This is a manuscript of an article that was accepted in February 2018 for publication in the Journal of Hydrology. The final publication is available at https://doi.org/10.1016/j.jhydrol.2018.02.009.

Evaluating the Potential for Site-Specific Modification of LiDAR DEM Derivatives to Improve Environmental Planning-Scale Wetland Identification using Random Forest Classification

Gina L. O'Neil[a], Jonathan L. Goodall[a,*], Layne T. Watson[b]

[a] Department of Civil and Environmental Engineering, University of Virginia, Charlottesville, Virginia 22904, USA.

[b] Department of Computer Science, University of Virginia, Charlottesville, Virginia 22904, USA; Departments of Computer Science, Mathematics, and Aerospace and Ocean Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, USA.

*Corresponding author. Tel.: (434) 243-5019

E-mail address: goodall@virginia.edu (J. Goodall)

Abstract

Wetlands are important ecosystems that provide many ecological benefits, and their quality and presence are protected by federal regulations. These regulations require wetland delineations, which can be costly and time consuming to perform. Computer models can assist in this process, but lack the accuracy necessary for environmental planning-scale wetland identification. In this study, the potential for improvement of wetland identification models through modification of digital elevation model (DEM) derivatives, derived from high-resolution and increasingly available Light Detection and Ranging (LiDAR) data, at a scale necessary for small-scale wetland delineations is evaluated. A novel approach of flow convergence modeling

is presented where Topographic Wetness Index (TWI), curvature, and Cartographic Depth-to-Water index (DTW), are modified to better distinguish wetland from upland areas, combined with ancillary soil data, and used in a Random Forest classification. This approach is applied to four study sites in Virginia, implemented as an ArcGIS model. The model resulted in significant improvement in average wetland accuracy compared to the commonly used National Wetland Inventory (84.9% vs. 32.1 %), at the expense of a moderately lower average non-wetland accuracy (85.6% vs. 98.0 %) and average overall accuracy (85.6% vs. 92.0%). From this, we concluded that modifying TWI, curvature, and DTW provides more robust wetland and non-wetland signatures to the models by improving accuracy rates compared to classifications using the original indices. The resulting ArcGIS model is a general tool able to modify these local LiDAR DEM derivatives based on site characteristics to identify wetlands at a high resolution.

KEYWORDS: wetlands, LiDAR, topographic indices, Random Forest

1. Introduction

Wetlands are important ecosystems that not only provide habitat for many plant and animal species, but also improve water quality, recharge groundwater, and ease flood and drought severity (Guo et al., 2017). Despite the ecological value of wetlands, their quality and presence are threatened by agricultural or development repurposing, pollutant runoff, and climate change (Klemas, 2011). Current estimates are that roughly 50% of wetlands have been lost globally since 1900 (Davidson, 2014) and approximately 53% of wetlands of the conterminous U.S. have been lost since the early 1600s (Dahl et al., 1991). The historic loss of wetlands and sustained threat to remaining wetlands has motivated increased efforts by scientists and government to protect and maintain these ecosystems.

U.S. federal regulations play an important role in the abatement of further wetland loss. One of the most important policies in support of this effort is Section 404 of the Clean Water Act, which protects the nation's waters, including wetlands. According to Page and Wilcher (1990), this law states that environmental planning entities must identify and assess environmental impact due to land development and water resource projects. This requires environmental planning entities, such as state departments of transportation (DOTs), to provide wetland delineations that are ultimately jurisdictionally confirmed by the U.S. Army Corps of Engineers (USACE). The USACE *Wetlands Delineation Manual* states that wetlands can be identified by environmental characteristics shared among the many wetland types. The USACE guidelines for wetland delineations use these common features and are based on the presence of hydrologic conditions that inundate the area, vegetation adapted for life in saturated soil conditions, and hydric soils (Environmental Laboratory, 1987).

Manual surveying by trained analysts will always be the most accurate method to delineate wetlands, however carrying out detailed field surveys can be time consuming and costly. According to estimates provided by representatives from the Virginia DOT (VDOT) Environmental Division, the costs of these delineations range from $60 to $140 per acre (~0.4 ha) (personal communication, November 28, 2017). These estimates are based on recent VDOT projects and can vary widely across projects. To offset these costs, the wetland permitting process could potentially be streamlined by supplementing and guiding the manual delineations with accurate digital wetland inventories. However, developing and updating wetland inventories can be expensive and technically challenging due to the complexity of wetland features (Kloiber et al., 2015). Furthermore, the existing national-scale wetland inventory in the U.S., the National Wetland Inventory (NWI), is not ideal for assisting in the permitting process. Despite being one of the most commonly used sources of wetland data in the U.S., NWI maps were never intended to map federally regulated wetlands (Cowardin & Golet, 1995; Environmental Laboratory, 1987) and research has shown that relying solely on the NWI may fail to protect a considerable fraction of wetlands (Morrissey & Sweeney, 2006). Thus, a wetland inventory with the reliability necessary to assist in the wetland permitting process is an unmet need.

Remote sensing has long been recognized as a powerful tool for identifying wetlands (Environmental Laboratory, 1987) and may offer an accurate and cost-effective way to fulfill this need (Guo et al., 2017; Lang et al., 2013; Lang & McCarty, 2014). Past studies have incorporated remote sensing data such as multispectral imagery, radar, and Light Detection and Ranging (LiDAR) for wetland identification. A review of wetland remote sensing studies of the past 50 years shows that most researchers incorporate multispectral imagery in wetland classifications (Guo et al., 2017). However, the incorporation of multispectral imagery can

weaken the potential for use during the wetland permitting process by introducing issues of resolution or accessibility. For example, the commonly used Landsat multispectral imagery is freely available on a national scale, but the 30 m resolution of this data can be too coarse to detect wetlands at a scale relevant to environmental planning entities, which can require a spatial accuracy of at least 1.5 m (VDOT Environmental Division, personal communication, November 28, 2017). While studies have shown higher resolution, multispectral data can result in accurate wetland classifications (e.g., Kloiber et al., 2015) these data can be inaccessible due to cost. Alternatively, LiDAR is remote sensing data that has been rapidly endorsed by the wetland science and management community for its growing availability and technological benefit to wetland mapping (Kloiber et al., 2015; Lang & McCarty, 2014). LiDAR sensors provide detailed information on the Earth's landscape and bare surface by collecting x, y, and z data that can then be interpolated to create digital elevation models (DEMs) (Lang & McCarty, 2014). LiDAR data availability has increased rapidly over the past 20 years, and although current coverage in the conterminous U.S. is at about one third, there is an ongoing effort by multiple federal agencies to hasten the collection of LiDAR data throughout the entire U.S. (Snyder & Lang, 2012). LiDAR derived DEMs have the ability to map wetlands by identifying areas of inundation based on topographic drivers of flow convergence and offer widely available, high-resolution data that could be utilized during the wetland permitting process. While conventional DEMs and their derivatives have been shown to be useful for wetland delineation (e.g., Hogg & Todd, 2007), LiDAR DEMs allow for more detailed mapping of topographic metrics (Lang & McCarty, 2014).

Previous research has shown that DEM derivatives have the potential to model spatial patterns of saturated areas, and that LiDAR DEM derivatives improve the ability of these metrics

to do so (e.g., Hogg & Todd, 2007; Lang et al., 2013; Millard & Richardson, 2013). Among the DEM derivatives found to be useful for this purpose are curvature, Topographic Wetness Index (TWI) and the Cartographic Depth-to-Water index (DTW) (e.g., Ågren et al., 2014; Lang et al., 2013; Murphy et al., 2009, 2011; Sangireddy et al., 2016). Curvature is defined as the second derivative of the input surface and can describe the degree of convergence and acceleration of flow (Moore et al., 1991). The TWI, developed by Beven and Kirkby (1979), relates the tendency of a site to receive water to the tendency of a site to evacuate water and is defined as

$$TWI = ln\left(\frac{\alpha}{\tan(\beta)}\right), \qquad (1)$$

where α is the specific catchment area, or contributing area per unit contour length, and tan(β) is the local slope. The DTW is a soil moisture index developed by Murphy et al. (2007) that is based on an assumption that soils very close in elevation to their assigned surface water are more likely to be saturated. The DTW model in grid form is calculated as

$$DTW\ (m) = \left[\sum\left(\frac{dz_i}{dx_i}\right)a\right] * \ x_c\ , \qquad (2)$$

where $\frac{dz}{dx}$ is the downward slope of a pixel, *i* is a pixel along a calculated least cost (i.e., slope) path to the assigned source pixel, *a* is 1 when the flow path is parallel to pixel boundaries or $\sqrt{2}$ when the flow crosses diagonally, and x_c is the pixel length (Murphy et al., 2007).

Although many studies have shown the benefit of using topographic indices to identify wetted areas, and the added benefit of deriving these indices at higher resolutions, there are unique challenges inherent to using LiDAR DEMs. Researchers have noted that LiDAR DEMs used for purposes related to modelling landform characteristics must be resampled to coarser resolutions and smoothed to overcome issues of increased “noise” from excessive topographic detail (MacMillan et al., 2003), with this topographic noise arising from DEMs on the order of 1

m pixel size (Richardson et al., 2009). Moreover, variations in DEM resolution result in significantly different spatial and statistical distributions of contributing areas and downslope flow path lengths (Woodrow et al., 2016), and at high resolutions, micro-topographic features can lead to highly variable slope values and provide unrealistic estimates of hydraulic gradients (Grabs et al., 2009; Lanni et al., 2011). Previous studies have acknowledged the negative effect that these micro-topographic features have on the ability of curvature (e.g., Sangireddy et al., 2016) and TWI (e.g., Sørensen & Seibert, 2007) to identify hydrologic features of interest. For example, Ågren et al. (2014) found that high-resolution DEMs (< 2 m) caused local TWI variations that are too strong to separate wetlands from uplands, whereas deriving the index from coarser (> 24 m) DEMs reduced these variations but resulted in poorly delineated flow channels and local depressions. In contrast, the researchers also concluded that DTW derivations are not sensitive to scale, but have suggested that the DTW could be further optimized (Ågren et al. 2014).

LiDAR DEM data and other remote sensing data are commonly used to map wetlands through supervised classification algorithms. Random Forest (RF) classification is a relatively new supervised classification method that is widely used for its ability to handle both continuous and categorical, high-dimensional data and produce descriptive variable importance measures (Millard & Richardson, 2015; Rodriguez-Galiano et al., 2012). RF has been shown to produce higher accuracies than other classification techniques, such as maximum likelihood, when incorporating multisource data (Duro et al., 2012; Miao et al., 2012; Rodriguez-Galiano et al., 2012). Furthermore, studies have shown that LiDAR DEM metrics are suitable input variables for the RF approach (e.g., Deng et al., 2017; Kloiber et al., 2015; Zhu & Pierskalla, 2016), and

that using this classifier has strong potential to improve mapping and imagery classification of wetlands (e.g., Millard & Richardson, 2013).

Many previous studies have relied primarily on ecological factors and spectral indices provided by multispectral imagery to classify wetlands, and fewer studies have evaluated the predictive power of LiDAR DEM data alone for this purpose. The primary objective of this study was to further advance the application of LiDAR DEM derivatives to wetland mapping by evaluating the potential of modified TWI, DTW, and curvature grids to address limitations noted by researchers and identify small (i.e., environmental planning-scale) wetlands across varying ecoregions. RF classifications of original and modified TWI, curvature, and DTW, where the TWI and curvature were modified via smoothing and the DTW was modified via adjustments to the input slope grid, along with ancillary national-scale soil data were assessed against field-mapped test data and compared to NWI maps to identify the best performing models. Accuracy assessments of these classifications provided a measure of the benefits and costs of modifying these input data. This approach was applied to four study sites across varying ecoregions of Virginia and implemented in ArcGIS with the potential for further refinement and utility by environmental planning entities.

2. Study Areas

The four sites in this study were selected due to availability of VDOT wetland delineations and LiDAR DEMs, and to have applications of this approach across varying ecoregions of Virginia. As seen in Figure 1, the study sites span five of the seven level III EPA ecoregions of Virginia: the Piedmont (45), the Mid-Atlantic Coastal Plain (63), the Northern Piedmont (64), the Southeastern Plains (65), and the Ridge and Valley (67). According to the

EPA (2013), the Piedmont ecoregion is considered the non-mountainous region of the Appalachians Highland and is comprised of transitional areas between the mountainous Appalachians to the northwest and the relatively flat coastal plain to the southeast. The soils in this region tend to be finer textured than in ecoregions 63 and 65. The Mid-Atlantic Coastal Plain is characterized by low, nearly flat plains with many swamps, marshes, and estuaries. The region has a mix of coarse and finer textured soils and poorly drained soils are common here. The Northern Piedmont consists of low rounded hills, irregular plains, and open valleys. It is a transitional region between the low mountains in ecoregion 66 and the flat coastal area of ecoregions 63 and 65. The Southeastern Plains are irregular and have a mosaic of cropland, pasture, woodland, and forest. The subsurface is predominantly sands, silts, and clays. The Ridge and Valley ecoregion is relatively low-lying and characterized by alternating forested ridges and agricultural valleys. Additional information describing the conditions of each study site can be found in Table 1.

