10.1007/s13157-019-01187-2, 2020.

Breiman, L.: Random Forests, Mach. Learn., 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Brinson, M. M.: A Hydrogeomorphic Classification for Wetlands, US Army Corps of Engineers, Waterways Experiment Station, Vicksburg, MS, USA, 101 pp., 1993.

Calhoun, A. J. K., Mushet, D. M., Bell, K. P., Boix, D., Fitzsimons, J. A., and Isselin-Nondedeu, F.: Temporary wetlands: challenges and solutions to conserving a "disappearing" ecosystem, Biol. Conserv., 211, 3–11, https://doi.org/10.1016/j.biocon.2016.11.024, 2017.

Cowardin, L. M., Carter, V., Golet, F. C., and LaRoe, E. T.: Classification of Wetlands and Deepwater Habitats of the United States, US Department of the Interior, Fish and Wildlife Service, Washington, DC, 103 pp., FWS/OBS-79/31, 1979.

Creed, I. F., Sanford, S. E., Beall, F. D., Molot, L. A., and Dillon, P. J.: Cryptic wetlands: integrating hidden wetlands in regression models of the export of dissolved organic carbon from forested landscapes, Hydrol. Process., 17, 3629–3648, https://doi.org/10.1002/hyp.1357, 2003.

Davidson, N. C.: How much wetland has the world lost? Long-term and recent trends in global wetland area, Mar. Freshwater Res., 65, 934–941, https://doi.org/10.1071/MF14173, 2014.

Davidson, N. C. and Finlayson, C. M.: Extent, regional distribution and changes in area of different classes of wetland, Mar. Freshwater Res., 69, 1525, https://doi.org/10.1071/MF17377, 2018.

Davidson, N. C., Dam, A. A. van, Finlayson, C. M., McInnes, R. J., Davidson, N. C., van Dam, A. A., Finlayson, C. M., and McInnes, R. J.: Worth of wetlands: revised global monetary values of coastal and inland wetland ecosystem services, Mar. Freshwater Res., 70, 1189–1194, https://doi.org/10.1071/MF18391, 2019.

De Reu, J., Bourgeois, J., Bats, M., Zwertvaegher, A., Gelorini, V., De Smedt, P., Chu, W., Antrop, M., De Maeyer, P., Finke, P., Van Meirvenne, M., Verniers, J., and Crombé, P.: Application of the topographic position index to heterogeneous landscapes, Geomorphology, 186, 39–49, https://doi.org/10.1016/j.geomorph.2012.12.015, 2013.

Dronova, I.: Object-Based Image Analysis in Wetland Research: A Review, Remote Sens., 7, 6380–6413, https://doi.org/10.3390/rs70506380, 2015.

Du, L., McCarty, G. W., Zhang, X., Lang, M. W., Vanderhoof, M. K., Li, X., Huang, C., Lee, S., and Zou, Z.: Mapping Forested Wetland Inundation in the Delmarva Peninsula, USA Using Deep Convolutional Neural Networks, Remote Sens., 12, 644, https://doi.org/10.3390/rs12040644, 2020.

Environmental Science Associates: TR-02 Wetland Assessment, Skagit River Hydroelectric Project, FERC NO. 553, Initial Study Report, Seattle City Light, Seattle, WA, 72 pp., 2022.

Fink, C. M. and Drohan, P. J.: High Resolution Hydric Soil Mapping using LiDAR Digital Terrain Modeling, Soil Sci. Soc. Am. J., 80, 355–363, https://doi.org/10.2136/sssaj2015.07.0270, 2016.

Halabisky, M.: Improved wetland identification for conservation and regulatory priorities, EPA Grant Number CD01J09401, Final Report, WA State Department of Ecology, 40 pp., 2018.

Halabisky, M.: WIP Map outputs with NWI for comparison, https://arcg.is/0HXjbq0 (last access: 18 October 2023), 2023.

Halabisky, M., Moskal, L. M., and Hall, S.: Object-based classification of semi-arid wetlands, J. Appl. Remote Sens., 5, 053511, https://doi.org/10.1117/1.3563569, 2011.

Halabisky, M., Babcock, C., and Moskal, L.: Harnessing the Temporal Dimension to Improve Object-Based Image Anal-

ysis Classification of Wetlands, Remote Sens., 10, 1467, https://doi.org/10.3390/rs10091467, 2018.

Harmon, M. E. and Franklin, J. F.: Tree Seedlings on Logs in *Picea-Tsuga* Forests of Oregon and Washington, Ecology, 70, 48–59, https://doi.org/10.2307/1938411, 1989.

