

16 - 0713

Project Documents for the National Wetlands Inventory Update in Minnesota – Phase 2

In Partial Fulfillment of the Environment and Natural Resources Trust Fund 2010 Work Program Final Report

July 31, 2011



Contents

<u>Tab</u>	<u>Description</u>
1	Abstract, Final Report, and Budget
2	News Articles
3	Publications <p style="margin-left: 40px;">Kloiber, S.M., Macleod, R.D., Smith, A.J., Knight, J.F., and Huberty, B.J. (2014) A semi-automated, multi-source data fusion update of a wetland inventory for east-central Minnesota, USA. <i>Wetlands, in review.</i></p> <p style="margin-left: 40px;">Corcoran, J.M., Knight, J.F., B. Brisco, S. Kaya, A. Cull, and Murhnaghan, K. (2011) The integration of optical, topographic, and radar data for wetland mapping in northern Minnesota. <i>Canadian Journal of Remote Sensing</i>, 27(5): 564-582.</p> <p style="margin-left: 40px;">Corcoran, J.M., Knight, J.F., and 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. <i>Remote Sensing</i>, 5(7): 3212-3238.</p> <p style="margin-left: 40px;">Knight, J.F., B. Tolcser, J. Corcoran, and Rampi, L. (2013) The effects of data selection and thematic detail on the accuracy of high spatial resolution wetland classifications. <i>Photogrammetric Engineering and Remote Sensing</i>, 79(7): 613-623.</p> <p style="margin-left: 40px;">Rampi, L.P., Knight, J.F., and Pelletier, K.C. (2014) Wetland mapping in the Upper Midwest United States: An object-based approach integrating lidar and imagery data. <i>Photogrammetric Engineering and Remote Sensing</i>. 80(5): 439-449.</p> <p style="margin-left: 40px;">Rampi, L.P., Knight, J.F., and Lenhart, C.F. (2014) Comparison of flow direction algorithms in the application of the CTI for mapping wetlands in Minnesota. <i>Wetlands</i>, 34(3): 515-525.</p>
Back Pocket	Example wetland map that can be produced from updated National Wetland Inventory data

2010 Project Abstract

For the Period Ending June 30, 2014

PROJECT TITLE: Updating the National Wetlands Inventory in Minnesota

PROJECT MANAGER: Steve Kloiber

AFFILIATION: Minnesota Department of Natural Resources

MAILING ADDRESS: 500 Lafayette Road North, Box 25

CITY/STATE/ZIP: St. Paul, MN 55155

PHONE: 651-259-5164

E-MAIL: steve.kloiber@state.mn.us

WEBSITE: http://www.dnr.state.mn.us/eco/wetlands/nwi_proj.html

FUNDING SOURCE: Environment and Natural Resources Trust Fund

LEGAL CITATION: M.L. 2010, Chp. 362, Sec. 2, Subd. 3b and M.L. 2013, Chapter 52, Section 2, Subdivision 17

APPROPRIATION AMOUNT: \$1,100,000

Overall Project Outcome and Results

Updated wetland maps were created for 13 counties in east-central Minnesota (7,150 square miles), encompassing the Twin Cities metropolitan area. Wetlands in Minnesota were originally mapped by the U.S. Fish and Wildlife Service in early 1980's as part of the National Wetlands Inventory (NWI). Although still widely used for land use planning, wetland permit screening and natural resource management, the original maps have grown increasingly out-of-date due to landscape alterations over the years. The data created for this project marks the first significant update to the NWI in Minnesota.

The new maps are much more accurate, capture more detail and provide more information than the original maps. Besides showing the location, size and type of each wetland, the updated map data include information on the wetland's landscape position and hydrologic characteristics, which can be useful in assessing the benefits provided, such as water quality improvement, flood storage, and fish and wildlife habitat. Updating the NWI is a key component of the State's strategy to monitor and assess wetlands in support of efforts to assure healthy wetlands and clean water for Minnesota. The DNR is planning to complete the NWI update for the entire state by 2020.