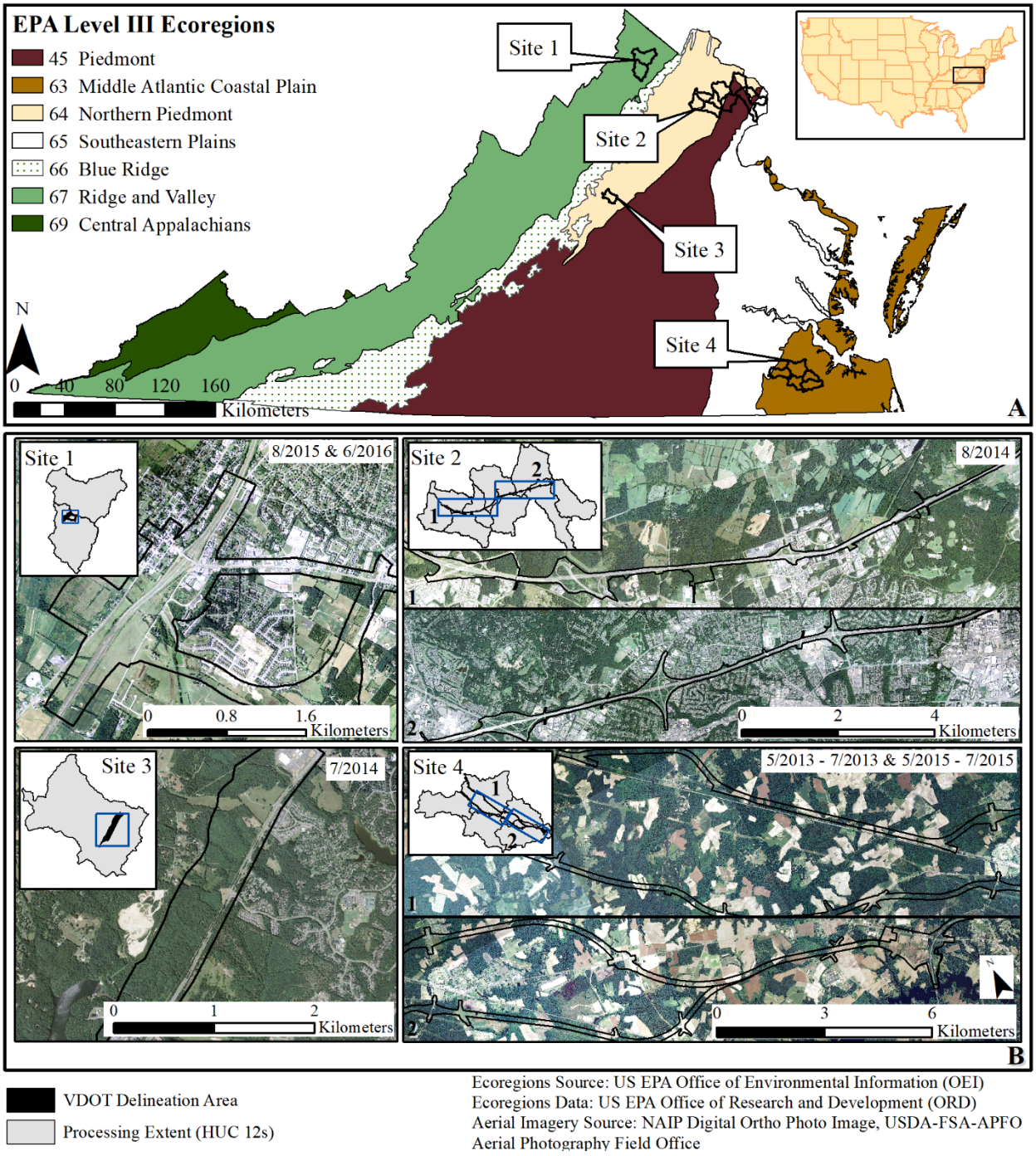

Figure 1. (a) Study site locations, outlined by watershed(s) used as site processing extent, spanning five of the seven ecoregions of Virginia, and (b) areas of each VDOT delineation site with orthoimagery corresponding to the time frame in which VDOT delineations were performed (M/YYYY).

Table 1. Conditions of the processing extent and VDOT delineation area for each study site; upper portion describes conditions of the processing extent and lower portion describes conditions of the VDOT delineation area.

	Site 1	Site 2	Site 3	Site 4
Processing Extent (HUC 12s) (km^2)	273	1208	65	547
LiDAR DEM Resolution (m)	1.00	1.50	0.76	0.76
HUC12 Max. Elevation (m)	458	417	223	37
HUC12 Min. Elevation (m)	140	0	96	0
HUC12 Mean Slope (%)	9.5	7.0	12.6	3.7
VDOT Delineation Total Area (km^2)	2.98	7.87	1.82	12.17
VDOT Delineation Max. Elevation (m)	241	147	178	34
VDOT Delineation Min. Elevation (m)	210	47	101	3
VDOT Delineation Mean Slope (%)	7.2	9.4	14.7	3.2
VDOT Wetland to Non-Wetland Ratio	0.02	0.02	0.02	0.42

3. Input Data

Freely available LiDAR elevation data, land cover data, national-scale hydrography data, national-scale soil data, and VDOT wetland delineations were used as inputs to the wetland identification model.

3.1. LiDAR Elevation Data

LiDAR-derived elevation data used in this study were provided by the Virginia Information Technologies Agency (VITA) in raster format (http://vgin.maps.arcgis.com). VITA LiDAR data products were freely available and included hydro-flattened, bare-earth DEMs. The LiDAR DEMs used in this study were collected and processed between 2010 and 2015 and have horizontal resolutions ranging from 0.76 m to 1.5 m. Tiles with different resolutions were merged and resampled to the coarsest resolution using the bilinear resampling method in ArcGIS, following the approach previously done by Ågren et al. (2014). Site 2 was unique in that LiDAR data were unavailable for approximately 230 km^2 (23%) of the processing extent and

0.8 km^2 (12%) of the VDOT delineation area. To fill the missing areas, 3 m elevation data from the National Elevation Dataset were used (https://viewer.nationalmap.gov) and resampled to 1.5 m to match the dominating LiDAR data. While resampling to finer resolutions is not ideal, maintaining consistency in the application of highest resolution LiDAR data across all study sites was prioritized over the error introduced in the relatively small portion of the processing extent, and even smaller portion of the delineation area.

3.2. Land Cover Data

Land cover data were used for post classification filtering. Land cover data used in this study were provided by VITA in raster format (http://vgin.maps.arcgis.com). VITA land cover data were derived from the Virginia Base Mapping Program 4 band orthophotography, collected between 2011 and 2014. These data provided 12 land cover classifications with 85-95% accuracy and have a horizontal resolution of 1 m (WorldView Solutions Inc., 2016).

3.3. National-Scale Datasets

National-scale soil and hydrography data were incorporated in the classification as ancillary data. Soil data used in this study were obtained from the Soil Survey Geographic database (SSURGO) and distributed by the Natural Resources Conservation Service's Web Soil Survey in polygon vector format (https://websoilsurvey.sc.egov.usda.gov). The SSURGO hydric rating, depth to water table, hydrologic soil group, surface texture, and soil drainage class were used as indicators of saturated conditions. According to the Soil Survey Staff (2017), the hydric rating attribute indicates the percentage of a map unit that meets the criteria for hydric soils. Hydric soils are characteristic of wetlands and are defined as soil that is formed under conditions of saturation, flooding, or ponding long enough during the growing season to develop anaerobic conditions in the upper horizon (Federal Register, 1994). The surface texture attribute describes

the representative texture class according to percentage of sand, silt, and clay in the fraction of the soil that is less than 2 mm in diameter. The drainage class attribute identifies the natural drainage conditions of the soil and refers to the frequency of wet periods without considering alterations of the water regime by human activities, unless they have significantly changed the morphology of the soil. The hydrologic soils group assignment is based on estimates of the rate of water infiltration when the soils are not protected by vegetation, are thoroughly wet, and receive precipitation from long-duration storms. The depth to water table attribute indicates the representative depth to the saturated zone in the soil.

Hydrography data used in this study were provided by the National Hydrography Dataset (NHD) in polygon vector format (https://viewer.nationalmap.gov). NHD HUC 12 watersheds intersected by the limits of VDOT delineations were combined to be used as the processing extent for each study site in order to encompass the hydrologically connected area around VDOT delineations. NHD streams and waterbodies within these processing extents were also used.

3.4. VDOT Wetland Delineations

Wetland delineations for each site were provided by VDOT and were used to create training and testing datasets. The VDOT delineations in Site 2, Site 3, and Site 4 were jurisdictionally confirmed by the USACE, and all study sites were produced through field surveys conducted by professional wetland scientists. For these reasons, the VDOT delineations were considered to be ground truth for the purpose of training and testing the wetland identification model. VDOT delineations were provided in polygon vector format and included both wetlands and streambeds. Both were included in subsequent processing because both are considered waters of the state and therefore must be delineated during the wetland permitting process. Although the delineations were categorized by wetland type by VDOT analysts, all

areas were merged into a single “wetland” category before application in this study. Additionally, limits of delineations were used to identify true non-wetland areas.

4. Methods

The workflow followed to implement the wetland identification approach consisted of three main parts: preprocessing, supervised classification, and post processing (Figure 2). The workflow was implemented in ArcGIS 10.4 and the ModelBuilder tool was used to automate processes that did not require user intervention. Outputs of the workflow were model predictions and confusion matrices used to assess the accuracy of those predictions. Components of the workflow are described in more detail in the following sections.

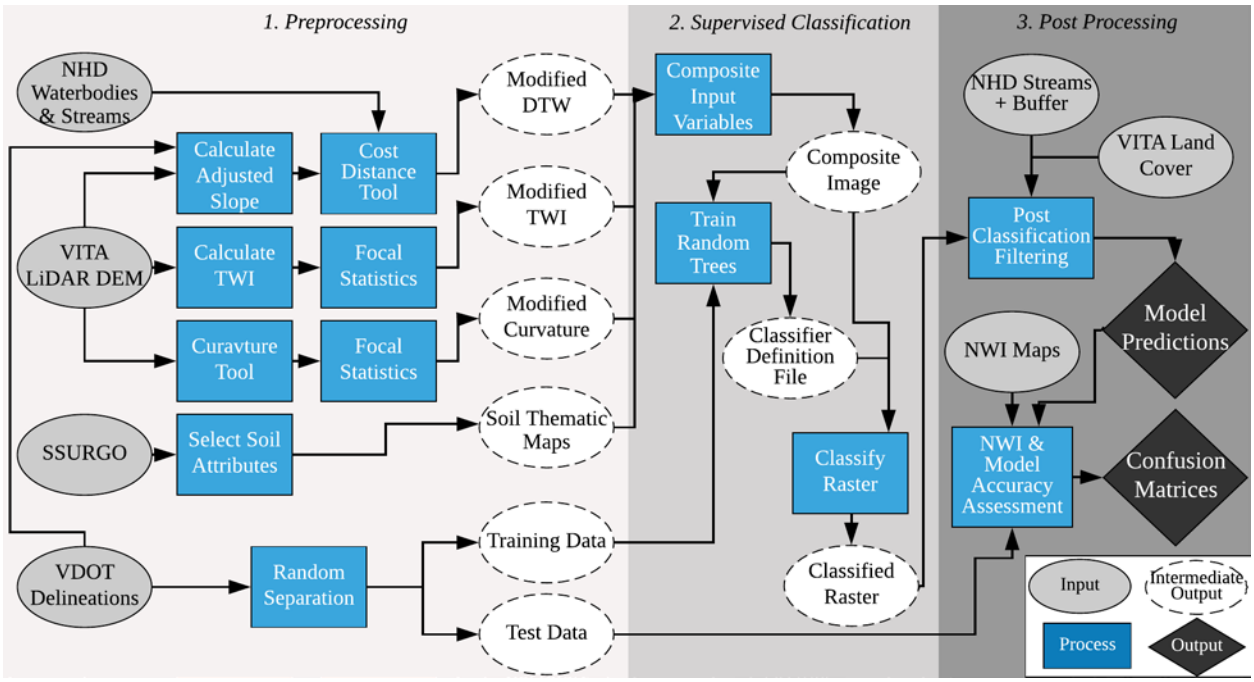

Figure 2. Workflow followed to implement the wetland identification approach as an ArcGIS model consisting of preprocessing, supervised classification, and post processing phases to create model predictions and confusion matrices used for accuracy assessment.