Janisch, J. E., Foster, A. D., and Ehinger, W. J.: Characteristics of small headwater wetlands in second-growth forests of Washington, USA, Forest Ecol. Manag., 261, 1265–1274, https://doi.org/10.1016/j.foreco.2011.01.005, 2011.

Kloiber, S. M., Macleod, R. D., Smith, A. J., Knight, J. F., and Huberty, B. J.: A Semi-Automated, Multi-Source Data Fusion Update of a Wetland Inventory for East-Central Minnesota, USA, Wetlands, 35, 335–348, https://doi.org/10.1007/s13157-014-0621-3, 2015.

Kopecký, M., Macek, M., and Wild, J.: Topographic Wetness Index calculation guidelines based on measured soil moisture and plant species composition, Sci. Total Environ., 757, 143785, https://doi.org/10.1016/j.scitotenv.2020.143785, 2021.

Lang, M., McCarty, G., Oesterling, R., and Yeo, I.-Y.: Topographic Metrics for Improved Mapping of Forested Wetlands, Wetlands, 33, 141–155, https://doi.org/10.1007/s13157-012-0359-8, 2013.

Lang, M. W. and McCarty, G. W.: Lidar intensity for improved detection of inundation below the forest canopy, Wetlands, 29, 1166–1178, https://doi.org/10.1672/08-197.1, 2009.

Maxwell, A. E., Warner, T. A., and Strager, M. P.: Predicting Palustrine Wetland Probability Using Random Forest Machine Learning and Digital Elevation Data-Derived Terrain Variables, Photogramm. Eng. Rem. S., 82, 437–447, https://doi.org/10.14358/PERS.82.6.437, 2016.

Maxwell, A. E., Warner, T. A., and Fang, F.: Implementation of machine-learning classification in remote sensing: an applied review, Int. J. Remote Sens., 39, 2784–2817, https://doi.org/10.1080/01431161.2018.1433343, 2018.

McFeeters, S. K.: The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features, Int. J. Remote Sens., 17, 1425–1432, https://doi.org/10.1080/01431169608948714, 1996.

Miller, D., Halabisky, M., Lorigan, D., and Brasel, T.: Wetland Intrinsic Potential (WIP) Tool, Zenodo [code], https://doi.org/10.5281/zenodo.10019936, 2023.

Miller, D. J. Programs for DEM Analysis. Landscape Dynamics and Forest Management, General Technical Report RMRS-GTR-101CD, USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO, USA, 38 pp., 2003.

Montgomery, J., Mahoney, C., Brisco, B., Boychuk, L., Cobbaert, D., and Hopkinson, C.: Remote Sensing of Wetlands in the Prairie Pothole Region of North America, Remote Sens., 13, 3878, https://doi.org/10.3390/rs13193878, 2021.

Murphy, P. N. C., Ogilvie, J., Connor, K., and Arp, P. A.: Mapping wetlands: A comparison of two different approaches for New Brunswick, Canada, Wetlands, 27, 846–854, https://doi.org/10.1672/0277-5212(2007)27[846:MWACOT]2.0.CO;2, 2007.

Newman, D. R., Lindsay, J. B., and Cockburn, J. M. H.: Evaluating metrics of local topographic position for multiscale geomorphometric analysis, Geomorphology, 312, 40–50, https://doi.org/10.1016/j.geomorph.2018.04.003, 2018.

Newman, D. R., Cockburn, J. M. H., Draguţ, L., and Lindsay, J. B.: Evaluating Scaling Frameworks for Multi-scale Geomorphometric Analysis, Geomatics, 2, 36–51, https://doi.org/10.3390/geomatics2010003, 2022.

Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A., Waterloo, M., and Saleska, S.: Height Above the Nearest Drainage – a hydrologically relevant new terrain model, J. Hydrol., 404, 13–29, https://doi.org/10.1016/j.jhydrol.2011.03.051, 2011.

O'Neil, G. L., Goodall, J. L., and Watson, L. T.: Evaluating the potential for site-specific modification of LiDAR DEM derivatives to improve environmental planning-scale wetland identification using Random Forest classification, J. Hydrol., 559, 192–208, https://doi.org/10.1016/j.jhydrol.2018.02.009, 2018.

O'Neil, G. L., Saby, L., Band, L. E., and Goodall, J. L.: Effects of lidar DEM smoothing and conditioning techniques on a topography-based wetland identification model, Water Resour. Res., 55, 4343–4363, https://doi.org/10.1029/2019WR024784, 2019.

