NCDOT Wetland Modeling Program: Development of Tidal Wetland Models using QL2 LiDAR

Final Report

P.I.: Sheng-Guo Wang, Professor, UNC Charlotte

UNCC WAM Research Team
with Axiom Research Team
for NCDOT RP 2016-19

Date: 11-20-2018

<table>
<tr><td>1. Report No.
NCDOT RP 2016-19</td><td>2. Government Accession No.
… …</td><td>3. Recipient's Catalog No.
… …</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
NCDOT Wetland Modeling Program: Development of Tidal Wetland Models using QL2 LiDAR</td><td>5. Report Date
Nov. 20, 2018</td></tr>
<tr><td>6. Performing Organization Code
… …</td></tr>
<tr><td colspan="2">7. Author(s)
Sheng-Guo Wang (PI) with Shanshan Jiang, Sandy Smith, Scott Davis</td><td>8. Performing Organization Report No.
… …</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
Dept. of Engineering Technology & Dept. of Computer Science
University of North Carolina – Charlotte
Charlotte, NC 28223-0001
with Axiom environmental, Inc.
218 Snow Ave., Raleigh, NC 27603</td><td>10. Work Unit No. (TRAIS)
… …</td></tr>
<tr><td>11. Contract or Grant No.
…NCDOT RP 2016-19…</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
NC Department of Transportation</td><td>13. Type of Report and Period Covered
Final Report
March 16, 2015 – August 15, 2018</td></tr>
<tr><td>14. Sponsoring Agency Code
NCDOT Project # 2016-19</td></tr>
<tr><td colspan="3">15. Supplementary Notes:
This project was supported by a grant from the U.S. Department of Transportation and the North Carolina Department of Transportation</td></tr>
</table>

16. Abstract

This Final Report is to summarize several main achievements of this project as follows:

(i) Automation Method and its Tools for the Tidal Wetland Identification and Analysis Process using QL2;
(ii) Method Development for Tidal Wetland Identification Process;
(iii) Reliability and Flexibility of Automation Tools and Methods; and
(iv) User Friendly deliverables.

These achievements fit the NCDOT research needs as: "while NCDOT has made significant advances with the concept, the process and tools of predicting wetlands using LiDAR is under-developed."

That also completes the goal of the project to provide an advanced QL2 LiDAR-based tidal wetland prediction method and automation tools based on ArcGIS for the NC coastal region.

The UNC Charlotte WAM Research Team with Axiom Research Team has successfully completed a number of valuable research topics related to tidal wetland prediction process, such as process automation, variables exploration, data mining, and statistical analysis, and best resolution selection.

The acclaimed results include the deliverable WAMAT-Tidal: WAM Automation Tools - Tidal and the Users' Guide to the Tools, tidal wetland prediction methods, and the best resolution determination method, as well as new WAMAT v4.4 & v5.1.

<table>
<tr><td colspan="2">17. Key Words
Wetland, Automation, Tidal Wetland, Modeling, Prediction, Analysis</td><td colspan="2">18. Distribution Statement
… …</td></tr>
<tr><td>19. Security Classif. (of this report)
Unclassified</td><td>20. Security Classif. (of this page)
Unclassified</td><td>21. No. of Pages
42</td><td>22. Price
… …</td></tr>
</table>

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

DISCLAIMER

The contents of this report reflect the views of the authors and not necessarily the views of the University. The authors are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the North Carolina Department of Transportation or the Federal Highway Administration at the time of publication. This report does not constitute a standard, specification, or regulation.

ACKNOWLEDGEMENTS

The research team thanks the North Carolina Department of Transportation for supporting and funding this project. We extend special thanks to the project Steering and Implementation Committee members:

- Morgan Weatherford (Chair)
- Philip S. Harris III
- LeiLani Paugh
- David Johnson
- James Mason
- Colin Mellor
- Sarah Schwarzer
- Neil Mastin
- John W. Kirby

The research team is indebted to the tremendous support provided by these committee members in helping advices and the scope of the project.

Special thanks are given to LeiLani Paugh, Morgan Weatherford, and John Kirby at NCDOT who provided us with valuable information and expert advice about the North Carolina Wetland Assessment Method (NC WAM), and strong support for the project. Sincere thanks are to Neil Mastin for his strong support.

The P.I. also wants to thank our partner Key Investigators Alexander P. (Sandy) Smith and Scott Davis at Axiom Environmental for their expert support, especially, for wetland field visits, Tidal Zone, and model testing. We worked so closely for the plan, variable set and its generation, field work, and frequent discussions.

Without the help of all of the above individuals, this project could not have such scientific results in such a successful manner, leading to our deliverable new automation tools WAMAT-Tidal and update WAMAT v4.4 and v5.1 as a useful product.

EXECUTIVE SUMMARY

The NCDOT has partnered with several federal agencies in funding the development of standard QL2 LiDAR elevation data for the North Carolina (NC) coastal region [7]. This effort follows both national and international recognition [6-10] of the importance in developing and integrating airborne LiDAR digital imagery and pattern-recognition technology into a GIS-based method for 21st century transportation and environment monitoring, measurement, and inventory. As part of this process, NCDOT noted that sufficient datasets depicting tidal wetlands are outdated and/or not accurate enough to use in the NEPA/LEDPA (National Environmental Policy Act/Least Environmentally Damaging Practicable Alternative) selection process. NCDOT has used prediction models in non-tidal portions of the state for palustrine wetlands [3, 8], but it is expected that different models will need to be developed for tidal wetlands. With the arrival of the new QL2 LiDAR, additional research will be needed to determine how to utilize and optimize the voluminous dataset [6].

Our goal for this project is to provide an advanced QL2 LiDAR-based tidal wetland prediction method and automation tools based on ArcGIS for the NC coastal region. Based on the NCDOT's needs [6], we have proposed a scope of work in this project as follows:

- Conduct a literature review and investigate the status of existing methods and models of LiDAR-based tidal wetland prediction and use of the QL2 standard LiDAR data;
- Determine the optimal resolution of DEM and subsequent terrain derivatives, and any other variables needed to predict wetlands via orthogonal test design approach [14] on QL2 LiDAR data and other related data;
- Develop appropriate methods to model tidal wetland boundary locations via first-hand experience of wetland scientists [4, 10, 5], regression method Logit (logistic regression), machine learning method RF (random forest) [15, 60, 68];
- Develop tools to automate the process of sampling, interpolation, variable creation, and model development and application where it is appropriate and feasible to do so;
- Validate our developing methods and models through field testing; and
- Prepare deliverable products including the proposed methods, models, algorithms, and tools [10.A– F].

The ultimate methods offer tidal wetland prediction models with machine learning (ML) methods for modeling and prediction. The automation tools vividly display the results of the process based on the GIS platform (ArcGIS and ArcMap) that NCDOT currently uses.

The PI and his research team at UNCC have worked closely with wetland scientists from Axiom Environmental, Inc. as a joint research team for this project.

This project has been successfully completed and can enhance identification and prediction of tidal wetlands using QL2 LiDAR data, machine learning, pattern recognition, and GIS, thereby significantly reducing the time and cost of field delineations. The results of this project will also provide a cost-effective source of potential wetland impacts that will improve the efficiency of initial project planning [8] and the NEPA process [9].

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

This Final Report is for the NCDOT Research Project RP 2016-19, titled “NCDOT Wetland Modeling Program: Development of Tidal Wetland Models using QL2 Lidar” during 04-01-2015 through 08-15-2018. It concludes several main achievements of this project as follows:

(i) Automation and its Tools of Tidal Wetland Identification and Analysis Process, which we call WAM Automation Tools – Tidal or WAMAT–Tidal in short;

(ii) Systematic Methods of Tidal Wetland Identification Process including Machine Learning Methods;

(iii) Reliability and Flexibility of the Developed Tools and Methods;

(iv) Best Resolution Determination Method along with Taguchi approach; and

(v) User Friendly Deliverables as listed in Attachments [10.A – 10.D, 10.F, 10.G]

This project is based on previous projects, e.g., the 2011 FHWA Environmental Excellence Awards (EEA) winner NCDOT and NCDENR “GIS-based Wetland and Stream Predictive Models” [8], and the 2015 National “Sweet Sixteen” High Value Research Award winner NCDOT Research Project 2013-13 “Improvements to NCDOT’s Wetland Prediction Model” [1-3, 5, 69].

As recognized nationally and internationally [6-8], there is a trend toward development and integration of airborne LiDAR, digital imagery [1-3, 6-8], and machine learning pattern recognition technology [2, 15, 17, 19, 68] for 21st century transportation and environmental monitoring, measurement, and inventory. This technology supports enhanced wetland prediction and enables reliable identification of wetland locations, thus reducing the time and cost of field delineations and providing early awareness of potential wetland impact areas in NC [8].

NCDOT has sponsored research into and development of an automated wetland prediction model to supplant the majority of field-based wetland delineations as part of a major streamlining initiative during the NEPA process. The results of the model give NCDOT the ability to compare alternatives of road projects while greatly decreasing the need for field delineated wetlands. However, much of that research has been focused on palustrine wetlands in the North Carolina interior. Additionally, NCDOT is making a significant investment to partially fund an update of the statewide LiDAR dataset collected at the QL2 standard.

The need definition of the NCDOT addressed by this project is to enhance research into and development of an automated wetland prediction model, especially an automated tidal wetland prediction model, to supplant the majority of field-based wetland delineations. Sufficient datasets depicting tidal wetlands are outdated and/or not accurate enough to use in the NEPA/LEDPA selection process. With the arrival of the new QL2 LiDAR, this research completes the need to determine how to utilize and optimize the voluminous dataset. The above-mentioned achievements fit the NCDOT research needs.

The goal of this project is to provide an advanced QL2 LiDAR-based tidal wetland prediction method and automation tools based on ArcGIS for the NC coastal region. The benefits to NCDOT include significantly reducing the time and cost of field delineations and providing early awareness of potential wetland impact areas in NC.

The significance of LiDAR implementation into wetland identification and modeling, as stated by the FHWA is to exemplify "how innovative technologies can be used to speed the environmental assessment process and ultimately advance transportation projects while protecting the environment" [8]. Therefore, this project research, e.g., [66 – 69], is important and highly needed. In addition, it contributes to NCDOT by keeping the leading status in this important area of research [10.A – 10.G], which can benefit NCDOT by innovative modeling and predicting automation tools and significant labor saving in the NEPA process [9].