4.1. Preprocessing

The preprocessing phase consisted of a combination of automated and semi-automated processes that required user intervention. Preprocessing steps not explicitly shown in Figure 2 include projection of input data to the appropriate North or South Virginia State Plane coordinate system, clipping data to the HUC 12 processing extent, rasterizing input data originally in polygon vector format by using the site LiDAR data as the pixel size constraint, and filling sinks within the LiDAR DEM. Rasterizing the polygon vector layers mapped at larger scales assumes that the information provided at the original scale (ranging from 1:24,000 to 1:12,000) is true for each pixel of the output grid (ranging from 0.76 to 1.52 m). The LiDAR DEM was filled using the depression filling algorithm of Planchon and Darboux (2002) that is implemented in ArcGIS. Intermediate outputs created by the preprocessing phase were calibrated input variables, training data, and testing data.

4.1.1. Modified Input Variable Creation

Input variables included the modified TWI, modified curvature, modified DTW, and selected soil thematic maps. Input variables were modified based on site characteristics and information provided by VDOT delineations in order to produce distinct wetland and non-wetland signatures, and user intervention was necessary to execute some of the calibration processes. Summarized modification parameters for topographic indices and information relevant to their calculation are shown in Table 2 and the methods used to calculate these parameters are described in the following sections.

Table 2. Modification parameters for topographic indices, and soil thematic maps determined to be relevant for each study site. Site characteristics relevant to the calculation of modification parameters are italicized and inclusion of a soil layer is indicated by an “X.”

	Site 1	Site 2	Site 3	Site 4
LiDAR DEM Resolution (m)	*1.0*	*1.52*	*0.76*	*0.76*
TWI Focal Statistic Window size (# pixels)	5	3	7	7
Curvature Focal Statistic Window size (# pixels)	5	3	7	7
TWI Focal Statistic Type	Median	Median	Median	Median
Curvature Focal Statistic Type	Mean	Mean	Mean	Mean
Maximum Underlying Wetland Slope Value (m/m)	*0.751*	*1.134*	*1.652*	*1.403*
Representative Wetland Slope (m/m)	*0.088*	*0.168*	*0.41*	*0.115*
DTW γ	11.42	5.95	2.44	8.70
DTW β	2	2	2	2
Hydrologic Soil Group	X	X		
Depth to Water Table	X			X
Surface Texture	X	X		
Hydric Rating		X		X
Soil Drainage Class		X		

4.1.1.1. TWI Modifications

The modified TWI grid is based on the TWI as defined in Equation (1). The TWI was created in ArcGIS as a Map Algebra expression. The inputs required for this calculation were a flow accumulation grid, to represent the α term, and a slope grid, to represent the tan(β) term, both derived from the filled LiDAR DEM. The D8 method (Jenson & Domingue, 1988) was used to generate flow direction and flow accumulation grids. A slope grid was generated with the ArcGIS slope tool, calculated as the steepest downhill descent from each pixel in units of m/m (Burrough & McDonell, 1998). A constant equal to 1 was added to flow accumulation grids so that every pixel received flow from itself as well as upslope pixels to avoid undefined TWI values, and a constant equal to 0.0001 (m/m) was added to slope grids to avoid dividing by zero. An example of the resulting TWI grid, overlaid with VDOT wetland areas, for a portion of Site 1 is shown in Figure 4 (panel A1). This TWI grid models the presence wetter areas (high TWI values) in locations of high flow accumulation and low slopes, and drier areas (low TWI values) in locations of steep slopes and less flow accumulation. Larger clusters of relatively high TWI values align with the VDOT delineated wetlands, however there is also a scattering of high TWI

values outside of these wetland boundaries, corroborating the challenges of high-resolution TWIs previously described in the literature (e.g., Ågren et al., 2014; Sørensen & Seibert, 2007). Some researchers recommend deriving TWIs from coarser DEMs (e.g., Ågren et al., 2014), but doing so would sacrifice the rich detail provided by LiDAR DEMs that may be needed to precisely model shape and size of environmental planning-scale wetlands.

Although these scatterings of relatively high TWI values may be modelling true micro-topographic features, their location outside of the field-mapped wetlands suggest these flow channels are not large enough to result in saturated conditions. Rather than lose hydrologic detail of the LiDAR data by resampling, anomalous local variations were smoothed by applying a low-pass filter over a moving NxN window to create the modified TWI. Applying a low-pass filter searches over a user-defined window in which every pixel is replaced with the statistical value from the surrounding pixels within the NxN window, as done by Ali et al. (2014), Buchanan et al. (2014), and Lanni et al. (2011). The window size for the smoothing operation is significant in that it is usually set with consideration of the average size of the feature of interest (Sangireddy et al., 2016). In this study we estimated that areas of interest must be at least 5 m in width based on the size of VDOT delineated wetlands. Therefore, window sizes were set to smooth over a total area of approximately 25 m^2 (5 m x 5 m) with this window size varying slightly across study sites depending on pixel length of the LiDAR data. Additionally, a median filter was chosen to perform smoothing rather than the mean filter. Visual assessment of both statistic types showed that the median filter better retained VDOT wetland edge features while removing scattered high TWI values outside of these boundaries. TWI smoothing was implemented in the ArcGIS model using the Focal Statistics tool. Window sizes used to calculate the modified TWI grid for each site are shown in Table 2, and an example of applying this modification for a

portion of Site 1 is shown in Figure 4, panel A2. Compared to the unmodified TWI (panel A1), this scene shows the larger cluster of relatively high TWI values within VDOT delineated wetlands were maintained, but the discrete, small flow channels outside of the true wetland boundaries have been smoothed via replacement of these pixels with relatively lower TWI values.

4.1.1.2. Curvature Modifications

Curvature grids, as defined by Moore et al. (1991) were created from the filled LiDAR DEM using the ArcGIS Curvature tool. Curvature has been shown to be a key component in the process of identifying likely channelized pixels indicating flow convergence (Ågren et al., 2014; Hogg & Todd, 2007; Kloiber et al., 2015; Millard & Richardson, 2013; Sangireddy et al., 2016). It was anticipated that the high resolution of the LiDAR-derived curvature grids would assist in separating small differences in concavity between nearly flat roadways and shallow local depressions. However, visual assessment of the LiDAR-derived curvature grids showed a similar issue of topographic noise as seen in the TWI, in that micro-topographic channels were also mapped. An example of the output curvature grid for a portion of Site 1 is shown in Figure 4, panel B1. This image shows negative and zero curvature values within VDOT wetland extents, which correspond to concave and flat areas, respectively.

Similar to modified TWI creation, the curvature was modified by applying a statistical smoothing process to curvature grids, following the approach of Sangireddy et al. (2016). When choosing the window size for this calculation, the assumption of the average size of features of interest was kept consistent with that of the TWI (i.e., at least 5 m in width). In this case a mean filter was chosen to smooth the curvature data rather than a median filter due to a visual inspection and perceived improvement in VDOT wetland edge retention resulting from the mean

smoothing. The modified curvature grid was created by applying the ArcGIS Focal Statistics tool. Window sizes used to calculate the modified curvature grid for each site are shown in Table 2 and an example of applying this modification for a portion of Site 1 is shown in Figure 4, panel B2. In this image one can see that the modified curvature grid has a smoother appearance but maintains significant areas of concavity.

4.1.1.3. DTW Modifications

The modified DTW grid is based on the DTW as defined in Equation (2). This iterative function finds the cumulative slope value along the least downward slope (i.e., “cost”) path to the nearest surface water (i.e., “source”) pixel with which it is most likely to be hydrologically connected (Murphy et al., 2009). To calculate the DTW, two input grids are required: a grid of slope values and a grid of areas of open water (Murphy et al., 2009). In this study, slope grids were derived from the filled LiDAR DEM using the ArcGIS slope function, as done in the original formulation of the DTW model (e.g., Murphy et al., 2007, 2009, 2011), and the source grids were created from rasterized NHD waterbodies and streams. While the publicly available NHD was chosen in this study to maintain consistency between the four sites, there are alternatives for researchers without publicly available open water data. The source grid can also be generated directly from elevation data by deriving streams based on a designated flow accumulation threshold (Murphy et al., 2009) or use of open source channel extraction software, such as GeoNet (Sangireddy et al., 2016). The effects and limitations of using the relatively coarsely mapped NHD as the source grid for the DTW are discussed in section 5.2. of this paper. The ArcGIS Cost Distance tool was used to evaluate Equation (2) within the model using the slope and NHD source grids as inputs. It was also necessary to add a small constant (0.0001 m/m) to all pixels in the slope grid to differentiate from source grid pixels, which are assigned a

value of zero for the calculation. An example of the resulting DTW grid for a portion of Site 1 is shown in panel C1 of Figure 4. As expected, low wetness (high DTW values) occurred in areas further and higher along the terrain from surface water, and high wetness (low DTW values) occurred in areas of low slopes that are closer to surface water. While wetted areas calculated by the DTW correspond to VDOT delineated wetlands, the transition from wet to dryer areas is gradual. We found this to result in lower non-wetland accuracy, or an overestimation of wetlands, when using only the original DTW formulation to identify wetland areas.

Therefore, a modified DTW was created to accelerate the gradual transition from wetlands to uplands in an effort to better distinguish wet from dry locations. The method outlined above was used to calculate the modified DTW, except that the input slope grid was replaced with an adjusted slope grid, defined as,

$$Y = \gamma * X^{\beta}, \qquad (3)$$

where X is the slope (with a small constant added to all values, as described earlier), and γ and β are calculated slope adjustment parameters. This adjustment to the slope values was intended to create two distinct ranges of low cost areas, where wetlands are likely to exist, and high cost areas, where wetlands are unlikely to exist, based on the observed distribution of wetland slope values in each site. The γ parameter allows users to control the cutoff between the low and high cost slope values, which corresponds to a designated representative wetland slope value. The β parameter allows users to control the rate of increase in cost as the slopes increase throughout the site. In this study, β was set to a value of 2 for all sites while γ was individually calibrated. We hypothesized that setting the wetland slope value equal to the 95th percentile of all underlying VDOT wetland slope values would result in a γ parameter that further flattens the terrain (i.e., reduces the cost) where most wetlands exist, disregarding assumed outliers, and further

steepening the terrain (i.e., increasing the cost) elsewhere. Representative slope values were calculated by extracting slope values within VDOT wetland boundaries, and calculating the 95th percentile of each array with the Numpy Python library. Figure 3 shows an example of this adjusted slope calculation and describes the effect of this adjustment for Site 1, where the 95th percentile was 0.088 m/m, which corresponded to a γ value of 11.42.

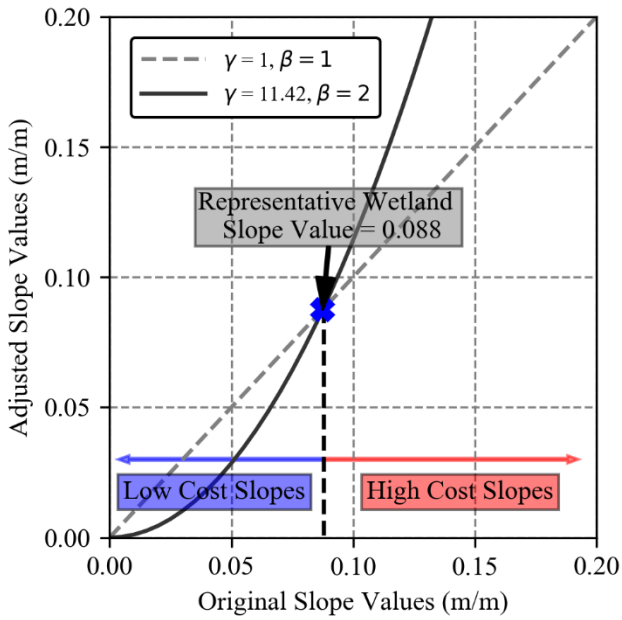

Figure 3. Example calculation of the adjusted slope grid (solid line) for Site 1 where the β was set to a value of 2 and γ was calculated to be 11.42, corresponding to a representative slope value taken to be the 95th percentile of all underlying wetland slopes. These adjustments decrease slopes that are originally below 0.088 and increase slopes that are originally above 0.088, relative to a slope grid (dashed line) where γ and β are both equal to 1.

Note: Although maximum wetland slope value in Site 1 was 0.751 m/m, a smaller range of values is shown here for clarity.

With the adjustments to the slope grid applied, Equation (2) becomes

$$Modified\ DTW\ (m) = \left[\sum \gamma \left(\frac{dz_i}{dx_i}\right)^2 a\right] * \ x_c, \qquad (4)$$

where γ and β=2 are introduced. Slope adjustment parameters and relevant site characteristics used to calculate these parameters are shown for each site in Table 2. An example of the effect of modifying the DTW in Site 1 using this calculation is shown in panel C2 of Figure 4. In this figure, the modified DTW (C2) shows relatively wetter areas within VDOT wetland boundaries and an accelerated increase to drier values moving away from VDOT wetlands, compared to the original DTW (C1).