O'Neil, G. L., Goodall, J. L., Behl, M., and Saby, L.: Deep learning Using Physically-Informed Input Data for Wetland Identification, Environ. Modell. Softw., 126, 104665, https://doi.org/10.1016/j.envsoft.2020.104665, 2020.

Pelt, R. V.: Forest Giants of the Pacific Coast, Global Forest Society in association with University of Washington Press, Seattle, WA, 200 pp., 2001.

Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., Tomasella, J., and Waterloo, M. J.: HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia, Remote Sens. Environ., 112, 3469–3481, https://doi.org/10.1016/j.rse.2008.03.018, 2008.

Riley, J. W., Calhoun, D. L., Barichivich, W. J., and Walls, S. C.: Identifying Small Depressional Wetlands and Using a Topographic Position Index to Infer Hydroperiod Regimes for Pond-Breeding Amphibians, Wetlands, 37, 325–338, https://doi.org/10.1007/s13157-016-0872-2, 2017.

Shi, X., Zhu, A.-X., Burt, J., Choi, W., Wang, R., Pei, T., Li, B., and Qin, C.: An Experiment Using a Circular Neighborhood to Calculate Slope Gradient from a DEM, Photogramm. Eng. Rem. S.., 73, 143–154, https://doi.org/10.14358/PERS.73.2.143, 2007.

Tanh, L., Anokye, M., Lee, I., and Racette, C.: Hawaii Climate: Utilizing Earth Observations to delineate wetland extents, model sea level rise inundation risk, and assess impacts on historic Hawaiian Lands, NASA DEVELOP National Program, Technical Report, 25 pp., 2022.

Tiner, R. W.: Use of high-altitude aerial photography for inventorying forested wetlands in the United States, Forest Ecol. Manag., 33–34, 593–604, https://doi.org/10.1016/0378-1127(90)90221-V, 1990.

Tiner, R. W.: Global Distribution of Wetlands, in: Encyclopedia of Inland Waters, edited by: Likens, G. E., Academic Press, Oxford, 526–530, https://doi.org/10.1016/B978-012370626-3.00068-5, 2009.

White, B., Ogilvie, J., Campbell, D. M. H. M. H., Hiltz, D., Gauthier, B., Chisholm, H. K. H., Wen, H. K., Murphy, P. N. C. N. C., and Arp, P. A. A.: Using the Cartographic Depth-to-Water Index to Locate Small Streams and Associated Wet Areas across Landscapes, Can. Water Resour. J., 37, 333–347, https://doi.org/10.4296/cwrj2011-909, 2012.

Woznicki, S. A., Baynes, J., Panlasigui, S., Mehaffey, M., and Neale, A.: Development of a spatially complete

floodplain map of the conterminous United States using random forest, Sci. Total Environ., 647, 942–953, https://doi.org/10.1016/j.scitotenv.2018.07.353, 2019.

Wu, Q. and Lane, C. R.: Delineating wetland catchments and modeling hydrologic connectivity using lidar data and aerial imagery, Hydrol. Earth Syst. Sci., 21, 3579–3595, https://doi.org/10.5194/hess-21-3579-2017, 2017.

Wulder, M. A., Loveland, T. R., Roy, D. P., Crawford, C. J., Masek, J. G., Woodcock, C. E., Allen, R. G., Anderson, M. C., Belward, A. S., Cohen, W. B., Dwyer, J., Erb, A., Gao, F., Griffiths, P., Helder, D., Hermosilla, T., Hipple, J. D., Hostert, P., Hughes, M. J., Huntington, J., Johnson, D. M., Kennedy, R., Kilic, A., Li, Z., Lymburner, L., McCorkel, J., Pahlevan, N., Scambos, T. A., Schaaf, C., Schott, J. R., Sheng, Y., Storey, J., Vermote, E., Vogelmann, J., White, J. C., Wynne, R. H., and Zhu, Z.: Current status of Landsat program, science, and applications, Remote Sens. Environ., 225, 127–147, https://doi.org/10.1016/j.rse.2019.02.015, 2019.

Zevenbergen, L. W. and Thorne, C. R.: Quantitative analysis of land surface topography, Earth Surf. Proc. Land., 12, 47–56, https://doi.org/10.1002/esp.3290120107, 1987.

Zhang, X., Liu, L., Zhao, T., Chen, X., Lin, S., Wang, J., Mi, J., and Liu, W.: GWL_FCS30: a global 30 m wetland map with a fine classification system using multi-sourced and time-series remote sensing imagery in 2020, Earth Syst. Sci. Data, 15, 265–293, https://doi.org/10.5194/essd-15-265-2023, 2023.