This project includes a number of valuable research topics related to wetland and tidal wetland prediction, such as process automation, variables exploration, data mining, machine learning, and statistical analysis. According to the project proposal [1], our goal for this project is to provide improved NCDOT LiDAR-based tidal wetland prediction models with *highly automated, reliable, and user-friendly tools* for NCDOT based on ArcGIS. In addition, this project provides a method to identify the best resolution for modeling and prediction. Therefore, we mainly concentrate on the topics of process automation and modeling and prediction methods for this project.

The rest of this report is organized in the following manner. Chapter 2 is to summarize our developed key deliverable tools: Tidal Wetland Prediction Automation Tools, called WAM Automation Tools – Tidal or WAMAT-Tidal in short. Chapter 3 describes the research results of our tidal wetland prediction models, including the tidal wetland variable set and two models of Logit and Random Forest (RF). Chapter 4 presents the process automation in tidal wetland prediction. In Chapter 5, case studies are conducted by applying our models and automation process to Brunswick and New Hanover counties, NC. Chapter 6 is about the best resolution research. Finally, Chapter 7 provides summary remarks of the project with the highlight of our deliverable research results. In addition, following the conclusions, the published papers and presentations are listed in Chapter 8, the References are listed in Chapter 9, and the Attachments are listed in Chapter 10 as Appendix.

This final report also includes the attached deliverables: automation tools package of WAMAT-Tidal for the Tidal Wetland Prediction Process Automation with its Users' Guide to the Tools directly to NCDOT, and the updated automation tools package of WAMAT v.4.4 and v.5.1 for the wetland prediction process automation with their Users' Guides.

2. WAMAT-Tidal: Tidal Wetland Prediction Automation Tools

This NCDOT project has a key deliverable that is the tidal wetland prediction automation tools package. It is developed based on our WAMAT (WAM Automation Tools) for the key task of this research project to complete the tidal wetland prediction process automation. This tools package is called WAMAT–Tidal in short. It includes the automation of the following processes:

(i) Tidal wetland variable generation process,

(ii) Tidal wetland model generation process,

(iii) Tidal wetland prediction process,

(iv) Tidal wetland evaluation process, and

(v) Full process of tidal wetland prediction including these above individual automation processes as a combined process for automatic run just by one click of the WAMAT-Tidal.

All automated processes are simple to run. In addition, WAMAT–Tidal has a function to easily remove individual variables, e.g., land cover or soils, and add new variables based on users' choices. Thus, it has flexibility not only in model selection, but also in variable selection.

The main structure of tidal wetland automatic prediction process WAMAT–Tidal is shown in Fig.1.

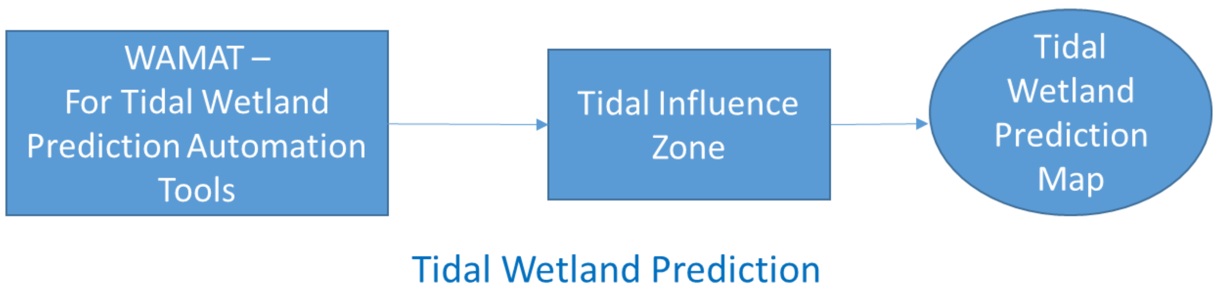

Figure 1. Key structure of WAMAT-Tidal for automatic tidal wetland prediction process

In WAMAT-Tidal, the tidal wetland variable set is based on QL2 LiDAR data and some special variables as described in the next chapter for this research project. As mentioned above, the WAMAT–Tidal has flexibility of its predictor variable selection.

The provided models including Logit and Random Forest (RF) based on the tidal wetland prediction variable set as described in Chapter 4. The prediction process can be run by either Logit model or RF model from the modeling process. After the prediction, the accuracy is evaluated in the evaluation process by the ground truth input data with colors.

During this project period, the UNCC WAM Research Team has further developed the WAMAT as the updated version v.4 (including v.4.0, v.4.1, v.4.2, v.4.3 and v.4.4) from the previous version v.3.2. Its function upgrade is summarized in Figure 2 below.

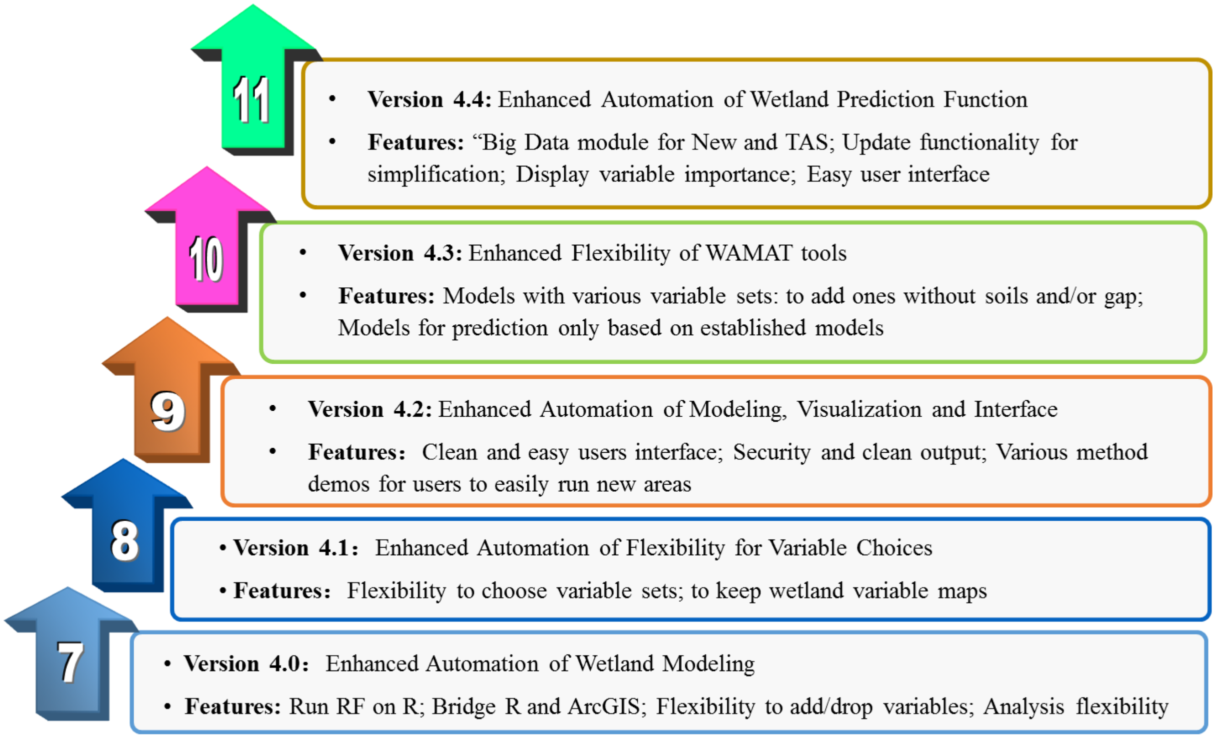

Figure 2. WAMAT Enhancement from v.3 to v.4 as v4.0 through v4.4

This updated WAMAT provides the NCDOT with enhanced automation and flexibility in variable selection (e.g., with both, either, or neither soil and/or land cover), model building RF on R outside ArcGIS, that leads to enhanced speed and accuracy. It also keeps the variable maps for the user to see and make use of. This makes it particularly easy for the user to run the model in new areas.

Further, the updated WAMAT adds a Big-Data (large area) Prediction ability for "New" and "TAS" approaches, while previously only the preferable "Regular" approach has set this Big-Data Prediction function. Moreover, the user interface is clean and easy; for example, there is a single interface where users can set their input files once, then click one button to run the whole process automatically.

The resultant prediction map is generated to show wetland and non-wetland areas in green and yellow, respectively, as well as the evaluation colors to depict prediction accuracy 1-1 as 1 (correct tidal wetland prediction, in dark green), 0-0 as 0 (correct non-wetland prediction, in grey), 2-2 as 2 (correct non-tidal wetland prediction, in green), and -1 (incorrect prediction, i.e., error, in red).

The WAMAT update provides a powerful base for the WAMAT-Tidal tools. The WAMAT-Tidal tools structure is shown in Fig. 3.

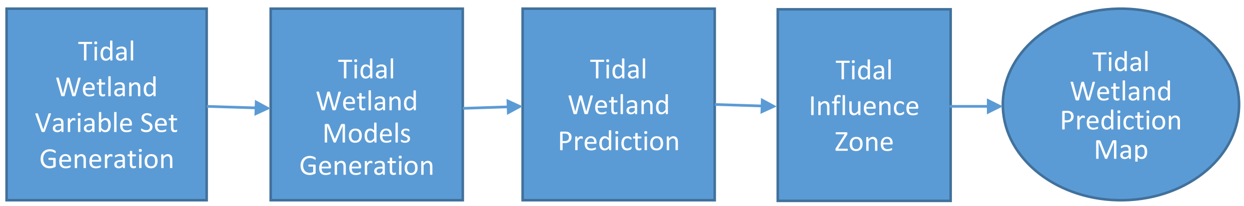

Figure 3. WAMAT-Tidal tools structure

Some additional information of WAMAT–Tidal for automation will be described further in Chapter 4. We shall describe the new tidal variable set, the models and the prediction methods in Chapter 3.

3. Tidal Wetland Prediction Models

In this chapter, we summarize the wetland prediction models and their methods we applied and developed with their performances by using QL2 LiDAR. We developed two models for the tidal wetlands prediction as follows.