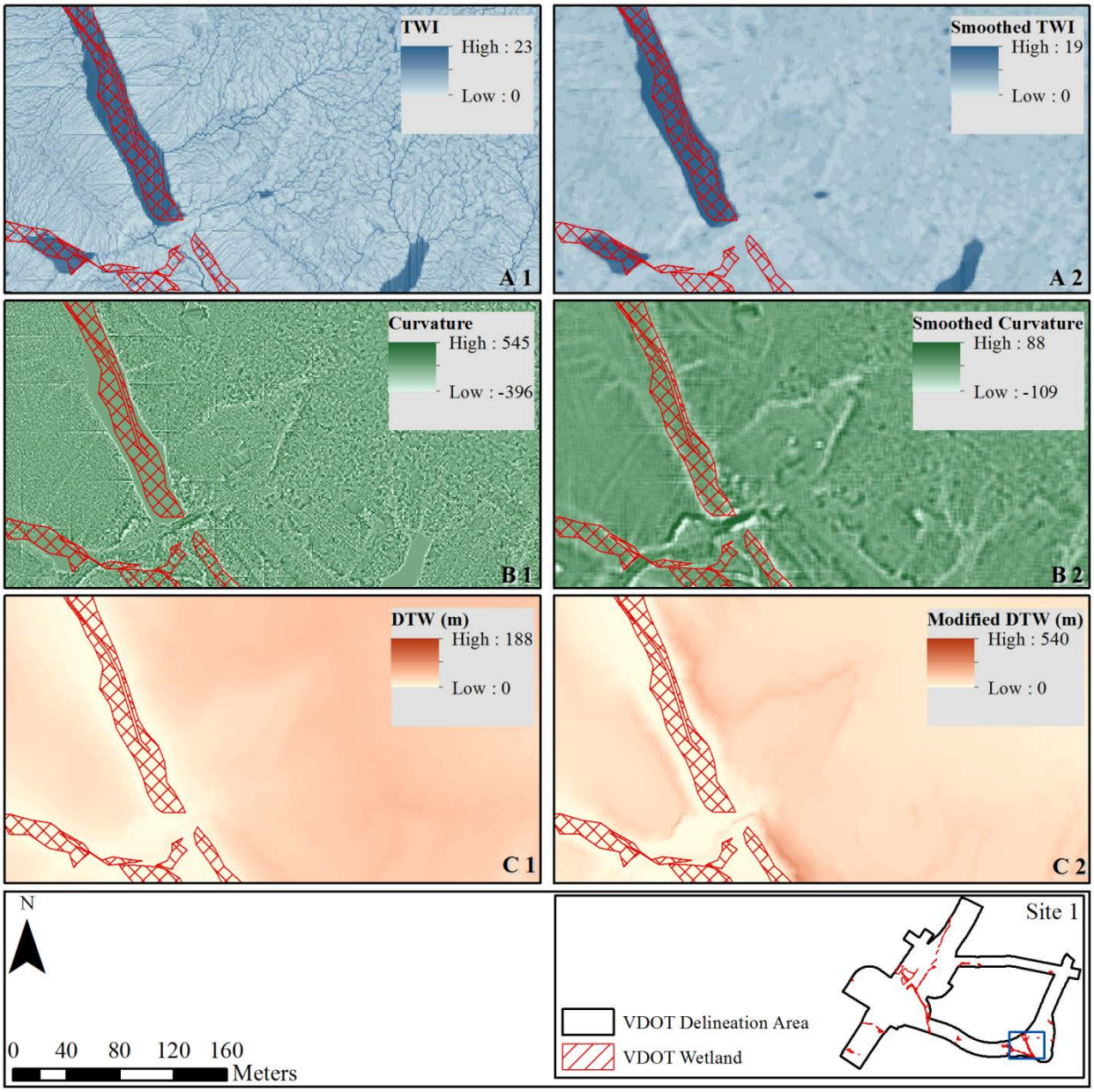

Figure 4. Topographic input variables in Site 1, original TWI (A1), curvature (B1), and DTW (C1), compared to modified versions each variable, shown in A2, B2, and C2, respectively. Modification parameters used to calculate the modified topographic indices in Site 1 are shown in Table 2.

Note: Panels A1 and B1 highlight anomalies in elevation data that are likely artifacts of LiDAR tile merging during original processing of raw data.

4.1.1.4. Soil Thematic Maps

The final input variables created in the preprocessing phase were soil thematic maps. Soil thematic maps were created from the extensive SSURGO database using the Soil Data Viewer ArcMap extension (NRCS, 2015). Although the Soil Data Viewer creates soil thematic maps automatically, combinations of soil layers were manually chosen for each site based on correspondence of the soil data to the current physical landscape. This correspondence was assessed by visual comparison to VDOT delineations and VITA land cover data. Soil layers that appeared too coarse, i.e. generally did not vary enough within the VDOT delineated area to describe features of interest, were not selected.

4.1.2. Training and Testing Data

An automated process was used to randomly designate 10% of VDOT delineation area to train the classifier and reserve the remaining 90% to test the classification results. It has been noted that statistical classifiers and machine learning algorithms may be sensitive to imbalanced training data or cases where rare classes are being classified (such as most cases of wetland identification), and the sensitivity of RF, specifically, to training class proportions was investigated by Millard and Richardson (2015). The researchers found that when training samples were disproportionately higher or lower than the true distribution of that feature, the final classification over or under predicted that class, respectively. They concluded that using a sampling strategy that ensures representative class proportions, and minimal spatial autocorrelation, minimized proportion-error in their results (Millard & Richardson, 2015). In this study we took into account the findings of Millard and Richardson (2015) when designing the methodology to randomly separate VDOT delineations into training and testing data. This process consisted of 4 steps: random point creation, point buffering, value extraction, and

training data separation (Figure 5). A stratified random sampling method was used in the first step to distribute a designated number training sample locations proportionately between wetland and non-wetland areas (panel A). These randomly generated points were then buffered to create circle polygons with an area of approximately 100 m^2 each (panel B). In the value extraction step (panel C), training data, composed of approximately 10% of the delineated area and with representative class proportions, were produced by rasterizing the buffered polygons with pixel values extracted from VDOT delineations to correct cases of buffered polygons falling into both wetland and non-wetland classes. The testing data were created by separating the training data from the VDOT delineations, leaving approximately 90% of the delineated area to be used for accuracy assessment (panel D). Statistics describing the training and testing datasets for each site are found in Table 3.

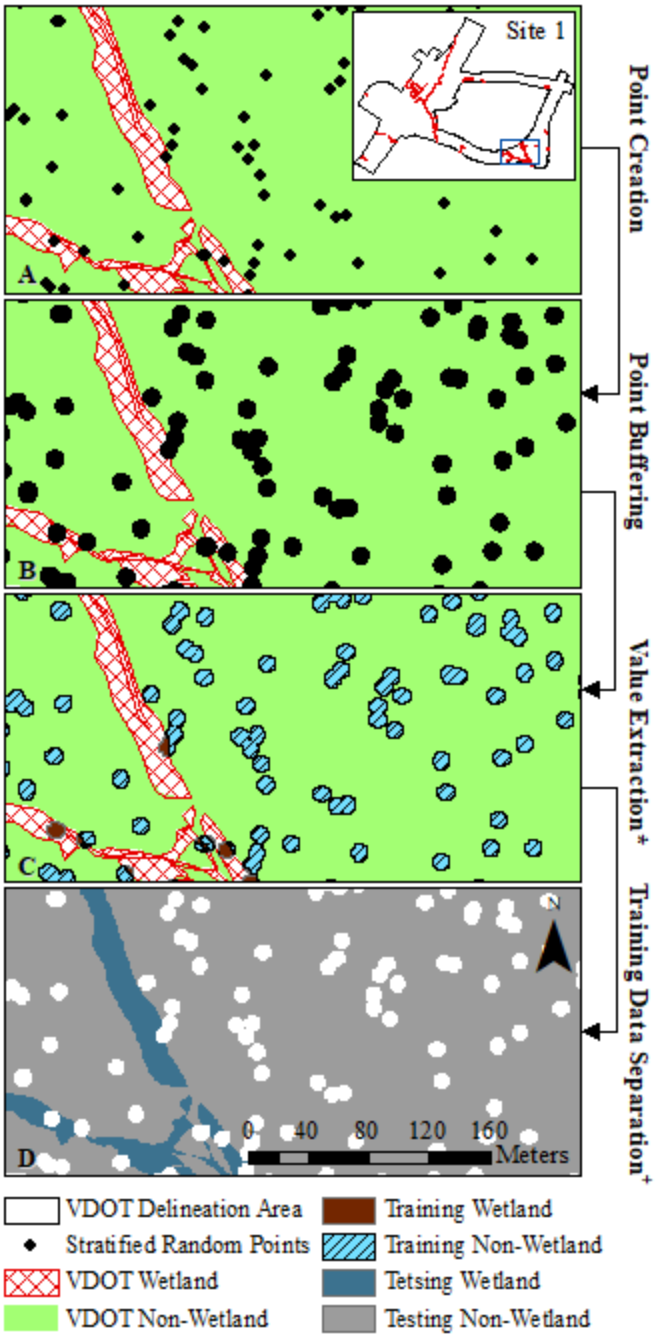

Figure 5. Example of the process, shown for Site 1, used to randomly separate VDOT delineations into training and testing datasets, consisting of four steps: (A) point creation, (B) point buffering, (C) value extraction, and (D) training data separation. Asterisk indicates the phase in which training data are created and superscript "+" indicates the phase in which testing data are created.

Table 3. Statistics describing the training and testing data for each site.

	Site 1	Site 2	Site 3	Site 4
Training Wetlands (km^2)	0.007	0.015	0.003	0.347
Training Non-Wetlands (km^2)	0.271	0.745	0.172	0.816
Training Wetland to Non-Wetland Ratio	0.03	0.02	0.02	0.43
Training Area to VDOT Delineation Area Ratio	0.09	0.09	0.10	0.10
Testing Area (VDOT Delineation - Training Area) (km^2)	2.71	7.11	1.65	11.00

4.2. Supervised Classification

In the first phase of the supervised classification portion of the workflow, the input variables created during preprocessing were combined into a multidimensional, composite image where each dimension stores an independent input variable. Wetland and non-wetland signatures were extracted from this composite image and used to perform the supervised classification. RF classification was chosen as the supervised classification algorithm for its noted advantages in similar studies, as described previously (e.g., Duro et al., 2012; Miao et al., 2012; Millard & Richardson, 2013; Rodriguez-Galiano et al., 2012). According to Breiman (2001), RF is an ensemble classifier that produces many Classification and Regression-like trees where each tree is generated from different bootstrapped samples of training data, and input variables are randomly selected for generating trees. This algorithm also produces variable importance information, which measures the mean decrease in accuracy when a variable is not used in generating a tree.

The RF classification was executed in ArcGIS with the Train Random Trees and Classify Raster tools (ESRI, 2016). The Train Random Trees tool utilizes the OpenCV implementation of the RF algorithm (Bradski, G., 2000). Using Train Random Trees, the training data were used to extract class signatures from the dimensions (i.e., input variables) of the composite image,

creating an ESRI Classifier Definition file with variable importance measures. The Classifier Definition file was subsequently used to classify the remainder of the composite image. The result of these operations is a grid where each pixel has been classified as wetland or non-wetland. As the focus of this study was to analyze the response of classification models to input data, the RF parameters were not varied or calibrated to study sites. For this reason, the default values of maximum number of trees, maximum tree depth, and maximum numbers of samples per class were held constant at the recommended default values of 50, 30, and 1000, respectively. Future work should perform a sensitivity analysis to test the effect of adjusting these parameters.

4.3. Post Processing

The first phase of post processing was post classification filtering. The objective of the post classification filtering was to account for areas that may be susceptible to water accumulation due to its local topography, but cannot be wetland areas due to impervious land cover. The post classification filtering algorithm first used a logical statement to determine if a classified wetland pixel overlaps VITA land cover designated as impervious. If this was false, the pixel classification was unchanged. If this was true, a second logical statement was used to account for cases where wetlands may exist under bridges by determining if classified wetland pixels are within 30 m of NHD streams. The 30-m buffer distance was an estimated value based on visual inspection, and more precise measurements would increase effectiveness of post classification filtering. If this second statement was false, the pixel was reclassified as non-wetland, otherwise it was left unchanged. This process produced the model predictions.

