

APPENDIX

Additional details regarding the data processing methods.

Lidar DEMs

The lidar data within Maryland were collected for the Maryland Department of Natural Resources (metadata hosted online: <http://dnrweb.dnr.state.md.us/gis/data/lidar/>). The lidar data within Delaware were collected as part of the 2007 Delaware coastal Program Lidar effort (metadata hosted online: <https://data.noaa.gov/harvest/object/7418610c-3842-4671-9040-1bac0faf9372/html>). These datasets were designed to meet or exceed the Federal Geographic Data Committee's National Standards for Spatial Data Accuracy for data at 1:2,400. The estimated horizontal positional accuracy of point returns exceeded 50 cm.

Mapping Wetland Depressions

In identifying wetland depressions using Whitebox Geospatial Analysis Tools, a modified version of the turning bands simulation technique (Matheron 1973) was used to introduce spatial autocorrelation into the potential error values (Ahrens 2012). Changes in the number of depressions identified were tested up to 50 iterations, but were found to stabilize after 20 iterations. Near-infrared lidar returns over water typically reflect the water surface elevation instead of the elevation of the ground below it (Lane and D'Amico 2010), necessitating consideration of wetness conditions at the time the lidar data was collected. The lidar data were collected on dates representing near normal or average wetness conditions. The Palmer Hydrological Drought Index (PHDI) ranged

from -1.8 to 3.6 and averaged 1.4 across the dates (NOAA NCDC 2016). The data were collected in spring months, a seasonally wet period.

Mapping the Stream Network

To map the total stream network, the existing semi-automated stream layer (Lang et al. 2012) was first burned into the DEM using Whitebox Geospatial Analysis Tools. Depressions were filled to hydrologically condition the DEM (Wang and Liu 2006) prior to calculating the flow accumulation values. Because drainage in the watershed has been heavily modified from agricultural activities, it was not always clear if a point was located on a naturally occurring stream, a ditch that replaced a naturally occurring stream, or a ditch. Eight of the points showed unrealistically low accumulation areas (<25,000) and using aerial imagery and differences between stream datasets, were found to most likely occur along non-topographically conforming ditches. These points were therefore excluded from the analysis prior to selecting an accumulation threshold.

Processing Radarsat-2 Imagery

The 3x3 covariance matrix averages the cross-polarization backscatter, and allows the polarization intensities (HH, VV, and HV) to be extracted for further analysis (Lee and Pottier 2009). The polarimetric decomposition of fully polarized data, in turn, can help maximize the ability of SAR data to distinguish physical features on the ground, including water (Baghdadi et al. 2001; Henderson and Lewis 2008). The traditional Kennaugh matrix is the linear transformation of the four-dimensional Stokes vector, which consists of the total intensity (K_0) and 15 linear coefficients of the transformation. This approach directly interprets and scales

the backscattering matrix elements themselves, deriving total intensity, as well as elements related to absorption, diattenuation, and retardance (Schmitt and Brisco, 2013). The normalized Kennaugh matrix (k) was derived by dividing the Kennaugh matrix by the total intensity (I), so that all elements range between -1 and 1.

$$[K] = \begin{bmatrix} K_0 & K_4 & K_5 & K_6 \\ K_4 & K_1 & K_9 & K_8 \\ K_5 & K_9 & K_2 & K_7 \\ K_6 & K_8 & K_7 & K_3 \end{bmatrix} = I \cdot \begin{bmatrix} 1 & k_4 & k_5 & k_6 \\ k_4 & k_1 & k_9 & k_8 \\ k_5 & k_9 & k_2 & k_7 \\ k_6 & k_8 & k_7 & k_3 \end{bmatrix} = I \cdot [k] \quad (1)$$

Random forest models are generally insensitive to collinearity among metrics, however the inclusion of correlated variables can deflate variable importance and overall variation explained, while the inclusion of a large number of variables can make interpretation difficult and introduce noise (Murphy et al. 2010). Because of this we implemented variable selection using random forests models in R (varSelRF package). We ran an initial random forest model with all metrics then a revised model that only included metrics selected by varSelRF. Accuracy statistics were generally improved by using the subset of metrics; therefore the final maps were derived using the subset of metrics, listed in Table A1 in the Appendix.

Table A1. Rasters included in the forested random forest model for each of the five image dates. Variables were selected using the varSelRF package in R. Only Radarsat-2 outputs included in at least one forest random forest model are listed.

Output	Raster	24-Mar-15	26-Mar-15	31-Mar-15	2-Apr-15	9-Apr-15
Covariance matrix	Covariance Matrix 1,1 ($S_{hh}S_{hh}^*$)	x	x	x	x	x
	Covariance Matrix 2,2 ($S_{hv}S_{hv}^*$)				x	
	Covariance Matrix 3,3 ($S_{vv}S_{vv}^*$)	x	x	x	x	
Kennaugh scattering matrix	Element 2, 2 of Kennaugh matrix (k_1)	x		x	x	x
	Element 3, 3 of Kennaugh matrix (k_2)	x	x	x	x	x
	Element 4, 4 of Kennaugh matrix (k_3)	x		x	x	

	Element 1, 2 of Kennaugh matrix (k_4)	x	x	x	x	
	Element 1, 3 of Kennaugh matrix (k_5)					x
	Element 1, 4 of Kennaugh matrix (k_6)					x
	Element 2, 4 of Kennaugh matrix (k_8)					x
	Element 3, 4 of Kennaugh matrix (k_7)	x				
Freeman-Durden decomposition	Power contributions due to double-bounce	x		x		x
	Power contributions due to volume scattering	x	x			x
Cloude-Pottier Decomposition	Entropy	x				
	Alpha Angle	x				x
	Beta Angle	x	x			x
	Eigenvalues - Lambda 1	x				
	Eigenvalues - Lambda 2	x				
	Eigenvalues - Lambda 3	x				x
	Real component of element 1 of Eigenvector 2	x				
Touzi Decomposition	Dominant Eigenvalue	x				
	Dominant Touzi Alpha_S Parameter			x	x	x
	Dominant Touzi Phase					
	Dominant Tau Angle (Helicity)					x
	Secondary Eigenvalue	x				x
	Tertiary Eigenvalue	x				x