(1) Logistic Regression model, and

(2) Random Forest model.

The first step is to determine the variable set for building prediction models. That is as described in the next Section (Section 3.1). After the model variable set has been determined, the next step is to build models by the following two methods as briefly described in Sections 3.2 and 3.3, followed by the intersection with the TIZ map for tidal wetland prediction described in Section 3.4.

3.1 Tidal Wetland Prediction Variables

For the tidal wetland prediction, we take the following predictor variable set as in Table 1.

Table 1. Tidal Wetland Prediction Variable Set used to build the models from QL2 Data

Variable	Full Name	Formula and illustration
TIZ	Tidal Influence Zone Map	A classification variable
TWA	Tidal Water Amplitude	*TWA = Max_elev - DEM*
MHHW_elev	Mean Higher High Water Elevation	A digital variable
Max_elev	Maximum water elevation	A digital variable
veg-l	Low Vegetation/Strata low	QL2 class 3 – low
veg-m	Medium Vegetation/Strata low	QL2 class 4 – medium
veg-h	High Vegetation/Strata high	QL2 class 5 – high
Qvcm	QL2 vegetation dominant class	$\mathrm{Qvcm}(x,y) = \begin{cases} 3, & \max(Pl, Pm, Ph) = Pl \\ 4, & \max(Pl, Pm, Ph) = Pm \\ 5, & \max(Pl, Pm, Ph) = Ph \end{cases}$ where Pl means number of low vegetation las points in one cell, Pm means number of medium vegetation las points in one cell, Ph means number of high vegetation las points in one cell.
vden	Vegetation density	Area & volume of all vegetation types in its neighborhood $\mathrm{vden}(x,y) = \frac{Pl + Pm + Ph}{area\ of\ one\ cell}$
Vl-l	Intensity of low vegetation returns	The classification of vegetation points is labeled as: Low vegetation – 3 Medium vegetation – 4 High vegetation – 5 High intensity values represent photosynthetically active vegetation, while lower intensity values are likely to represent wet surface condition or less photosynthetically active vegetation
Vl-m	Intensity of medium vegetation returns	A digital variable
Vl-h	Intensity of high vegetation returns	A digital variable
water	QL2 water class	QL2 class 9
bldg	QL2 building class	QL2 class 6
rw	QL2 road	QL2 class 13
elev	Elevation	Elevation of each cell: z(x, y)
soil	Soil data	Soil types in Axiom's soil table: mineral -- 1 organic -- 3 Other --2

Table 1. Tidal Wetland Prediction Variable Set used to build the models from QL2 Data (Continued)

Variable	Full Name	Formula and illustration
slp	Slope	In degree: $$slp(x,y) = 57.29578 \times \operatorname{atan}\left(\sqrt{\left(\frac{dz}{dx}\right)^2 + \left(\frac{dz}{dy}\right)^2}\right)$$
cv	Curvature	$$cv(x,y) = 57.29578 \times \operatorname{atan}\left(\sqrt{\left(d\frac{slp}{dx}\right)^2 + \left(d\frac{slp}{dy}\right)^2}\right)$$
curv5	Smooth curvature	Each cell gets mean value of curvature from its 5*5 neighbors. $$curv5(x,y) = \frac{\sum_{i=i1}^{i25} cv(i)}{25}$$
prcv	Profile curvature	Curvature on vertical (y) direction
plcv	Plan curvature	Curvature on horizontal (x) direction
wei	Wetness Elevation Index	Series of increasingly larger neighborhoods used to determine the relative landscape
weiRe	Reclassification of wei	Wei value of each cell will be reclassified as 0 if original value is bigger than a predefined threshold, else is reclassified as 1
asp	Aspect	$$asp(x,y) = 57.29578 \times \operatorname{atan2}\left(\left[\frac{dz}{dy}\right], -\left[\frac{dz}{dx}\right]\right)$$
mdec	Maximum Downslope Elevation Change	Maximum difference of z(x,y) between the cell and its neighbor cells. mdec = Max(z_i - z) z_i is the elevation of a neighbor cell
batwi	Ratio of slope and drainage area	batwi = slp / drainage contributing area (calculated with breached DEM)
gap	Land Cover Data	Categorized land use types

Table 1 lists the variables for building the models by their features via Logit method and RF method. Here, we point out that these variables are derived and generated from QL2 LiDAR data, except the Tidal Influence Zone (TIZ) and Tidal Water Amplitude (TWA), which are two important variables for the tidal wetland modeling and prediction generated with the help of our partner Axiom Environmental. TIZ and TWA will be described below in detail. Also, these variables are listed in the Attachment [B] "WAMAT-Tidal Users' Guide, v.4.1". And TIZ is documented in the attachment [H] "Tidal Influence Zone Dataset". Compared with the original WAMAT tools, this variable set includes newly proposed tidal influence zone, tidal water amplitude, vegetation variables (high, medium, low), etc. The other regular variables are the same as the ones that our WAMAT uses for wetland identification.

The new TIZ variable is described briefly here. Our partner Axiom key investigators have provided the UNCC Team with a Tidal Influence Zone (TIZ) map for the coastal region of North Carolina. The map has been developed using National Oceanographic and Atmospheric Administration (NOAA) data found at the following web site: https://coast.noaa.gov/slr/. The TIZ map utilizes predicted daily and wind-driven tidal water elevations to predict tidal wetland extent. TIZ generation generally consisted of correlating maximum water elevations and depths provided by NOAA with QL2 LiDAR-derived elevation data (2014, Phase 1 and 2015, Phase II LiDAR). Areas of equal elevation were identified and grouped by 14-digit Hydrologic Unit (HUs) that were separated where appropriate to more accurately define changes in maximum water elevation. Subsequently, Axiom has continued investigations for refining the TIZ map including field-verifying mapped water levels.

The updated TIZ map provides a more precise estimate of the areas in North Carolina affected by astronomical and/or wind tides than currently available data. Its main improvements include the following:

- The addition of new and useful attributes:
    - elevations of tidal water,
    - potential coastal island locations,
    - influence of salt or fresh water;
- Division of the Tidal Influence Zone into three parts based on geographic location;
- Assignment of individual Hydrologic Unit identifiers; and
- More precise tidal data modified by field work.

Axiom's new TIZ data are divided into three areas:

(1) TIZ Area A

Area A consists of the areas draining to and adjacent to the Albemarle Sound, including barrier islands. It encompasses the northern portion of the NC TIZ, generally from the NC-Virginia border to Oregon Inlet.

(2) TIZ Area B

Area B consists of the areas draining to and adjacent to the Pamlico Sound, including barrier islands. It encompasses the central portion of the NC TIZ, generally from Oregon Inlet to Beaufort Inlet.

(3) TIZ Area C

Area C consists of the areas draining to and adjacent to the Cape Fear River and the southern coast, including barrier islands. It encompasses the southern portion of the NC TIZ, generally from Beaufort Inlet to the NC-South Carolina border.

The relationship between tidal wetlands and astronomical and wind tides are provided below:

a. The TIZ occupies the area within the Maximum Elevation that water reaches due to astronomical tides (or astronomical tides plus wind tides, where applicable).

b. Astronomical tides occur daily, and the highest average elevation that it reaches (the average of the higher of the two daily high tides) is the Mean Higher High Water Elevation (MHHW_elev).

    - All areas that are inundated daily (i.e. in the TIZ and at or lower than the MHHW_elev) are predicted to be wetlands.

c. Wind tides occur occasionally and can push the water above the MHHW level. The highest level they can normally reach is the Maximum Water Elevation (Max_elev). In areas not subject to wind tides, Max_elev = MHHW_elev.

    - Areas within the TIZ but at an elevation higher than the MHHW elevation may or may not be wetlands, but the wetlands that are found here are considered tidal.

Another variable, Tidal Water Amplitude (TWA) can be derived at individual sites as a function of the Maximum Water Elevation minus the Site Elevation. For example, at the inland (maximum) extent of the TIZ, the TWA is zero and increases moving seaward.

These two variables together with others as listed in Table 1 are applied to our tools. These variable layers are included with the tools. The TIZ variable plays a key role to delineate the tidal influence regions and non-tidal influence regions, similar to the riparian variable to delineate the riparian regions and non-riparian regions. Thus, we utilize the TIZ in the variable set, but also in the final intersection with the predicted wetland map in the coastal areas, which helps to generate the tidal wetlands and non-tidal wetlands.

It is to be emphasized that the methods, models and tools are valid for all various areas when the DEM LiDAR data and TIZ data are available.

In addition, we show some new variables as vden (Vegetation density) and qvcm (QL2 vegetation dominant class) in following Figure 4 and Figure 5, respectively. Our tools automatically generate these two variables vden and qvcm from the input LiDAR data by the formulas as listed in Table 1. These two variables are calculated based on the data of QL2 class 3 (veg-l), class 4 (veg-m) and class 5 (veg-h). Thus, the classified point cloud is required as the input data. Currently, we have successfully run our tools for the test areas in Brunswick County. In the future, we will further test for the maximum sized area that the model can be applied to in view of possible computational restriction and ArcGIS limit.