The second phase of post processing was accuracy assessment. The model predictions and NWI map for the study area were assessed for accuracy in terms of agreement with the test

dataset. Accuracy assessments were evaluated with confusion matrices, which summarized the areas of wetland agreement, non-wetland agreement, false negative predictions (cases where true wetland areas were predicted to be non-wetland), and false positive predictions (cases where true non-wetland areas were predicted to be wetland). Confusion matrices for the model predictions and NWI maps were used to calculate wetland accuracy, non-wetland accuracy, and overall accuracy using Equations 5-7,

$$Wetland\ Accuracy = \frac{wetland\ agreement\ (km^2)}{\sum test\ (actual)\ wetland\ (km^2)} \tag{5}$$

$$NonWetland\ Accuracy = \frac{nonwetland\ agreement\ (km^2)}{\sum test\ (actual)\ nonwetland\ (km^2)} \tag{6}$$

$$Overall\ Accuracy = \frac{wetland\ agreement\ (km^2) + nonwetland\ agreement\ (km^2)}{\sum test\ (actual)\ area\ (km^2)}. \tag{7}$$

The use of these metrics to assess wetland classifications is common in literature (e.g., Ågren et al., 2014; Millard & Richardson, 2013).

5. Results and Discussion

5.1. Highest Performing Models

To determine the highest performing models, classifications varying only topographic inputs were first performed and assessed, and the input data that resulted in highest overall accuracy were combined with relevant soil layers, if any. In the coming sections, the following results are discussed: (1) scenes for each site comparing highest performing models and their level of agreement with VDOT delineations, compared to NWI maps, (2) variable importance of highest performing input data, and (3) the accuracy assessment of highest performing models compared to the NWI. The input data used to produce the best performing models and the importance of these inputs according to the ESRI Classifier Definition file are listed in Table 4. Although

accuracy assessments for each site only extend to testing dataset limits, scenes depicting predictions and VDOT delineations prior to the separation process are shown for clarity.

Table 4. Input data that produced the highest performing wetland identification model in each site, in terms of overall accuracy, as well as variable importance and rank of each input variable according to the ESRI Classifier Definition file. Topographic inputs with an asterisk indicate the application of modifications using parameters from Table 2.

	Input 1	Input 2	Input 3	Input 4	Input 5	Input 6	Input 7	Input 8
Site 1	TWI*	Curvature*	DTW*	HSG[1]	Depth to WT[2]	ST[3]	-	-
VI[+]	0.087	0.111	0.333	0.131	0.182	0.156	-	-
Rank	6	5	1	4	2	3	-	-
Site 2	TWI*	Curvature*	DTW	HSG[1]	-	ST[3]	HR[4]	DC[5]
VI[+]	0.078	0.107	0.156	0.208		0.126	0.177	0.150
Rank	7	6	3	1		5	2	4
Site 3	TWI*	Curvature*	DTW*		-	-	-	-
VI[+]	0.158	0.325	0.516					
Rank	3	2	1		-	-	-	-
Site 4	TWI*	Curvature*	DTW*		Depth to WT[2]		HR[4]	
VI[+]	0.076	0.114	0.215		0.338		0.257	
Rank	5	4	3		1		2	

[+]Variable Importance; [1]Hydrologic soil group; [2]Depth to water table; [3]Surface texture; [4]Hydric rating; [5]Drainage Class

5.1.1. Site 1 Results

Wetland predictions and NWI data for Site 1 are shown in Figure 6. Both of the NWI scenes (A1 and B1) exemplify the tendency of the NWI to underestimate the size of VDOT delineated wetlands by mapping wetlands primarily along streams. While the narrow NWI wetlands precisely map the wetland areas that are in agreement with VDOT delineations, the NWI fails to match the contours or the size of larger wetland zones. These larger wetland zones were more fully mapped by wetland predictions produced by the model (A2 and B2). However the model also produced relatively higher overestimation of wetlands. Overestimation of

wetlands is especially prevalent in location 1. Underlying input variables indicated that overestimation here was due to a depression that was filled to become a large, zero-slope area. This flat zone resulted in a corresponding generalized area of high wetness values in the modified TWI and modified DTW. In addition, the surface texture input indicated that silty clay loam, which have relatively slow infiltration rates (~0.5 cm/h) (Soil Survey Staff, 2017), was also present in this overestimated area, likely contributing to the wetland predictions here. It is possible that the results in this site could be improved by using an alternative to the pit filling (i.e., ArcGIS Fill) algorithm to avoid creation of generalized, flat areas, more severe adjustments to the slope grid for the modified DTW, or higher resolution SSURGO data. Panel B2 shows more precise model wetland predictions, represented by conformity of predicted wetlands to the curvature of VDOT delineated wetlands. This panel encompasses the scene in Figure 4 (C2) where the modification to the DTW was shown to more precisely map wetland areas. For that reason, we attribute the relatively precise mapping of wetlands in B2 in part to the modifications used for the DTW in this site. Location 2 shows one small wetland that was undetected by the model. This may indicate a wetland formed due to conditions more strongly driven by vegetation rather than topography or proximity to surface water.

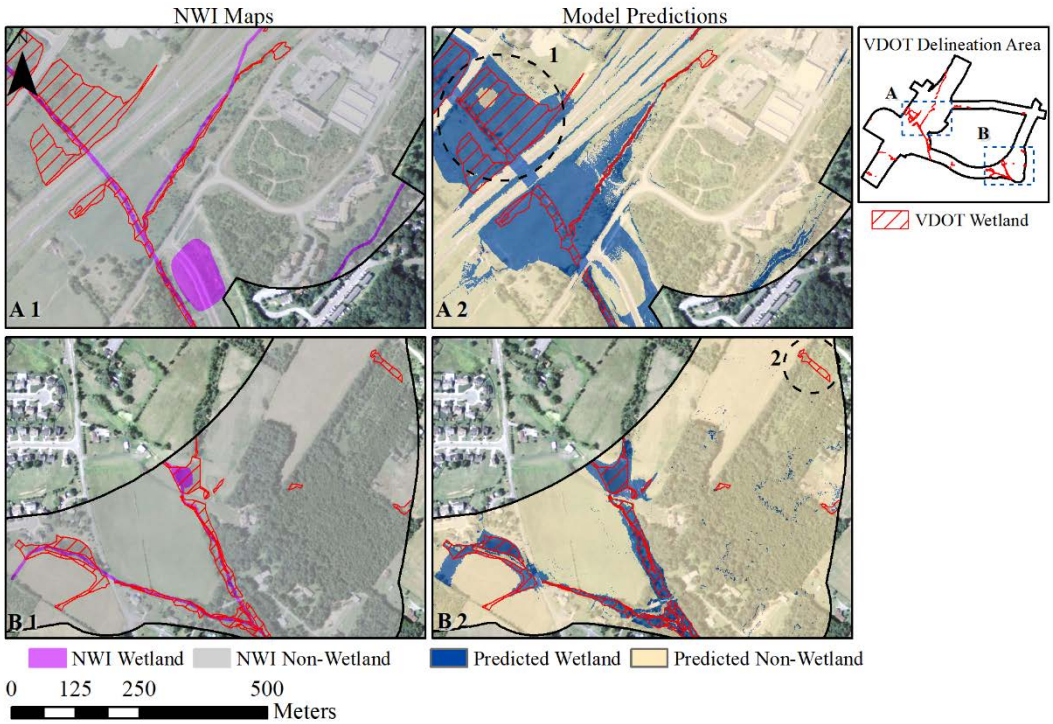

Figure 6. Examples of NWI maps (A1 and B1) and model predictions (A2 and B2) for Site 1, both compared to VDOT delineations.

5.1.2. Site 2 Results

Two scenes of the model predictions and NWI maps for Site 2 are shown in Figure 7. In panels A1 and A2, the NWI dataset and model predictions both show similar overestimation of wetland area, although the model resulted in higher overestimation. The false positive predictions in this area were due to flow convergence indicated by the topographic inputs, and the presence of hydric soils indicated by the SSURGO data. Also, many false positive predictions in this site were in locations overlapping road features (e.g., location 1). This may indicate a need for alternate modifications to topographic inputs, especially curvature, to better differentiate channelized areas due road features from channelized areas that are wetlands, as proposed by Sangireddy et al. (2016). Panel B1 shows another example of NWI wetland delineations following along streams, but failing to capture the extents of larger wetland zones.

For this same area, the model predicted wetlands further from the streambeds due to the gradual slopes surrounding them and better encompassed VDOT delineated wetlands (locations 2 and 3).

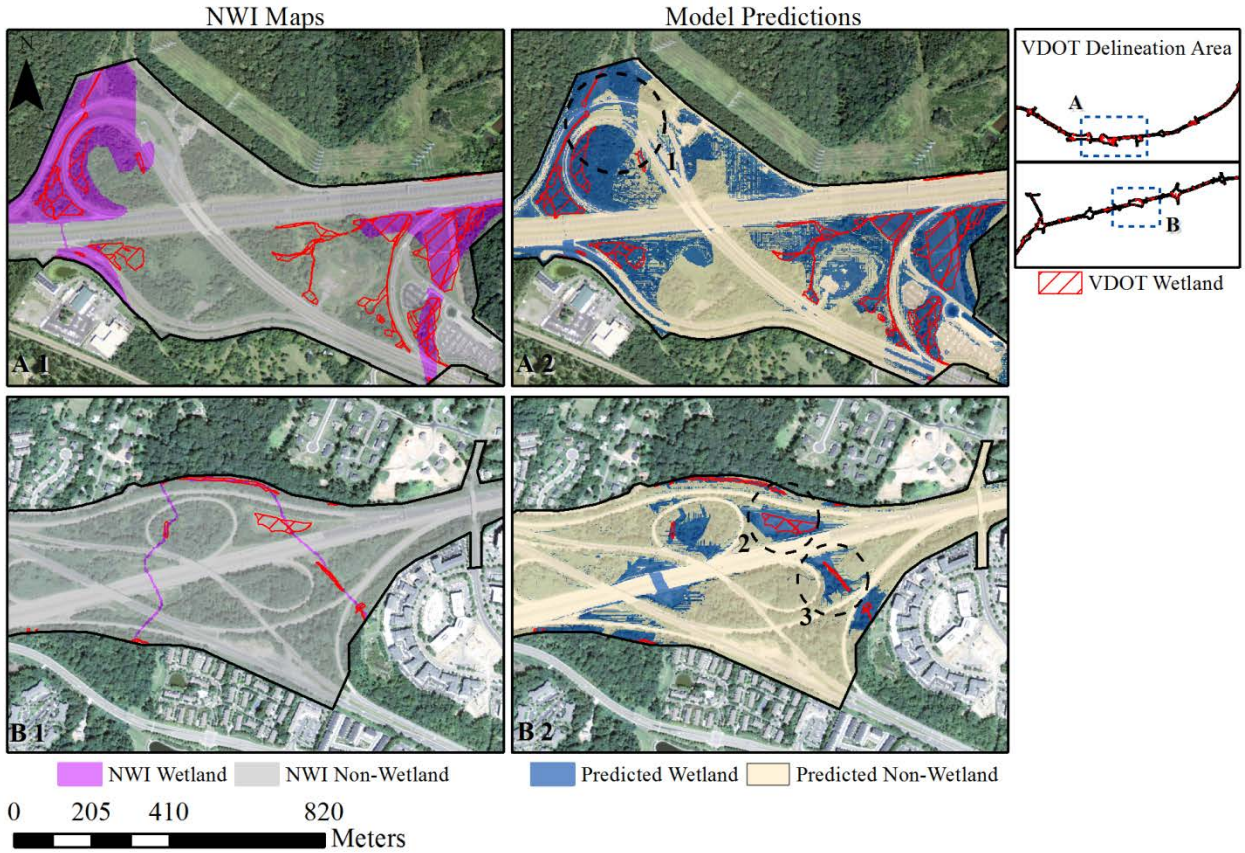

Figure 7. Examples of NWI maps (A1 and B1) and model predictions (A2 and B2) for Site 2, both compared to VDOT delineations.

5.1.3. Site 3 Results

Examples of model predictions and NWI data for Site 3 are shown in Figure 8. As seen in Table 4, Site 3 was unique in that no soil layers were included in the best performing model. Visual assessment of relevant soil layers in this area showed that the SSURGO data did not vary in a way that effectively differentiated between features of interest. Site 3 was also unique for its wetlands which were typically narrow and located along small flow channels, rather than in larger wetland zones. The NWI data shown either do not conform to the bends along the length

of wetlands (A1), or failed to map a number of wetlands in these channelized areas (B1). The model predicted a larger portion of the VDOT delineated wetlands in both scenes, however the wetland predictions often extended too far on either side of the narrow wetlands (A2). Location 1 shows another example of a local depression filled to become a generalized, flat area, resulting in an overestimation due to the modified TWI and modified DTW indices. Additionally, both scenes A2 and B2 show that the model detected road edges and road medians as wetland areas. This is a shortcoming of the model that was observed in other sites, such as Site 2, and indicates a need for further modification to topographic indices.

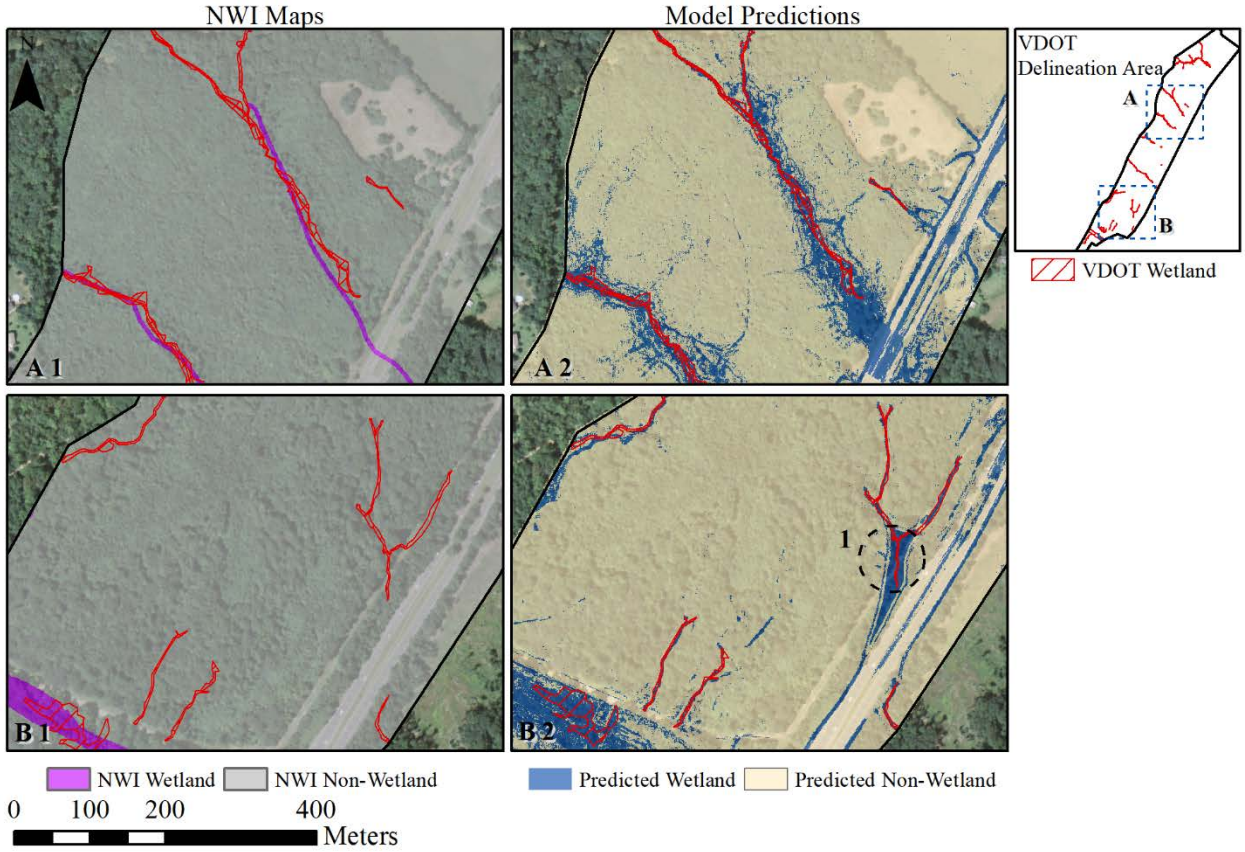

Figure 8. Examples of NWI maps (A1 and B1) and model predictions (A2 and B2) for Site 3, both compared to VDOT delineations.

5.1.4. Site 4 Results

Figure 9 shows three scenes from the NWI maps and model predictions for Site 4, which was the largest site studied. Site 4 was also unique for having the largest distribution of VDOT delineated wetlands, covering more than 40% of the surveyed area, as well as the mildest average slope (see Table 1). NWI maps underestimated a large portion of VDOT delineated wetlands, and the portions of these wetlands that were mapped were delineated with less precision than typically seen by the NWI (e.g., location 2). The model predictions also resulted in a large number of false negative predictions and imprecise wetland delineations. The well-defined contours of model predictions (e.g., locations 1, 3, and 4) exemplify the heavy reliance of the model on soil thematic layers. In these scenes, the primary drivers for wetland prediction were the presence of hydric soils and shallow depth to water table, which both outlined the same contours as these wetland predictions. The relatively lower reliance on topographic indices in this site is likely due to the unchanging topography of the area, which is characteristic of the Mid-Atlantic Coastal Plain, as there was often little to no flow convergence indicated by the topographic indices where VDOT delineated wetlands were mapped. It is possible that alternative filtering techniques or more severe adjustments to the slope grid could increase the effectiveness of topographic indices to detect wetted areas, however the correspondence of the model to the soil layers used and the relatively high occurrence of false negative predictions imply that vegetation data would also be valuable in this region.

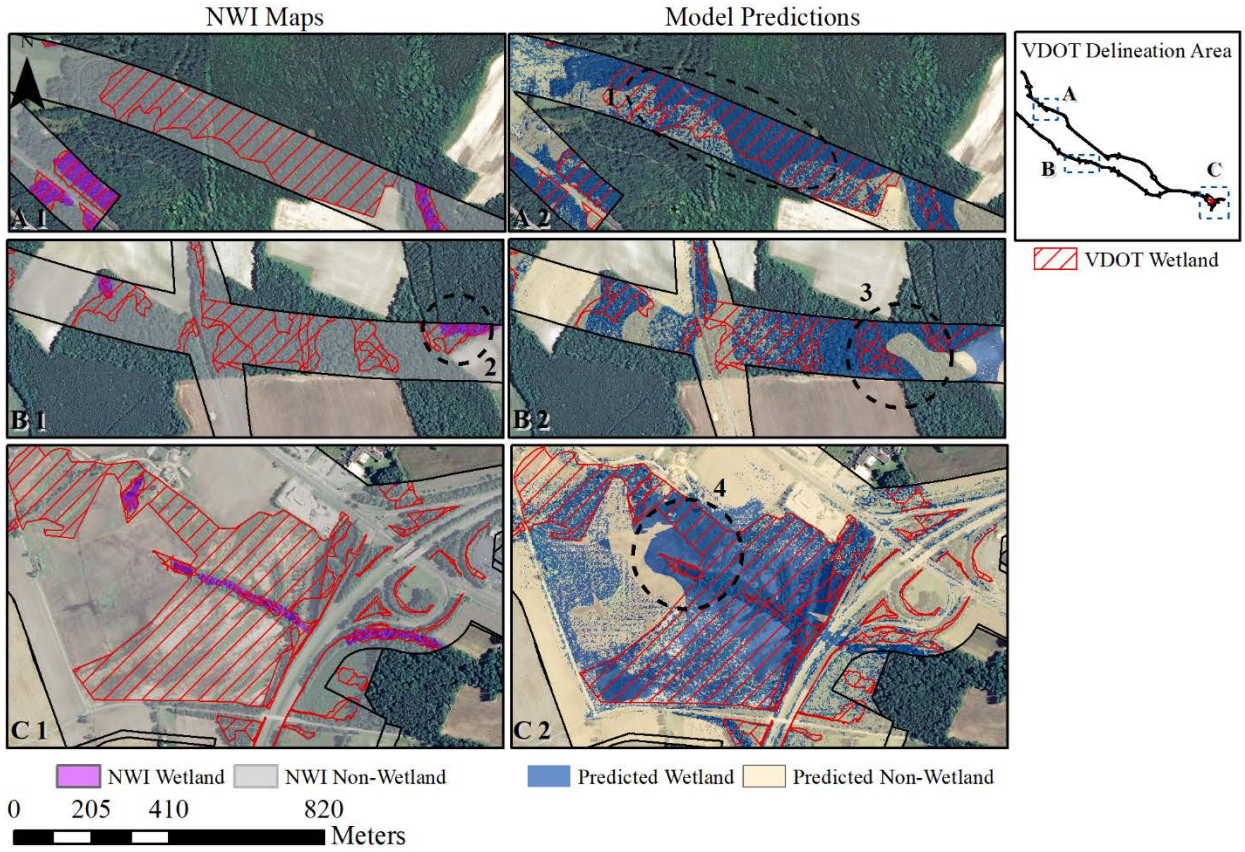

Figure 9. Examples of NWI maps (A1, B1, and C1) and model predictions (A2, B2, and C2) for Site 4, both compared to VDOT delineations.

5.1.5. Variable Importance

An important output from the RF classification was the ESRI Classifier Definition file, which provided the variable importance of each input used in classifications (see Table 4). Variable importance measures were used to gauge the ability of input variables to provide unique, significant information to the classifier. Table 4 shows that in Site 1, Site 3, and Site 4, the modified DTW was the most important topographic index, and in Site 2 the original DTW was the most important topographic index. In contrast, the modified TWI was the overall least important input variable in every study site. The low ranking of the modified TWI relative to the modified and original DTW suggests that some information was duplicated by these inputs, but

that the modified DTW provided more robust wetland and non-wetland signatures. This corresponds to the findings of previous studies (e.g., Ågren et al., 2014; Murphy et al., 2009) which stated that wet TWI values were restricted to discrete lines of flow accumulation within wetted areas, whereas the DTW model effectively encompassed wetted areas as a whole and was therefore more robust. For this same reason, it was unexpected that for Site 3 the modified DTW ranked higher than the modified TWI, as the VDOT delineated wetlands here were primarily restricted to narrow lines of flow accumulation. Soil data were among the most important variables in all sites that included them. In Site 1 and Site 2, this is likely due to the heavy presence of road features and the ability of the soil information to better distinguish these from wetland features relative to the topographic indices, which were observed to detect water accumulation near these features. The higher importance of soil layers in Site 4 is likely due to the flat terrain, and is in line with the wetland predictions seen in Figure 9, which were dictated primarily by areas of hydric soil and shallow depth to water table. The low importance of the topographic indices in Site 4 also reinforces the claim that topographic indices that are static and assume the local slope is an adequate proxy subsurface flow patterns, such as the TWI and DTW, are less suitable in flat areas due to undefined flow directions that are likely to change over time (Grabs et al., 2009). The lower importance of modified curvature relative to DTW inputs in all sites may indicate that our application of the curvature was limited by the ArcGIS fill operation and smoothing, which generalized potentially significant terrain features, since curvature has been shown to strongly determine flow convergence in flat topography (Sangireddy et al., 2016).

5.1.6. Accuracy Assessment

The accuracy of model predictions was assessed using the testing data, and compared to the accuracy achieved by the NWI maps. Table 5 shows the confusion matrices produced for the

best performing model and the NWI maps across all study sites. In each confusion matrix, test data are represented along columns and NWI and model predictions are represented along rows. Categorized pixels (expressed as total km^2) in Table 5 were used to calculate wetland accuracy, non-wetland accuracy, and overall accuracy using Equations 5-7. It is important to note that the accuracy assessment only extended to the limits of the testing data, which as previously described, are randomly selected subsets of the original VDOT delineations, and the effect of varying testing and training data separation on model accuracy was not assessed.

Table 5. Confusion matrices used to assess the accuracy of NWI maps (left) and best performing model predictions (right) compared to the test data, where predicted values are represented along rows and actual values are represented along columns. Wetland, non-wetland, and overall accuracy rates are derived from values in the confusions matrices using Equations 5-7.

Site 1

NWI (predicted)	**Test Data (actual)**			**Model (Predicted)**	**Test Data (actual)**		
	Wetland (km^2)	Non-Wetland (km^2)	∑=		Wetland (km^2)	Non-Wetland (km^2)	∑=
Wetland (km^2)	0.012	0.034	*0.05*	Wetland (km^2)	0.056	0.202	*0.26*
Non-Wetland (km^2)	0.053	2.605	*2.66*	Non-Wetland (km^2)	0.009	2.441	*2.45*
∑=	*0.07*	*2.64*	*2.7*	∑=	*0.07*	*2.64*	*2.7*

Site 2

NWI (predicted)	**Test Data (actual)**			**Model (Predicted)**	**Test Data (actual)**		
	Wetland (km^2)	Non-Wetland (km^2)	∑=		Wetland (km^2)	Non-Wetland (km^2)	∑=
Wetland (km^2)	0.064	0.280	*0.34*	Wetland (km^2)	0.127	1.038	*1.16*
Non-Wetland (km^2)	0.084	6.673	*6.76*	Non-Wetland (km^2)	0.021	5.912	*5.93*
∑=	*0.15*	*6.95*	*7.1*	∑=	*0.15*	*6.95*	*7.1*

Site 3

NWI (predicted)	**Test Data (actual)**			**Model (Predicted)**	**Test Data (actual)**		
	Wetland (km^2)	Non-Wetland (km^2)	∑=		Wetland (km^2)	Non-Wetland (km^2)	∑=
Wetland (km^2)	0.010	0.022	*0.03*	Wetland (km^2)	0.026	0.203	*0.23*
Non-Wetland (km^2)	0.020	1.592	*1.61*	Non-Wetland (km^2)	0.004	1.411	*1.41*
∑=	*0.03*	*1.61*	*1.6*	∑=	*0.03*	*1.61*	*1.6*

Site 4

NWI (predicted)	**Test Data (actual)**			**Model (Predicted)**	**Test Data (actual)**		
	Wetland (km^2)	Non-Wetland (km^2)	∑=		Wetland (km^2)	Non-Wetland (km^2)	∑=
Wetland (km^2)	1.052	0.116	*1.16*	Wetland (km^2)	2.648	1.717	*4.37*
Non-Wetland (km^2)	2.220	7.596	*9.81*	Non-Wetland (km^2)	0.625	6.005	*6.63*
∑=	*3.27*	*7.71*	*11.0*	∑=	*3.27*	*7.71*	*11.0*

Note: Values shown are rounded for clarity.