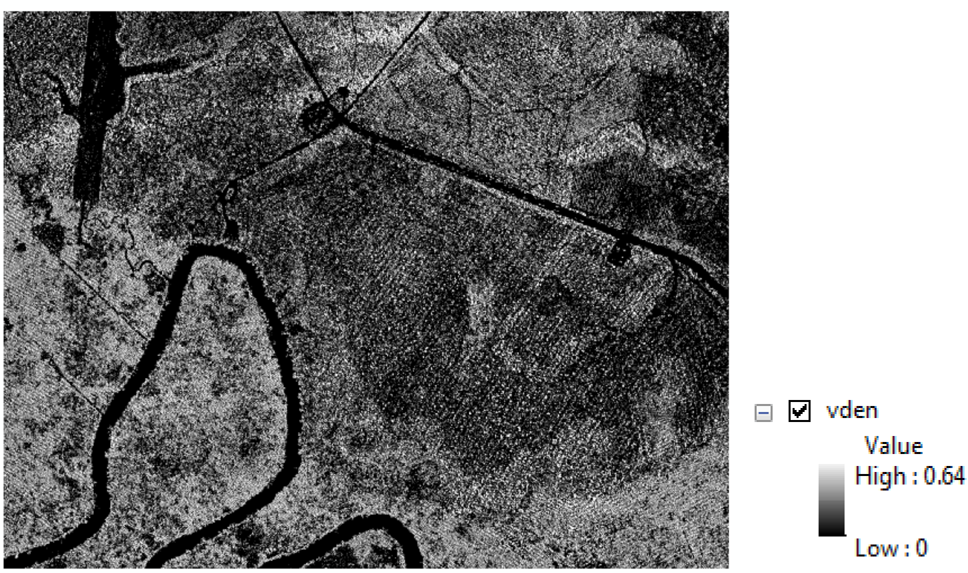

Figure 4. Variable vden (cells with higher values contain higher amounts of vegetation)

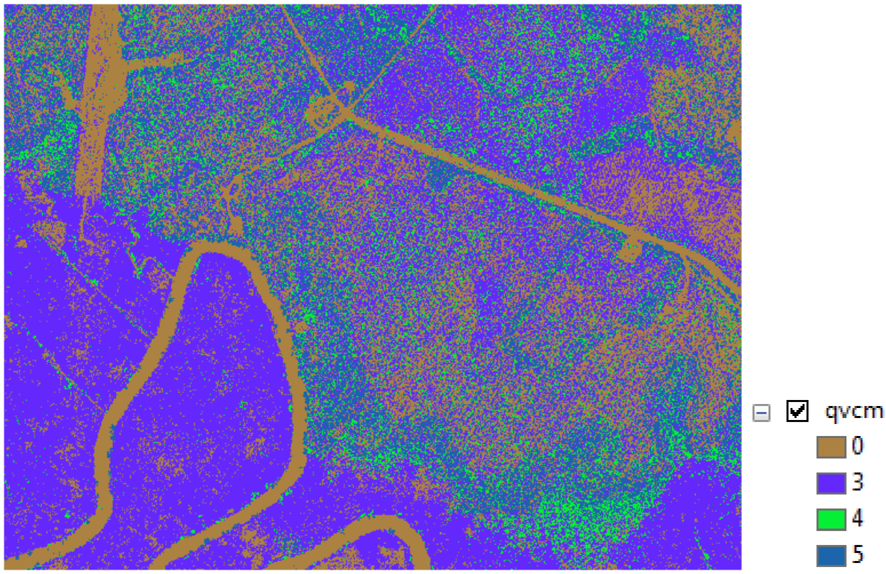

Figure 5. Variable qvcm (0 indicates no vegetation)

In the next two sections, the methods of Logit and RF used to run modeling and prediction of tidal wetlands identification are described. They are also described as in our final report of NCDOT RP 2013-13.

3.2 Logistic Regression (Logit) Model

First, we have applied the logistic regression model to classify the landscape into two categories (wetland and non-wetland) for tidal wetland identification. Before we describe the logistic regression model, let's first describe a linear regression model as in (1), which predicts the occurrence of wetland as a function $y(\boldsymbol{x})$ of the selected explanatory variable vector $\boldsymbol{x}$ at a data point as

$$y(\boldsymbol{x}) = \boldsymbol{\beta}^T \boldsymbol{x} + \varepsilon \tag{1}$$

where $\boldsymbol{x}$ is the wetland variables vector $\boldsymbol{x} = [x_1, x_2, \cdots, x_m]^T$, y is a response variable as the prediction result, $\boldsymbol{\beta}$ is the coefficient vector as a "weighting factor" for the variable vector, and ε is an estimator/noise error or adjustment of this linear estimator. In a prediction area, each point (e.g., 20×20 feet^2 as a point), the variable vector $\boldsymbol{x}$ can be arranged in a matrix X, and the corresponding response variable y can be presented as a vector $\boldsymbol{y}$, where each row represents a data point. Then we have the following linear regression model in a matrix-vector format as

$$\boldsymbol{y}(X) = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2}$$

Because the response vector should be a binary-valued vector, i.e., the prediction model is a two-category classification; therefore, a binary-valued model is used with a logistic function transform to (1) and called logistic regression. Logistic regression is just to take a transform on the continuous-valued response variable to predict a binary response with a “probability” value in [0, 1]. In statistics, the probability describing the possible outcomes of a single trial is modeled as a function of predictor variables, using a logistic function

$$p(\boldsymbol{x}) = F(t) = \frac{e^t}{1+e^t} = \frac{1}{1+e^{-t}} \tag{3}$$

where $t = \boldsymbol{\beta}^T\boldsymbol{x} + \varepsilon$, i.e., to transform a continuous response $y(x)$ in (1) to a binary response. After the logistic function transform, we may have a generalized linear model for binary response in probability as

$$\hat{y} = logit\ (E[y|\boldsymbol{x}\,]) = logit\ (p) = \ln\left(\frac{p}{1-p}\right) = t = \boldsymbol{\beta}^T\boldsymbol{x} + \varepsilon \tag{4}$$

$$p = E[y|\boldsymbol{x}\,] = \frac{1}{1+e^{-\boldsymbol{\beta}^T x-\varepsilon}} \tag{5}$$

Sometimes, it is simply written as a new response variable *y* as follows

$$y = \frac{1}{1+e^{-\boldsymbol{\beta}^T x-\varepsilon}} \tag{6}$$

Also, please notice that the Logit model may be extended for multi-category classification.

3.4 Random Forest (RF)

In order to reduce the sensitivity to data noise and the overfitting problem, we have applied a decision tree-based classification method Random Forest (RF), a machine learning method, with the derivative variables in Table 1 to identify wetlands for tidal wetland prediction. Random trees in RF are built by a set of rules that uses a bagging technique to randomly select sub-datasets and optimization technique to determine the best decision tree nodes from a randomly selected sub-set of variables [15, 68]. Thus, it leads to a random forest. Then, in the prediction process, RF can recursively partition the data into categories.

The classification tree analysis (CTA), also referred to as classification and regression trees (CART), is a typical tree-based classification method. RF aims at improving predictive ability by taking the majority vote result from the prediction results of multiple trees in classification mode, or taking the average result of the prediction results of multiple trees in regression mode. Thus, this method is not sensitive to noise or overtraining, as resampling is not based on weighting. Furthermore, it is computationally much lighter than methods based on boosting and somewhat lighter than simple bagging. In the literature, it is used for land cover classification [32], and recently used for the first time for wetland identification in our publications [68].

Here, we have developed and applied the RF model for the tidal wetland classification by using new QL2 LiDAR, especially using the newly listed variable set in Table 1 for modeling and prediction. For prediction, each tree in the forest generates a class result based on randomly selected input data and a randomly selected sub-set of variable features. Then the method collects the voting results from the resulting trees. It is described in Figure 6.

In the selection of the variables at each node, one of the optimal searches is to calculate the decrease of Gini index (an impurity measure) and another is to calculate the decrease in error, every time a new variable is introduced. These are used for building decision trees.

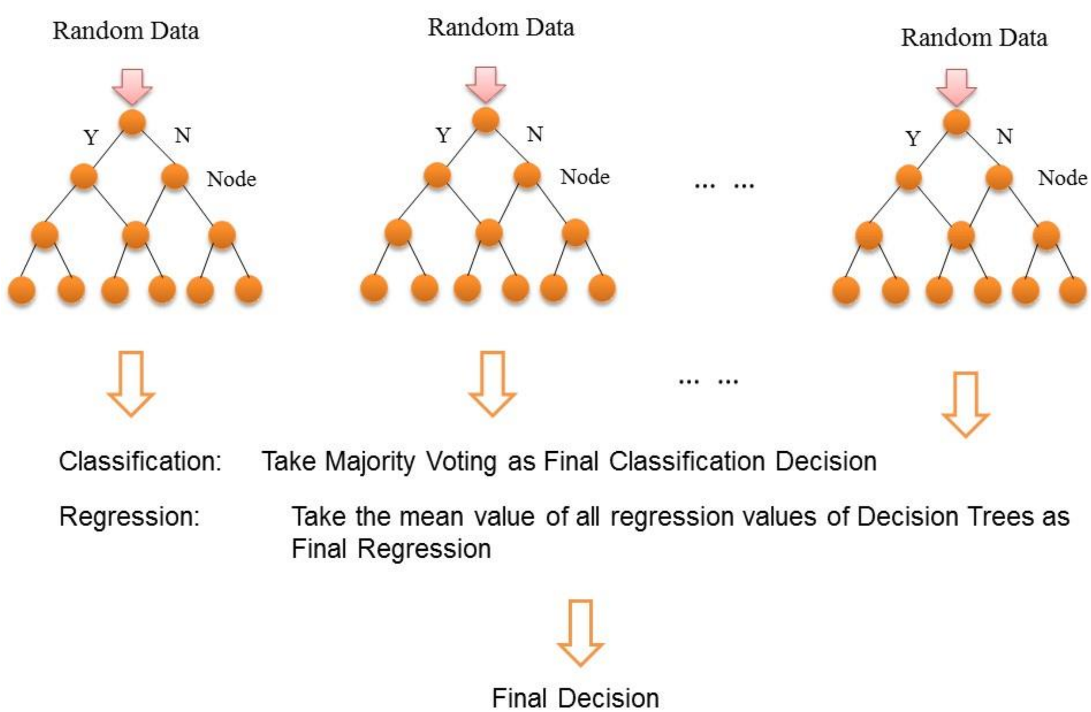

Figure 6. Random Forest Method

Similarly, please notice that the RF model may also work for multi-category classification problems, e.g., tidal wetlands, non-tidal wetlands, and non-wetlands.

4. Automation Process

In this Chapter, we summarize the automation process and the tools we have developed for WAMAT–Tidal, similar to [2, 3, 68] as developed for WAMAT (RP 2013-13) [5]. The detail of WAMAT-Tidal (WAM Automation Tools–Tidal) [10.A] can be found in the Users' Guide as Appendix [10.B].

These tools can be flexibly and automatically run to implement several tasks related to tidal wetland prediction. The main tasks of tidal wetland prediction include:

(a) Data pre-processing of QL2, especially for TIZ and TWA;

(b) Model training;

(c) Predicting;

(d) Wetland mapping;

(e) Intersection with TIZ to generate tidal wetlands, non-tidal wetlands, and non-wetlands;

(f) Model performance evaluation; and

(g) Tidal wetland map display.