Figure 10 summarizes the accuracy achieved by the best performing model predictions and NWI maps. In the context of the wetland permitting process, it is important to have high values for all accuracy metrics. To uphold the objective of protecting existing wetlands, wetland accuracy is of high importance, and in order to provide realistic estimates of potentially impacted wetland areas in transportation and environmental planning, non-wetland accuracy is also

necessary. However, it is important to be aware of the potential for overall accuracy, which measures the portion of the entire area that is correctly classified regardless of class, to be misleading due to the uneven distribution of landscape classes. For example, the consistently conservative wetland mapping by the NWI is reflected by the high average non-wetland accuracy (98.0%). Due to the uneven distribution of wetland and non-wetland classes in all but one of the study sites, the conservative nature of the NWI predictions also translated into high average overall accuracy (92.0%), despite an average wetland accuracy of 32.1 %. In contrast, the model predictions resulted in significantly higher average wetland accuracy (84.9%), but at the expense of moderately lower average non-wetland and overall accuracy (85.6% and 85.6%, respectively). As previously discussed, Site 4 was the lowest performing site. The low wetland accuracy here may be due to a lack of vegetative signatures to distinguish wetland from upland area, especially in this excessively flat area where terrain indices were found to be less important.

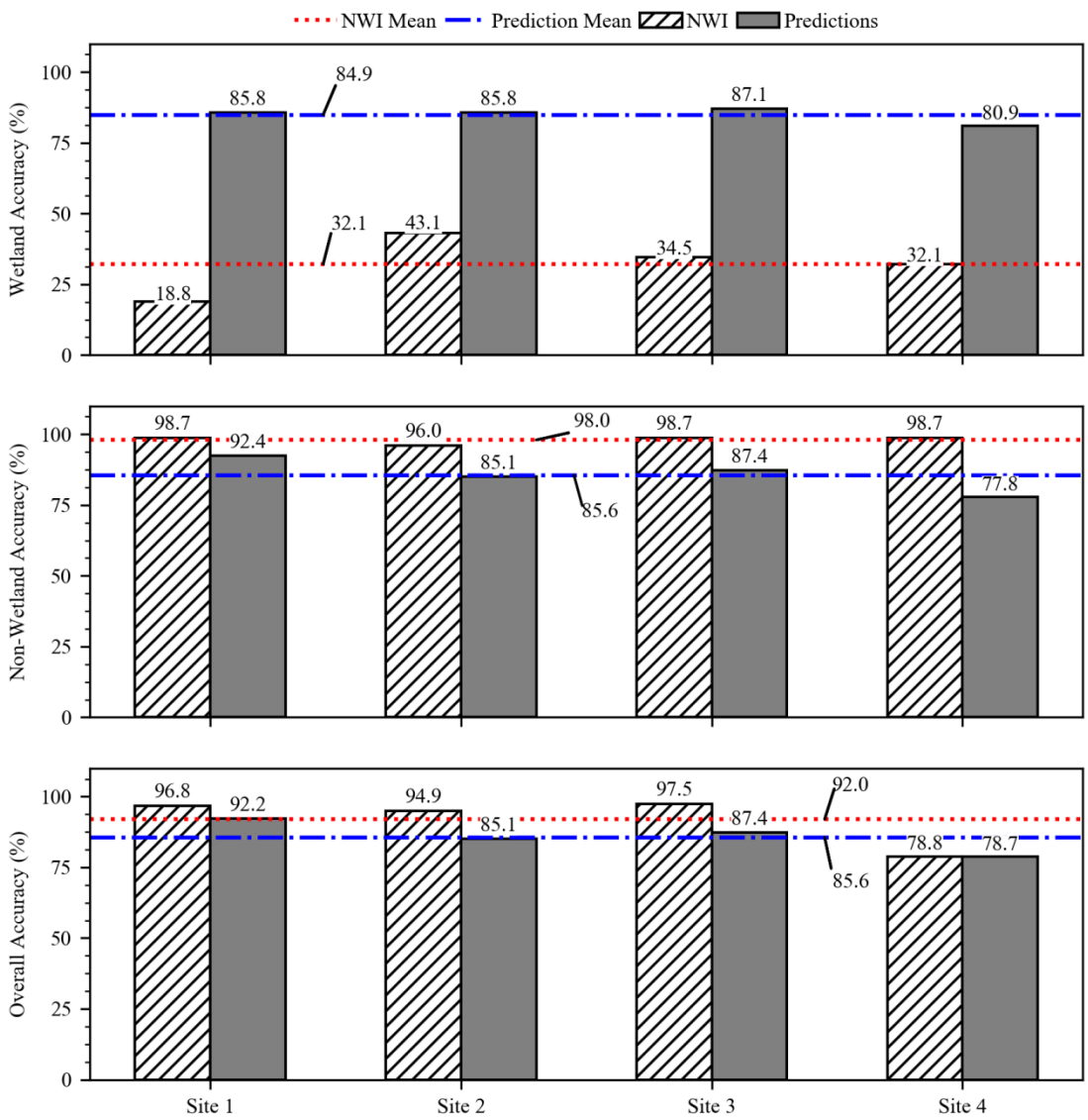

Figure 10. Wetland, non-wetland, and overall accuracy produced by the best performing model predictions, compared to accuracy produced by NWI maps.

5.2. Response of Model to Input Data Modification

Iteration results in terms of wetland, non-wetland, and overall accuracy highlight the benefit and cost of applying the modifications described here, as well as including the coarser mapped (1:24,000 to 1:12,000) SSURGO data. Results of the analysis of model responses to classification iterations are shown in Table 6, where the highest performing iteration per

accuracy metric, not including iteration 5 which built off of top performing topographic inputs, is indicated with a "+" superscript and modified topographic inputs are indicated with an asterisk.

Table 6. Wetland, non-wetland, and overall accuracy achieved by iterations of RF classification for each site. Asterisk indicates modifications with parameters from Table 2 were applied and "+" superscript indicates highest performing iteration per accuracy metric.

	Iteration:	1	2	3	4	5
	Input Data:	TWI, Curvature, DTW	TWI*, Curvature*, DTW*	TWI, Curvature, DTW*	TWI*, Curvature*, DTW	Best Performing of 1-4, plus soils
Site 1	Wetland Accuracy (%)	86.26^{+}	83.65	84.47	85.97	85.84
	Non-Wetland Accuracy (%)	88.34	90.45^{+}	87.77	89.15	92.36
	Overall Accuracy (%)	88.29	90.29^{+}	87.69	89.08	92.20
Site 2	Wetland Accuracy (%)	67.57	69.85	71.33^{+}	69.50	85.78
	Non-Wetland Accuracy (%)	83.58	83.87	81.14	84.26^{+}	85.06
	Overall Accuracy (%)	83.25	83.58	80.94	84.13^{+}	85.08
Site 3	Wetland Accuracy (%)	82.72	87.12	83.88	88.10^{+}	-
	Non-Wetland Accuracy (%)	85.20	87.40^{+}	83.49	86.72	-
	Overall Accuracy (%)	85.16	87.40^{+}	83.50	86.74	-
Site 4	Wetland Accuracy (%)	55.15	57.11	62.67^{+}	60.74	80.91
	Non-Wetland Accuracy (%)	69.31	78.03^{+}	64.44	71.97	77.76
	Overall Accuracy (%)	65.09	71.80^{+}	63.91	68.63	78.70

Shown in Table 6, non-wetland accuracy and overall classification accuracy from iteration 1, where the original versions of all indices were used, improved in every site as a result of modifying all topographic indices (iteration 2). In addition, for three of the four sites, modifying all topographic indices resulted in the highest overall accuracy. These results suggest there is a benefit to applying the modifications presented here rather than using the indices as they are traditionally calculated, where this benefit is a reduction in false positive predictions and increase in overall accuracy. Furthermore, in every site that relevant soil layers were applicable, the inclusion of these soil layers with top performing topographic indices (i.e., iteration 5) further improved the RF classification. From this, we conclude that in these sites, the soil data provided important information to the classifier, despite its relatively coarse scale. Both Site 2 and Site 4 saw relatively high increases in wetland accuracy resulting from iteration 5, which suggests the

topographic indices were not effective in encompassing flow convergence or subsurface moisture conditions in order to detect wetlands. Iterations 3 and 4 were performed to determine the effect of individual modifications on the classification. Note that for this evaluation, modified TWI and modified curvature were generalized into a single category of modifications because of their similar adjustment parameters and methods.

The purpose of modifying topographic indices was largely to reduce false positive predictions in that TWI and curvature grids were modified to reduce unrealistic flow convergence due to excess topographic detail, and the DTW was modified to accelerate the transition from wetland to upland areas. Results in Table 6 show that the effect of modifying only the TWI and curvature grids (iteration 4 vs. iteration 1) was an increase in non-wetland accuracy in every study site, as well as an increase in wetland accuracy in all but Site 1. The decrease in wetland accuracy in this site may indicate unintentional smoothing of some features of interest (i.e., too large of a window size), and it is possible that in this study site a mean filter or smaller window would have performed better. In sites 2, 3 and 4, results of iteration 4 suggest the statistic type and window size were effective. Despite the improvements to classifications with these modifications, the modified TWI and curvature grids can be further advanced. The current approach should be expanded to test the effects of varying window sizes of smoothing filters and statistic type, as well as the TWI formulation.

The effect of modifying only the DTW (iteration 3 vs. iteration 1) appeared to be an increase in wetland accuracy in sites 2, 3, and 4, and an unexpected decrease in non-wetland accuracy in every site. This suggests that while the modified DTW was effective in increasing non-wetland accuracy when combined with modified TWI and modified curvature, the DTW modification alone may not be sufficient for reducing false positive predictions. The limited

improvements provided by the DTW modification could be due to the designation of the representative wetland slope value, which may not apply an effective cut off between low and high cost areas. Additionally, improvements to the original DTW calculation before applying modifications may enhance the results of iteration 3. The DTW calculation can be improved first through slope calculation on a DEM corrected with an alternate method, and second by deriving the source grid by extracting surface water features directly from the LiDAR data. In this study, DTW source grids were generated from rasterized NHD data, which are mapped at a coarser scale (1: 24,000 – 1: 12,000) compared to the LiDAR data and therefore, do not capture precise curvature and locations of streams and open water.

6. Conclusions

This study evaluated the potential for modification of LiDAR DEM derivatives, combined with ancillary national-scale soil data, to improve a RF classification of wetland areas at a scale relevant for the wetland permitting process, over four study sites in Virginia. The approach was implemented as a model in ArcGIS and performed a RF classification of input variables that were modified to provide distinct wetland and non-wetland signatures. Model predictions were assessed against field-mapped testing data, provided by the Virginia DOT, and compared to NWI maps. Accuracy assessments showed that compared to NWI maps, the highest performing models produced significantly higher average wetland accuracy (84.9% and 32.1%, respectively), while resulting in moderately lower average non-wetland accuracy (85.6% and 98.0%, respectively) and overall accuracy (85.6% and 92.0%, respectively).

Through multiple iterations of input variable combinations, we concluded that there is potential to improve classifications through modification of topographic indices. In every site,

the highest performing models included modified topographic indices, and the addition of available soil layers further improved these classifications. Assessment of the variable importance of the highest performing models showed that DTW inputs were of higher importance, compared to the modified TWI in all study sites. This finding supports conclusions of previous studies (e.g., Ågren et al., 2014; Murphy et al., 2009), which state the DTW model provides more robust flow convergence information compared to the TWI. The low variable importance of the TWI relative to the DTW also suggests that there is duplicate information provided between these two indices. In addition, the heavy reliance of the model in Site 4 on soil data reinforces previous findings that topographic indices like the TWI and DTW are less effective in flat areas due to undefined flow directions that are likely to change over time, whereas these indices typically model static conditions and assume local slope describes subsurface flow patterns (Grabs et al., 2009; Murphy et al., 2009). Through classification iterations, we found that non-wetland and overall classification accuracy increased in all sites when all topographic indices were modified, compared to the accuracy achieved by using the original versions of these indices. While modifications to the DTW alone did not reduce false positive predictions, modifications to only the TWI and curvature did have this effect. However, we believe the DTW modification approach could be further improved on. In addition, iteration accuracies varied by small margins in many cases, and it is important to note that that RF parameters and training and testing data separation were not varied or calibrated to sites in this study. Completing this additional calibration step may produce different outcomes of iteration comparisons.

Results from this study offer a starting point to the enhancement of the model implementation in ArcGIS to include the capability of modifying LiDAR DEM derivatives based

on site characteristics to map small-scale wetlands in support of environmental planning and conservation efforts. The results while successful, have also highlighted shortcomings that should be addressed to further enhance the approach and model implementation. We found that the topographic indices were limited by the use of the ArcGIS fill function, which removed local depressions in the LiDAR DEM by creating larger areas of flat terrain. Studies have shown that high-resolution elevation data could be filtered with more sophisticated methods (e.g., Besl et al., 1989; Haralick et al., 1983; Lindsay et al., 2016; Mainguy et al., 1995; Sangireddy et al., 2016), and exploring these methods could improve the accuracy of the topographic indices, especially in low relief areas. The TWI modification can be further advanced on by assessing model responses to alternate TWI formulations such as the D-infinity method for deriving flow accumulation (Tarboton, 1997) and the Soil Topographic Index formulation which has been shown to improve modelling of soil moisture patterns through inclusion of relevant soil properties (e.g., Buchanan et al., 2014; Lanni et al., 2011). Alternate curvature modifications should also be explored, as this index has been shown to effectively model flow convergence in low-relief and engineered landscapes by applying automated filtering techniques (Sangireddy et al., 2016). Improvements to the DTW modification should include deriving source data directly from LiDAR DEMs through calibrated flow initiation thresholds, as shown by Ågren et al. (2014), and deriving flow accumulation using the D-infinity method (Murphy et al., 2009, 2011), or incorporating the use of other channel extracting software, such as GeoNet (Sangireddy et al., 2016). Furthermore, variable importance indicated that the DTW and TWI may provide duplicate information in many cases, and efforts should be made to effectively combine these indices through a mathematical relationship to reduce feature space for the classifier. Future work should also address the excessive computation times needed to process the high-resolution LiDAR data.

Implementing this approach using parallel computing could allow for reductions in runtime needed to calculate γ and β parameters through an iterative calibration to study sites in the DTW modification process. Alternative implementations of the RF algorithm should be tested as well, as the ArcGIS implementation is limited in output data provided to users. Lastly, the approach presented here should be applied to additional study areas to begin to identify modification parameters that can be effectively generalized by site characteristics. While the prototype model has produced more accurate wetland predictions for the study sites compared to NWI, these improvements would strengthen the potential for this approach to be a useful tool for wetland identification in support of environmental planning decision making in areas where wetland maps are currently unavailable.

7. Acknowledgements

The authors wish to thank Michael Fitch, J. Cooper Wamsley, Daniel Redgate, and Caleb Parks, from the Virginia Department of Transportation (VDOT), for providing valuable feedback throughout the course of this project and for their support in collecting important data. Funding for this project was provided by the Virginia Transportation Research Council (VTRC).

8. References

Ågren, A. M., Lidberg, W., Strömgren, M., Ogilvie, J., & Arp, P. A. (2014). Evaluating digital terrain indices for soil wetness mapping -- a Swedish case study. Hydrology & Earth System Sciences, 18(9), 3623-3634. doi:10.5194/hess-18-3623-2014 .

Ali, G., Birkel, C., Tetzlaff, D., Soulsby, C., McDonnell, J. J., & Tarolli, P. (2014). A comparison of wetness indices for the prediction of observed connected saturated areas under contrasting conditions. Earth Surface Processes and Landforms, 39(3), 399-413. doi: 10.1002/esp.3506 .

Besl, P. J., Birch, J. B., & Watson, L. T. (1989). Robust window operators. Machine Vision and Applications, 2(4), 179-191. doi:10.1007/BF01215874 .

Beven, K., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrological Sciences Journal, 24(1), 43-69.

Bradski, G. (2000). The OpenCV Library. Dr. Dobb's Journal: Software Tools for the Professional Programmer, 25(11), 120-123.

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

Buchanan, B. P., Fleming, M., Schneider, R. L., Richards, B. K., Archibald, J., Qiu, Z., & Walter, M. T. (2014). Evaluating topographic wetness indices across central New York agricultural landscapes. Hydrology and Earth System Sciences, 18(8), 3279. doi:10.5194/hess-18-3279-2014

Burrough, P. A., and McDonell, R. A., 1998. Principles of Geographical Information Systems (Oxford University Press, New York), 190 pp.

Cowardin, L., & Golet, F. (1995). US Fish and Wildlife Service 1979 Wetland Classification: A Review. Vegetatio, 118(1/2), 139-152.

Dahl, T. E., Johnson, C. E., & Frayer, W. E. (1991). Wetlands, status and trends in the conterminous United States mid-1970's to mid-1980's. US Fish and Wildlife Service.

Davidson, N. C. (2014). How much wetland has the world lost? Long-term and recent trends in global wetland area. Marine and Freshwater Research, 65(10), 934-941. doi: 10.1071/MF14173 .

Deng, J., Smith, A. S., Davis, S., Weatherford, M., Paugh, L., & Wang, S. G. (2017). Identification of NC Wetland Types by LiDAR Data and Tree Based Machine Learning Methods (No. 17-01199).

Duro, D. C., Franklin, S. E., & Dubé, M. G. (2012). A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sensing of Environment, 118, 259-272. doi:10.1016/j.rse.2011.11.020 .

Environmental Laboratory. (1987). Corps of Engineers wetlands delineation manual.

ESRI. (2016). Train Random Trees. Retrieved June 8, 2017, from http://desktop.arcgis.com/en/arcmap/10.4/tools/spatial-analyst-toolbox/train-random-trees-classifier.htm.

Federal Register. July 13, 1994. Changes in hydric soils of the United States.

Grabs, T., Seibert, J., Bishop, K., & Laudon, H. (2009). Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model. Journal of Hydrology, 373(1), 15-23. doi:10.1016/j.jhydrol.2009.03.031 .

Guo, M., Li, J., Sheng, C., Xu, J., & Wu, L. (2017). A Review of Wetland Remote Sensing. Sensors, 17(4), 777. doi:10.3390/s17040777 .

Haralick, R. M., Watson, L. T., & Laffey, T. J. (1983). The topographic primal sketch. The International Journal of Robotics Research, 2(1), 50-72.

Hogg, A. R., & Todd, K. W. (2007). Automated discrimination of upland and wetland using terrain derivatives. Canadian Journal of Remote Sensing, 33, S68-S83. doi:10.5589/m07-049 .

Jenson, S. K., & Domingue, J. O. (1988). Extracting topographic structure from digital elevation data for geographic information system analysis. Photogrammetric Engineering and Remote Sensing, 54(11), 1593-1600.

Klemas, V. (2011). Remote sensing of wetlands: case studies comparing practical techniques. Journal of Coastal Research, 27(3), 418-427. doi:10.2112/JCOASTRES-D-10-00174.1 .

Kloiber, S. M., Macleod, R. D., Smith, A. J., Knight, J. F., & Huberty, B. J. (2015). A semi-automated, multi-source data fusion update of a wetland inventory for east-central Minnesota, USA. Wetlands, 35(2), 335-348. doi:10.1007/s13157-014-0621-3 .

Lang, M. W., & McCarty, G. (2014). Light Detection and Ranging (LiDAR) for Improved Mapping of Wetland Resources and Assessment of Wetland Conservation Practices USDA, Natural Resources Conservation Service.

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

Lanni, C., McDonnell, J. J., & Rigon, R. (2011). On the relative role of upslope and downslope topography for describing water flow path and storage dynamics: a theoretical analysis. Hydrological Processes, 25(25), 3909-3923. doi:10.1002/hyp.8263 .

Lindsay, J. B. (2016). Efficient hybrid breaching-filling sink removal methods for flow path enforcement in digital elevation models. Hydrological Processes, 30(6), 846-857. doi: 10.1002/hyp.10648 .

MacMillan, R. A., Martin, T. C., Earle, T. J., & McNabb, D. H. (2003). Automated analysis and classification of landforms using high-resolution digital elevation data: applications and issues. Canadian Journal of Remote Sensing, 29(5), 592-606. doi:10.5589/m03-031 .

Mainguy, Y., Birch, J. B., & Watson, L. T. (1995). A robust variable order facet model for image data. Machine vision and applications, 8(3), 141-162. doi:10.1007/BF01215810 .

Miao, X., Heaton, J. S., Zheng, S., Charlet, D. A., & Liu, H. (2012). Applying tree-based ensemble algorithms to the classification of ecological zones using multi-temporal multi-source remote-sensing data. International Journal of Remote Sensing, 33(6), 1823-1849. doi: 10.1080/01431161.2011.602651 .

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.

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. doi:10.5589/m13-038 .

Moore, I. D., R. B. Grayson, and A. R. Landson. 1991. Digital Terrain Modelling: A Review of Hydrological, Geomorphological, and Biological Applications. Hydrological Processes 5: 3–30.

Morrissey, L., & Sweeney, W. (2006). Assessment of the national wetlands inventory implications for wetland protection. Proceedings of the 2006 AWRA Spring Specialty conference: GIS and water resources IV. DR Maidment & JS Grounds III, Houston, TX, 139 papers/abstracts (May 8–10, 2006)(TPS-06-1)(ISBN 1-882132-70-X) CD-ROM (Specialty Conference).

Murphy, P., Ogilvie, J., & Arp, P. (2009). Topographic modelling of soil moisture conditions: a comparison and verification of two models. European Journal of Soil Science, 60(1), 94-109. doi:10.1111/j.1365-2389.2008.01094.x .

Murphy, P. N., Ogilvie, J., Connor, K., & Arp, P. A. (2007). Mapping wetlands: a comparison of two different approaches for New Brunswick, Canada. Wetlands, 27(4), 846-854. doi: 10.1672/0277-5212(2007)27[846:MWACOT]2.0.CO;2 .

Murphy, P. N., Ogilvie, J., Meng, F. R., White, B., Bhatti, J. S., & Arp, P. A. (2011). Modelling and mapping topographic variations in forest soils at high resolution: A case study. Ecological Modelling, 222(14), 2314-2332. doi:10.1016/j.ecolmodel.2011.01.003 .

Natural Resources Conservation Service (NRCS). (2015). Soil Data Viewer 6.2 User Guide. Unpublished manuscript.

Page, R., & Wilcher, L. (1990). Memorandum of Agreement Between the Environmental Protection Agency and the Department of the Army concerning the determination of mitigation under the Clean Water Act, Section 404 (b)(1) Guidelines. Washington, DC, USA.

Planchon, O., & Darboux, F. (2002). A fast, simple and versatile algorithm to fill the depressions of digital elevation models doi:10.1016/S0341-8162(01)00164-3 .

Richardson, M. C., Fortin, M. J., & Branfireun, B. A. (2009). Hydrogeomorphic edge detection and delineation of landscape functional units from lidar digital elevation models. Water Resources Research, 45(10). doi:10.1029/2008WR007518 .

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing, 67, 93-104. doi:10.1016/j.isprsjprs.2011.11.002 .

Sangireddy, H., Stark, C. P., Kladzyk, A., & Passalacqua, P. (2016). GeoNet: An open source software for the automatic and objective extraction of channel heads, channel network, and channel morphology from high resolution topography data. Environmental Modelling & Software, 83, 58-73. doi: 10.1016/j.envsoft.2016.04.026 .

Snyder, D., & Lang, M. (2012). Significance of the 3D Elevation Program to wetlands mapping. National Wetlands Newsletter, 34(5), 11-15.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Retrieved August 7, 2017, from https://websoilsurvey.nrcs.usda.gov/.

Sørensen, R., & Seibert, J. (2007). Effects of DEM resolution on the calculation of topographical indices: TWI and its components. Journal of Hydrology, 347(1), 79-89. doi:10.1016/j.jhydrol.2007.09.001 .

Tarboton, D. G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. Water resources research, 33(2), 309-319. doi: 10.1029/96WR03137 .

U.S. Environmental Protection Agency (EPA). (2013). Primary Distinguishing Characteristics of Level III Ecoregions of the Continental United States.

Woodrow, K., Lindsay, J. B., & Berg, A. A. (2016). Evaluating DEM conditioning techniques, elevation source data, and grid resolution for field-scale hydrological parameter extraction. Journal of Hydrology, 540, 1022-1029. doi:10.1016/j.jhydrol.2016.07.018 .

WorldView Solutions Inc. (2016). Technical Plan of Operations: Virginia Statewide Land Cover Data Development. Richmond, VA.

Zhu, J., & Pierskalla, W. P. (2016). Applying a weighted random forests method to extract karst sinkholes from LiDAR data. Journal of Hydrology, 533, 343-352. doi: 10.1016/j.jhydrol.2015.12.012 .