The automation tools are developed based on ArcGIS 10.1.

During the project period, we have provided NCDOT five new major versions of our WAMAT as v.4.0 through v.4.4, with their User Guides [10.C, 10.D]. It is as summarized above in Figure 2, Chapter 2. In addition, recently we developed new version WAMAT v.5.1 to fit the NCDOT special requirement to overcome the computation limitation in the current GIS [*10.F, *10.G].

4.1. Advantages of the Automation Tool WAMAT–Tidal

There are some important features of the WAMAT-Tidal based on WAMAT new versions as summarized below.

(1) Flexible:

WAMAT-Tidal has all flexibility from WAMAT. In addition, it has flexibility to easily add or remove the predictor variables for building models and running predictions.

(2) Efficient:

The algorithm and tools are both enhanced to be able to predict large areas based on WAMAT. We efficiently divide the data in big areas and then combine their results in the algorithm, thus it can be quickly calculated and run well.

(3) User friendly:

The simple interface is more straightforward and applicable. Users can easily change their data files, such as linking them to the files in different folds for different areas for running tidal wetland modeling and prediction in different areas.

In addition, we applied the automation tools for tidal wetland detection using input with different resolutions to test and identify the best resolution in modeling and prediction.

5. Case Study and Field Validation

This Chapter describes the field visit and the case studies. The field visit areas are in Brunswick County.

5.1 Tidal wetland prediction of Brunswick and New Hanover Counties

We have implemented the automation process of tidal wetland prediction for several areas in Brunswick and New Hanover counties. Our prediction models are built by the sampled training QL2 data from areas provided by Axiom and NCDOT (Figure 7) and the TIZ map (Figure 8). It is emphasized that the predicted areas are not in the training areas, but are extended areas along two directions as NW and E from the two wide-sides of the training area (Fig. 7).

Figure 7. Wetland training area

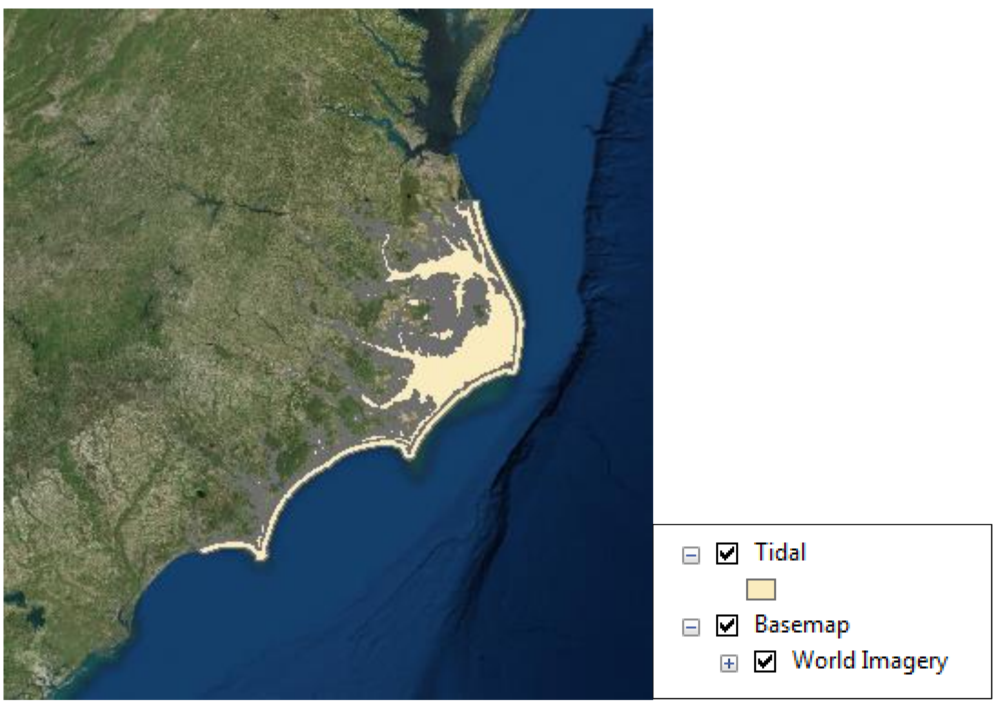

Figure 8. Tidal influence zone (TIZ)

5.2. Model construction

We have run the following process of building our models and predicting tidal wetlands. It runs our models including two machine learning methods: (1) Logistic Regression method, and (2)

Random Forest method. We have also run the process via an "approach A" and "dynamic resolution" for study.

Approach A. To build wetland model with wetland types,

(i) To predict wetland types by the above-built model,

(ii) To combine predicted wetland types into a combined wetlands prediction,

(iii) To run intersection of the combined wetlands prediction with the TIZ for resultant tidal wetlands (and types if needed).

Dynamic Resolution is to let the source data have various resolutions for modeling and prediction.

5.3. Field validation

Axiom has visited appropriate sites within the TIZ to field-verify the results of TIZ generation and tidal water extents in various areas. Wetland areas have been delineated across a representative sample of ecoregions, and the data have been provided to UNCC team for analysis and model refinement.

With expert Sandy Smith at Axiom Environmental, our team executed a validation visit to Brunswick and New Hanover counties on August 1 and 2, 2018, two full days. Expert Scott Davis at Axiom also provided useful maps for the team to run this field visit.

The goals of this field trip are: (1) to validate automated wetland identification digital maps generated by using Logistic Regression (Logit) model and Random Forest (RF) model; (2) to differentiate the tidal and non-tidal wetlands; and (3) to collect wetland types for future further studies and applications.

Methodology has been developed with the machine learning-based RF method and a regression Logit method. We first identify the wetland areas via approach A and dynamic resolution by using our developing WAMAT–Tidal tools, then overlay the identified wetland areas on the TIZ map to determine if the predicted wetland is potentially tidal.

A brief summary of the field test is as follows.

- Study area

– Wetland training area: it is located in Brunswick County provided by Axiom, as shown in Figure 7 in above Section 5.1.

– Wetland and transect area: it is also shown in Figure 7 above.

– Tidal Influence Zone: as provided by Axiom, it is shown in Figure 8 in above Section 5.1.

– Areas for modeling verification: It takes the intersection of the training area and the TIZ, as shown below on Figure 9. We then verified the prediction result for this tidal wetland.

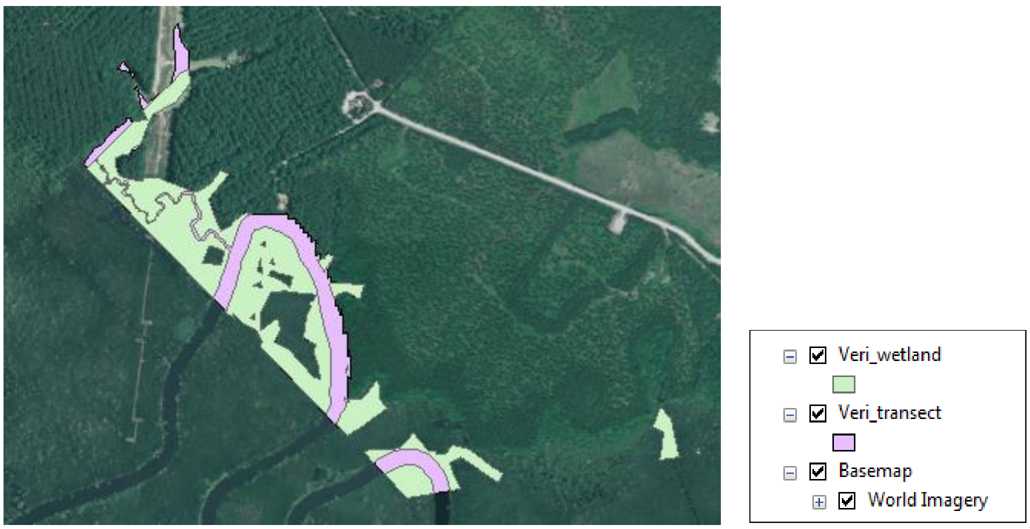

Figure 9. Tidal wetland verification area for model building

– Regions for tidal wetland prediction verification: We investigated the four extended regions, as shown in Figure 10, where we visited 8 sites. Then, we verified the prediction results for the tidal wetlands during the field visit.

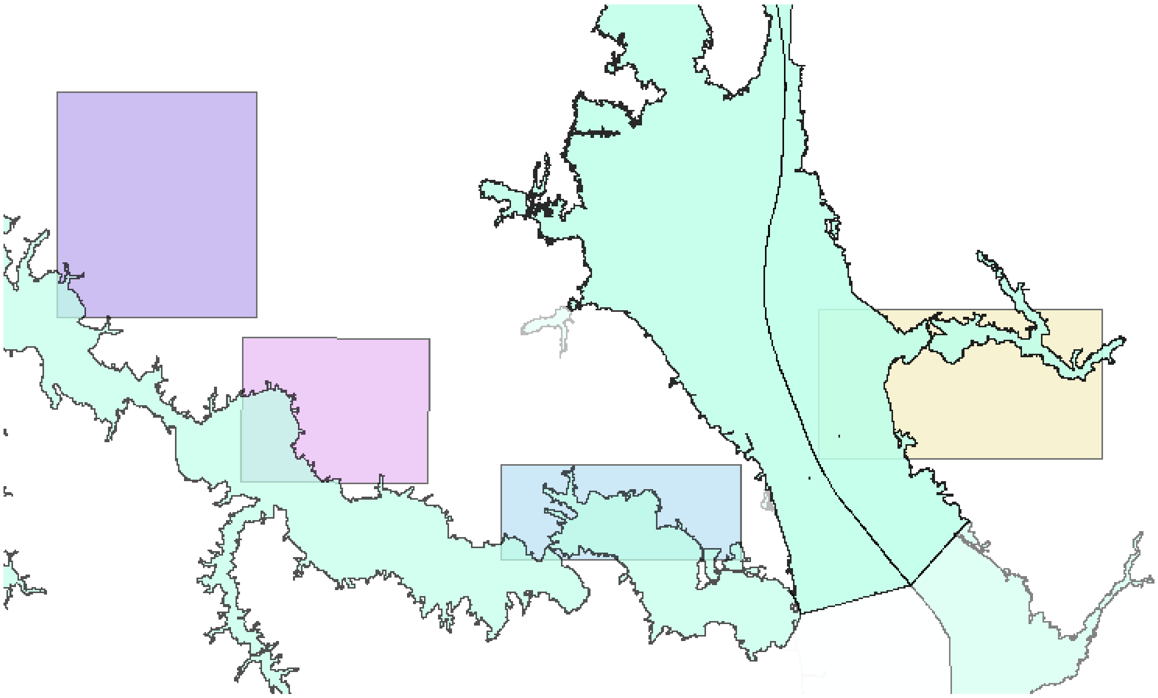

Figure 10. Regions used for tidal wetland prediction and verification

The field validation results in this field test show that

(i) In modeling, RF has better results than Logit; i.e., RF results in less modeling error.

(ii) In non-tidal wetland prediction, Logit usually gives more accurate wetland predictions than RF.

(iii) In tidal wetland prediction, both Logit and RF models provide the same very highly accurate tidal wetland predictions.

(iv) Within the TIZ, RF results are better because RF shows better performance in excluding roads/water

A summary for tidal wetland predictions of two models, Logit and RF in Brunswick and New Hanover counties is described in detail in the Attachment [10.E].

6. Method for Best Resolution Identification and Test

6.1 Best resolution determination method

For the best resolution determination in modeling and prediction, the PI has introduced a new method for a dynamic multi-resolution scheme test and analysis of tidal wetland modeling and prediction by the orthogonal experiment design using the Taguchi method [14, 24]. Based on that new method, we have run experiments for the best resolution selection/identification as listed in Table 2 below.

The object of this experiment is to find resolutions of input and output that achieve the higher accuracy, i.e., to find the best resolution.

The study area is in Brunswick County with the training data from Axiom's field work is shown in Figure 7. We applied the Taguchi orthogonal method to reduce number of tests to obtain the best resolution solution.

6.2 Best resolution test of QL2 data for tidal wetland prediction

Digital Elevation Models (DEMs) representing the QL2 data have been provided by the NC Division of Emergency Management (NCDEM) with resolutions of 5, 10, 20, and 50 feet, respectively. So, what is the best resolution of QL2 data for tidal wetland prediction? The goal is to determine the best resolution among a combination of data files for the best accuracy of prediction. Here, we consider choices among the 5, 10, and 20-foot DEMs in view of 50 feet is too large for accurate prediction, thus there are three levels for each factor.

The setting of resolution parameters test is from the Taguchi method, and its experiments lead to the following best resolution set for Logit and RF respectively as shown in Table 2.

Table 2. Best Resolution Recommendations

	logit	RF
DEM	5	20
Soil	10	20
Vegetation	20	20
Output normalized	10	5

We ran a validation analysis for the final suggested best resolution from Table 2 based on Taguchi method and the PI's dynamic multi-resolution test. The results are shown in the following figures.

```
**********************************************************
error 0/1: 177/192900=0.0917573872473%
error 1/0: 109/83551=0.130459240464%
0/0: 192723/192900=99.9082426128%
1/1: 83442/83551=99.8695407595%
error total: 286/276451=0.103454138346%
**********************************************************
```

Figure 11. Accuracy and error rate validation of the recommended RF resolution

```
**********************************************************
error 0/1: 549/48220=1.13853172957%
error 1/0: 181/20857=0.867814163111%
0/0: 47671/48220=98.8614682704%
1/1: 20676/20857=99.1321858369%
error total: 730/69077=1.05679169622%
**********************************************************
```

Figure 12. Accuracy and error rate validation of the recommended logit resolution

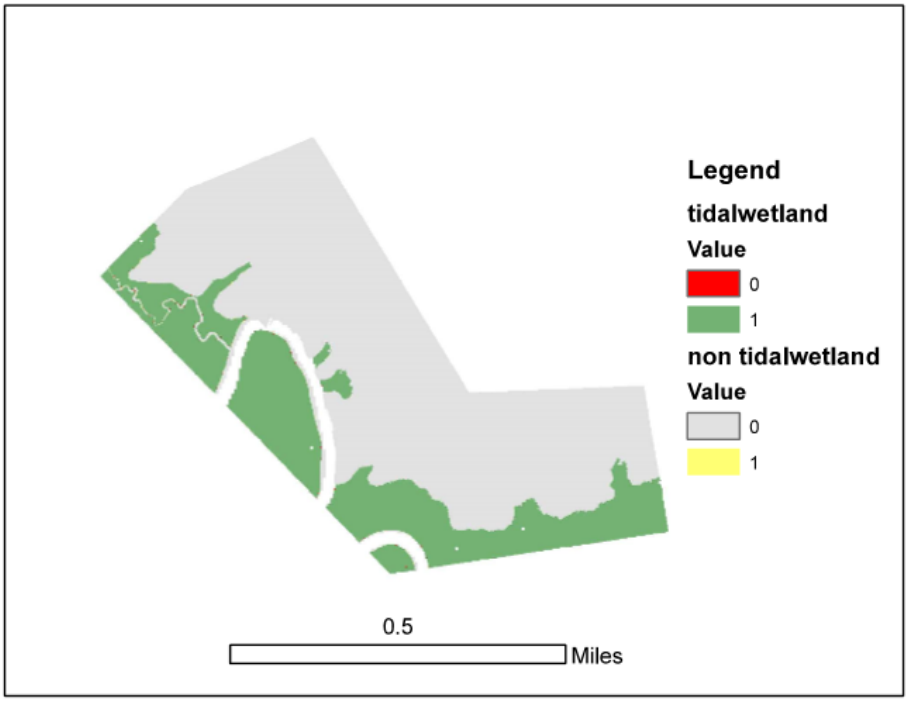

Figure 13. Map for RF with the recommended resolution
[Green color for 1-1, i.e., tidal-wetland – predicted tidal-wetland as correct; Red color for 1-0, i.e., tidal-wetland – predicted non-tidal-wetland as missing; Grey color for 0-0, i.e., non-tidal-wetland – predicted non-tidal-wetland as correct; Yellow color for 0-1, i.e., as over predicted tidal-wetland.]

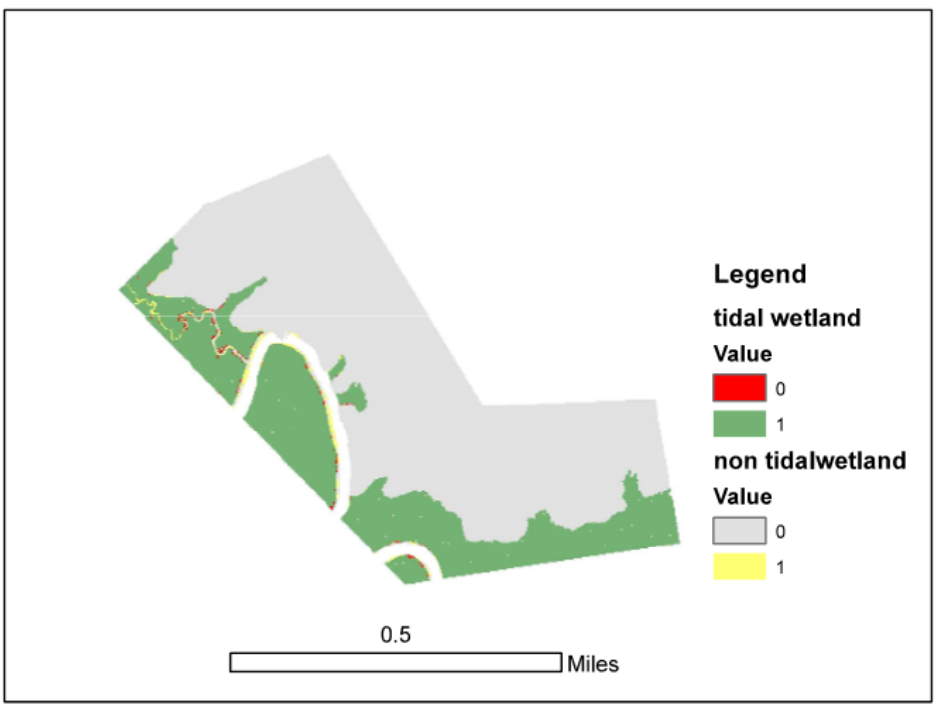

Figure 14. Map for Logit with the recommended resolution

7. Conclusion

This project mainly focuses on the following major objectives:

(a) To develop an effective predictor variable set for tidal wetland prediction;

(b) To develop effective methods for modeling tidal wetlands by using QL2 LiDAR data;

(c) To develop new automated practical tools for tidal wetland identification and prediction by using QL2 LiDAR data based on the developed methods;

(d) To run a field test to validate and evaluate the developed methods and tools;

(e) To develop the best resolution determination method; and

(f) To have deliverable automated tidal wetland prediction tools.

According to the results, we summarize this project completion status as follows:

(1) We have successfully completed this important project for the NCDOT needs of tidal wetland modeling and prediction.

(2) During this project period, we have further developed and updated our WAMAT (patent supported) to v.4.4 with extended functions and easy run interface, which is easy to install and user-friendly to use with a full process automation and/or a module process automation as user's choice. That helps the development of tools for tidal wetland prediction.

(3) We have successfully developed tidal wetland prediction automation tools, WAMAT-Tidal, as a deliverable product for NCDOT to use internally. The Users' Guide of WAMAT-Tidal is also ready for deliverable with the tools together.

(4) Two systematic models are presented and developed with the automation. They are logistic regression model (Logit) and Random Forest model (RF).

(5) The models with automation have been applied to predict wetlands and tidal wetlands for Brunswick County. The resultant data and digital maps are delivered to NCDOT as attachment [10.E].

(6) A field visit to Brunswick and New Hanover counties has been conducted with Axiom

Environmental support. Our prediction results are mainly based on the QL2 data with soil and TIZ data, which may change over time. But the tools can be run based on updated data.

(7) Further research and study in this important research area and direction is needed to advance our developed system and the NCDOT's excellent NC WAM work to continue leading in the nation.

The deliverable products include:

(i) WAMAT-Tidal v.4.1,

(ii) WAMAT-Tidal v.4.1 Users' Guide,

(iii) WAMAT v.4.4, (as well as v.4.1 ~ 4.3),

(iv) WAMAT v.4.4 Users' Guide, (as well as v.4.1 ~ 4.3 User's Guides),

(v) Systematic Logit model and RF model for tidal wetland prediction in automation tools,

(vi) Digital tidal wetland maps from the above models for Brunswick County regions,

(vii) WAMAT v.5.1, and

(viii) WAMAT v.5.1 Users' Guide.

During this project period, we have published 3 papers as listed in the next Chapter. Among them are one at the International Conference on Ecology and Transportation, held in Raleigh, NC, 2015 [8.1], and another two at Transportation Research Board (TRB) Annual Meetings, 2017 and 2018 respectively [8.2 – 8.3].

The PI and NCDOT were invited to present our research of NCDOT RP 2013-13 with demos as the 2015 Sweet Sixteen High Value Research awarded project at the 2016 TRB Annual Meeting in Washington, D.C., January 2016 [8.4].

Furthermore, just recently, our research result has led to a US Patent issued by USPTO on 07-17-2018 [8.5].

A summary for that is listed in the next Chapter as follows.

8. Papers Published and Patent Awarded in the Project Period

Published Papers and/or Presentations:

[8.1] S.-G. Wang, J. Deng, M.-Z. Chen, M. Weatherford, and L. Paugh, "Random Forest Classification and Automation for Wetland Identification based on DEM Derivatives", *2015 ICOET* (*International Conference on Ecology and Transportation*), paper 778, session 408-2, Raleigh, US, pp.1-17, Sept. 2015.

[8.2] J. Deng, A.S. Smith, S. Davis, M. Weatherford, L. Paugh, and S.-G. Wang*, "Identification of NC Wetland Types by Lidar Data and Tree Based Machine Learning Methods", *the TRB 96th Annual Meeting, National Academies of Sciences-Engineering-Medicine*, Paper No.17-01199, pp.1-16, Washington DC, Jan. 2017.

[8.3] J. Deng, S.-G. Wang*, A.S. Smith, S. Davis, M. Weatherford, L. Paugh, and S. Jiang, "Scale Analysis of a Wetland Classification Model based on LiDAR Data and Machine Learning Methodology", TRB 97th Annual Meeting, National Academies of Sciences-Engineering-Medicine, Paper No. 18-01812, pp.1-16, Washington DC, Jan. 2018.

[8.4] S.-G. Wang†, M. Weatherford, L. Paugh, N. Mastin†, and J. Kirby, "Improvements to NCDOT's Wetland Prediction Model", *State Department of Transportation High Value Research, 2015 AASHTO-RAC Awarded Sweet 16 High Value Research Project, at the TRB 95th Annual Meeting, National Academies of Sciences-Engineering-Medicine*, Washington DC, Jan. 10-14, 2016. (†Invited Presenters/Speakers)

* Corresponding Author

US Patent issued by USPTO:

[8.5] S.-G. Wang, L. Bai, J. Deng, M. Jia, M. Weatherford, L. Paugh, W. Tang, M. Chen and S. Chen, "Wetland Modeling and Prediction", Invention Documents, UNC Charlotte, April 18, 2014. US 10,026,221, 07-17-2018. (14/724,787, 05-28-2015)

9. References

[1] S.-G. Wang, A.P. Smith, and S. Davis, “NCDOT Wetland Modeling Program: Development of Tidal Wetland Models using QL2 LiDAR”, Proposal of NCDOT RP 2016-19, 2014.

[2] S.-G. Wang, L. Bai, J. Deng, M. Jia, M. Weatherford, L. Paugh, W. Tang, M. Chen and S. Chen, “Automation Process Method of Generating Wetland Predictive Variables”, Invention Documents, UNC Charlotte, April 18, 2014. (USPTO 62/003,869, 05/28/2014)

[3] S.-G. Wang, L. Bai, L. Paugh and M. Weatherford, “Automation Process Method of Wetland Modeling and Prediction”, Invention Documents, UNC Charlotte, April 18, 2014. (USPTO 62/003,887, 05/28/2014)

[4] S. Davis and A.P. (Sandy) Smith, NCWAM North Carolina Wetland Type Flow-Chart, Axiom Notes, March 2014.

[5] S.-G. Wang (PI), “Draft Final Report on Project – Improvements to NCDOT’s Wetland Prediction Model”, NCDOT RP 2013-13, UNCC, 08-14-2014.

[6] M. Weatherford and P. Harris, (2014). “NCDOT Wetland Modeling Program: Development of Tidal Wetland Models using QL2 Lidar”, Call for New Research Needs, 6107, NC DOT, 2014.

[7] H. Morgan and W. Draper, (2014). “Living LiDAR Data Collection at QL2”, http://proceedings.esri.com/library/userconf/proc14/papers/386_161.pdf. April 9, 2014.

[8] FHWA, 2011 Environmental Excellence Awards, GIS-based Wetland and Stream Predictive Models – For Excellence in Environmental Research, http://environment.fhwa.dot.gov/eea2011/environment_research.htm, 2011.

[9] National Environmental Policy Act (NEPA), US Environmental Law, 1970.

[10] N.C. Wetland Functional Assessment Team, (2010). Wetland Assessment Method (NC WAM) User Manual, version 4, Oct. 2010.

[11] Environmental Laboratory, (1987). “Corps of Engineers Wetlands Delineation Manual”, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS, 1987.

[12] Environmental Laboratory, (2010). “Regional Supplement to the Corps of Engineers Wetland Delineation Manual: Atlantic and Gulf Coastal Plain Region”, U.S. Army Engineer Research and Development Center, Vicksburg, MS, 2010.

[13] North Carolina Floodplain Mapping Program, www.ncfloodmaps.com, Retrieved 06-06-2014.

[14] R.K. Roy, (2001). Design of experiments using the Taguchi approach: 16 steps to product and process improvement, John Wiley & Sons.

[15] L. Breiman, (2001). "Random forests", Machine learning, vol. 45, pp. 5-32.

[16] Allen, T., Wang, Y., Gore, B., Swords, J., & Newcomb, D. (2011). Coastal Wetland mapping Using Time Series SAR Imagery and LiDAR: Alligator River National Wildlife Refuge, North Carolina. Paper presented at the Proceedings, Pecora 18 symposium, Herndon, Virginia.

[17] Baker, C., Lawrence, R., Montagne, C., & Patten, D. (2006). Mapping wetlands and riparian areas using Landsat ETM+ imagery and decision-tree-based models. Wetlands, 26(2), 465-474.

[18] Boyd, J. (2002). Compensating for Wetland Losses under the Clean Water Act. Environment: Science and Policy for Sustainable Development, 44(9), 43-44.

[19] Castañeda, C., & Ducrot, D. (2009). Land cover mapping of wetland areas in an agricultural landscape using SAR and Landsat imagery. Journal of Environmental Management, 90(7), 2270-2277.

[20] D. E. Chapple, P. Faber, K.N. Suding, and A.M. Merenlender, (2017). "Climate Variability Structures Plant Community Dynamics in Mediterranean Restored and Reference Tidal Wetlands," *Water*, vol. 9, p. 209, 2017.

[21] Connor, W.H., K.W. Krauss, and T.W. Doyle. (2007). Ecology of Freshwater Forests in Coastal Deltaic Lousiana and Northeastern South Carolina. pp. 223-253 *in* Ecology of Tidal Freshwater Forested Wetlands of the Southeastern United States, Conner, W.H., T.W. Doyle, and K.W. Krauss (eds.)

[22] Copeland, B.J., R.G. Hodson, S.R. Riggs, and J.E. Easley, Jr. (1983). The ecology of Albemarle Sound, North Carolina: an estuarine profile. U.S. Fish and Wildlife Service, Division of Biological Services, FWS/OBS-83/01. 68 pp.

[23] Corcoran, J.M., Knight, J.F., & Gallant, A.L (2013). Influence of Multi-Source and Multi-Temporal Remotely Sensed and Ancillary Data on the Accuracy of Random Forest Classification of Wetlands in Northern Minnesota, Remote Sensing, 5(7), 3212-3238.

[24] Design of Experiments (DOE) Using the Taguchi Approach, http://nutek-us.com/DOE_Topic_Overviews35Pg.pdf

[25] EPA, U.S. (2012, Octorber 05, 2012). Section 404 of the Clean Water Act: how wetlands are defined and identified. Retrieved 0529, 2014, from http://water.epa.gov/type/wetlands/outreach/fact11.cfm.

[26] S. Fagherazzi, M.L. Kirwan, S.M. Mudd, G.R. Guntenspergen, S. Temmerman, A. D'Alpaos, *et al.*, (2012). "Numerical models of salt marsh evolution: Ecological, geomorphic, and climatic factors," *Reviews of Geophysics,* vol. 50, 2012.

[27] Freeman, E., Frescino, T., & Moisen, G. (2009). ModelMap: An R package for modeling and map production using Random Forest and Stochastic Gradient Boosting. USDA Forest Service, Rocky Mountain Research Station, 507.

[28] Friedman, J.H. (2002). Stochastic gradient boosting. Computational Statistics & Data Analysis, 38(4), 367-378.

[29] Friedrichs, C.T, and D.B. Aubrey. (1988). Non-linear tidal distortion in shallow well-mixed estuaries: a synthesis. Estuarine, Coastal and Shelf Science, v. 27: 521-545.

[30] Genç, L., B. Dewitt, and S. Smith. (2004). “Determination of wetland vegetation height with LiDAR.” Turkish Journal of agriculture and forestry, v.28(1), pp.63-71.

[31] Giese, G.L, H.B. Wilder, and G.G. Parker. (1979). Hydrology of major estuaries and sounds of North Carolina. U.S. Geological Survey, Water Resources Investigations 79-46. 175 pp.

[32] Gislason, P.O., Benediktsson, J.A., & Sveinsson, J.R. (2006). Random forests for land cover classification. Pattern Recognition Letters, 27(4), 294-300.

[33] Gorham, E. (1991). Northern peatlands: role in the carbon cycle and probable responses to climatic warming. Ecological applications, 1(2), 182-195.

[34] Hackney, C.T., G.B. Avery, L.A. Leonard, Martin Posey, and Troy Alphin. (2007). Biological, Chemical, and Physical Characteristics of Tidal Freshwater Swamp Forests of the Lower Cape Fear River/Estuary, North Carolina. Pp. 183-221 *in* Ecology of Tidal Freshwater Forested Wetlands of the Southeastern United States, Conner, W.H., T.W. Doyle, and K.W. Krauss (eds.).

[35] Henderson, F.M., & Lewis, A.J. (2008). Radar detection of wetland ecosystems: a review. International Journal of Remote Sensing, 29(20), 5809-5835.

[36] Hess, L.L., Melack, J.M., Novo, E.M., Barbosa, C.C., & Gastil, M. (2003). Dual-season mapping of wetland inundation and vegetation for the central Amazon basin. Remote Sensing of Environment, 87(4), 404-428.

[37] Hogg, A., & Todd, K. (2007). Automated discrimination of upland and wetland using terrain derivatives. Canadian Journal of Remote Sensing, 33(S1), S68-S83.

[38] Hope, Morgan, and Draper Will. "Living LiDAR Data Collection at QL2 - Proceedings.esri.com." Web. 04 Feb. 2016. http://proceedings.esri.com/library/userconf/proc14/papers/386_161.pdf.

[39] James, B.V, C.C. Trettin, and T.J. Callahan. (2012). Hydrologic influences within a tidal freshwater forested wetland. Proceedings of the 2012 South Carolina Water Resources Conference, held October 10-11, 2012 at the Columbia Metropolitan Convention Center. 6 pp

[40] Johnston, Keith. "Quality Level 2 (QL2) LIDAR Utilization at NCDOT." NC MicroStation Local Users Group, 16 Dec. 2015. Web. 23 Mar. 2016. http://www.nclug.com/uploads/2/8/0/0/2800961/ql2_lidar_to_nclug_4-3_20151216.pdf

[41] Keddy, P. A. (2010). Wetland ecology: principles and conservation: Cambridge University Press.

[42] Kennedy, G., & Mayer, T. (2002). Natural and constructed wetlands in Canada: An overview. Water Quality Research Journal of Canada, 37(2), 295-325.

[43] Kirwan, M. L. and J. P. Megonigal, (2013). "Tidal wetland stability in the face of human impacts and sea-level rise," Nature, vol. 504, pp. 53-60, 2013.

[44] Lawrence, R., Bunn, A., Powell, S., & Zambon, M. (2004). Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. Remote Sensing of Environment, 90(3), 331-336.

[45] Lawrence, R.L., & Wright, A. (2001). Rule-based classification systems using classification and regression tree (CART) analysis. Photogrammetric Engineering and Remote Sensing, 67(10), 1137-1142.

[46] Leck, M.A., A.H. Baldwin, V.T. Parker, Lisa Schile, and D.F. Whigham. (2009). Plant communities of tidal freshwater wetlands of the continental USA and Canada. Pp. 41-58 *in* Tidal Freshwater Wetlands, Barendregt, A, D.F. Whigham, A.H. Baldwin (eds.). Backhuys Publishers, Leiden, The Netherlands.

[47] Li, J., & Chen, W. (2005). A rule-based method for mapping Canada's wetlands using optical, radar and DEM data. International Journal of Remote Sensing, 26(22), 5051-5069.

[48] Loh, W. Y. (2011). Classification and regression trees. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 1(1), 14-23.

[49] Nuttle, W.K. (1988). The extent of lateral water movement in the sediments of a New England salt marsh. Water Resources Research, Vol 24 (12): 2077-2085.

[50] Odum, W.E., T.J. Smith III, J.K. Hoover,and C.C. McIvor. (1984). The ecology of tidal freshwater marshes of the United States east coast: a community profile. U.S. Fish and Wildlife Service, FWS/OBS-83/17. 177 pp.

[51] Olhan, E., Gun, S., Ataseven, Y., & Arisoy, H. (2010). Effects of agricultural activities in Seyfe Wetland. Scientific Research and Essays, 5(1), 9-14.

[52] Pal, M., & Mather, P. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sens. Environ, 86, 554–565.

[53] Rebelo, L.-M., Finlayson, C., & Nagabhatla, N. (2009). Remote sensing and GIS for wetland inventory, mapping and change analysis. Journal of Environmental Management, 90(7), 2144-2153.

[54] Rodríguez-Galiano, V., Abarca-Hernández, F., Ghimire, B., Chica-Olmo, M., Atkinson, P., & Jeganathan, C. (2011). Incorporating spatial variability measures in land-cover classification using Random Forest. Procedia Environmental Sciences, 3, 44-49.

[55] Team, R. D. C. (2012). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2012: ISBN 3-900051-07-0.

[56] Thompson, Gary. "North Carolina LiDAR - Federal Geographic Data Committee." Mar. 2015. Web. 23 Mar. 2016. <https://www.fgdc.gov/ngac/meetings/march-2015/north-carolina-lidar-ngac-march-2015.pdf>

[57] Tiner, R.W. (2013). Tidal Wetlands Primer: An Introduction to Their Ecology, Natural History, Status, and Conservation. University of Massachusetts Press, Amherst and Boston, 508 pp.

[58] Toner, M., & Keddy, P. (1997). River hydrology and riparian wetlands: a predictive model for ecological assembly. Ecological applications, 7(1), 236-246.

[59] Wang, L., Lyons, J., Kanehl, P., & Bannerman, R. (2001). Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental Management, 28(2), 255-266.

[60] Wang, S.-G., Bai, L., Deng, J., Jia, M., Weatherford, M., Paugh, L., Tang, W., Chen, M., & Chen, S. (2015). Wetland Modeling and Prediction, Patent application, US 14/724,787, USPTO, 05-28-2015.

[61] Weatherford, Morgan. "NCDOT Wetland Modeling Program Updates." 16 Oct. 2014. Web. 04 Feb. 2016. <https://connect.ncdot.gov/resources/Environmental/Merger Process Meetings/NCDOT LiDAR Presentation.pdf>

[62] Weatherford, M. and Harris, P. (2015). Improvements to NCDOT's Wetland Prediction Model, Call for New Research Needs, 3115, NC DOT, 2011.

[63] Wiegert, R. G. and B.J. Freeman. 1990. Tidal salt marshes of the southeast Atlantic coast: a community profile. U.S. Fish and Wildlife Service Biological Report 85 (7.29), 70 pp.

[64] Wright, C., & Gallant, A. (2007). Improved wetland remote sensing in Yellowstone National Park using classification trees to combine TM imagery and ancillary environmental data. Remote Sensing of Environment, 107(4), 582-605.

[65] Zomer, R., Trabucco, A., & Ustin, S. (2009). Building spectral libraries for wetlands land cover classification and hyperspectral remote sensing. Journal of Environmental Management, 90(7), 2170-2177.

[66] J. Deng, A.S. Smith, S. Davis, M. Weatherford, L. Paugh, and S.-G. Wang*, (2017). "Identification of NC Wetland Types by Lidar Data and Tree Based Machine Learning Methods", the TRB 96th Annual Meeting, National Academies of Sciences-Engineering-Medicine, Paper No.17-01199, pp.1-16, Washington DC, Jan. 2017.

[67] J. Deng, S.-G. Wang*, A.S. Smith, S. Davis, M. Weatherford, L. Paugh, and S. Jiang, (2018). "Scale Analysis of a Wetland Classification Model based on LiDAR Data and Machine Learning Methodology", TRB 97th Annual Meeting, National Academies of Sciences-Engineering-Medicine, Paper No. 18-01812, pp.1-16, Washington DC, Jan. 2018. http://amonline.trb.org/2017trb-1.3983622/t010-1.3999179/514-1.3999392/18-01812-1.3999398/18-01812-1.3999399?qr=1

[68] S.-G. Wang, J. Deng, M.-Z. Chen, M. Weatherford, and L. Paugh, (2015). "Random Forest Classification and Automation for Wetland Identification based on DEM Derivatives", *2015 ICOET* (*International Conference on Ecology and Transportation*), paper 778, session 408-2, Raleigh, US, pp.1-17, Sept. 2015.

[69] S.-G. Wang†, M. Weatherford, L. Paugh, N. Mastin†, and J. Kirby, (2016). "Improvements to NCDOT's Wetland Prediction Model", State Department of Transportation High Value Research, 2015 AASHTO-RAC Awarded Sweet 16 High Value Research Project, at the TRB 95th Annual Meeting, National Academies of Sciences-Engineering-Medicine, Washington DC, Jan. 10-14, 2016. (†Invited Presenters/Speakers)

[70] S.-G. Wang, L. Bai, J. Deng, M. Jia, M. Weatherford, L. Paugh, W. Tang, M. Chen, and S. Chen, (UNCC & NCDOT) (2018). "Wetland Modeling and Prediction", US Patent 10,026,221 07-17-2018.

*Corresponding Author

10. Appendix – Deliverables (submitted separately)

[A] WAMAT-Tidal v.4.1, (2018). S.-G. Wang (PI) and S. Jiang. (Delivered in Oct. 2018)

[B] WAMAT-Tidal Users' Guide, v.4.1, (2018). S.-G. Wang (PI) and S. Jiang. (Delivered in Oct. 2018)

[C] WAMAT (WAM Automation Tools) v.4.4, (10-05-2017). S.-G. Wang (PI) and S. Jiang. (Delivered in Oct. 2017)

[D] WAM Automation Tools (WAMAT) – QUICK START GUIDE, v.4.4, (10-05-2017). S.-G. Wang (PI), and S. Jiang. (Delivered in Oct. 2017)

[E] Field Test Validation at Brunswick County for Tidal Wetland Prediction, (2018). Sheng-Guo Wang (PI), Sandy Smith, Scott Davis, Shanshan Jiang, and Yinan He. (Delivered in Oct. 2018)

[F] *WAMAT (WAM Automation Tools) v.5.1, (09-21-2018). S.-G. Wang (PI) and S. Jiang. (Delivered in Oct. 2018)

[G] *WAM Automation Tools (WAMAT) – QUICK START GUIDE, v.5.1, (10-04-2018). S.-G. Wang (PI), and S. Jiang. (Delivered in Oct. 2018)

[H] Tidal Influence Zone Dataset, Scott Davis, Axiom, 2018.