Accomplishments for this project phase also include; acquiring high-resolution, spring, leaf-off, digital aerial imagery for 23,900 square miles of southern Minnesota, acquiring field validation data for southern Minnesota, and developing wetland mapping procedures for the agricultural region of Minnesota.

Project Results Use and Dissemination

Imagery acquired for this project is available to the public through the Minnesota Geospatial Information Office (http://www.mngeo.state.mn.us/chouse/wms/geo_image_server.html). The MnGeo imagery service receives about one million page requests per month for the southern Minnesota imagery. This is the first publicly available leaf-off imagery data for southern Minnesota since 1991.

The updated wetland map data are available through an interactive mapping application on the DNR's website at: <http://www.dnr.state.mn.us/eco/wetlands/map.html>. The data can also be downloaded, free of charge, for use in geographic information system applications through the DNR's data deli at: <http://deli.dnr.state.mn.us/>. The data will eventually be incorporated into the national "Wetland Mapper" application maintained by the U.S. Fish and Wildlife Service.

The wetland mapping procedures and accuracy results for the 13-county updated NWI data are presented and discussed in a manuscript that has been submitted to the journal *Wetlands*, a publication of the Society of Wetland Scientists (SWS). Information from this project was also presented at the SWS annual conference in Duluth, MN in 2013. In addition, a press release was distributed regarding the updated NWI data and the story was published on several online news websites.

Researchers at the University of Minnesota Remote Sensing and Geospatial Analysis Laboratory conducted an extensive study of the effects of digital elevation model (DEM) preprocessing and mapping methods on the accuracy of wetlands maps in three different physiographic regions of Minnesota. This research covered two study sites in agricultural areas including the Minnesota River Headwaters (Big Stone County) and Swan Lake (Nicollet County) as well as a comparison site from northern Minnesota (St. Louis and Carlton Counties). The results of this effort were compiled and submitted for publication in several peer-reviewed scientific journals along with results from the earlier phase of the NWI update project. Three hard copies and one electronic copy of these publications have been submitted with the final report to LCCMR. There have also been numerous presentations at professional conferences.

Publications

Corcoran, J.M., Knight, J.F., B. Brisco, S. Kaya, A. Cull, and Murhnaghan, K. (2011) The integration of optical, topographic, and radar data for wetland mapping in northern Minnesota. *Canadian Journal of Remote Sensing*, 27(5): 564-582.

Corcoran, J.M., Knight, J.F., and 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.

Knight, J.F., B. Tolcser, J. Corcoran, and Rampi, L. (2013) The effects of data selection and thematic detail on the accuracy of high spatial resolution wetland classifications. *Photogrammetric Engineering and Remote Sensing*, 79(7): 613-623.

Rampi, L.P., Knight, J.F., and Pelletier, K.C. (2014) Wetland mapping in the Upper Midwest United States: An object-based approach integrating lidar and imagery data. *Photogrammetric Engineering and Remote Sensing*, 80(5): 439-449.

Rampi, L.P., Knight, J.F., and Lenhart, C.F. (2014) Comparison of flow direction algorithms in the application of the CTI for mapping wetlands in Minnesota. *Wetlands*, 34(3): 515-525.

Environment and Natural Resources Trust Fund (ENRTF) 2010 Final Report

Date of Report: July 31, 2014
Final Report July 31, 2014
Date of Work Program Approval: January 5, 2010
Project Completion Date: June 30, 2014

I. PROJECT TITLE: Updating the National Wetlands Inventory: Phase 2

Project Manager: Steve Kloiber
Affiliation: Minnesota Dept. of Natural Resources
Mailing Address: 500 Lafayette Road, Box 25
City / State / Zip: St. Paul, MN 55155
Telephone Number: 651-259-5164
E-mail Address: steve.kloiber@state.mn.us
Fax Number: 651-296-1811
Web Site Address: http://www.dnr.state.mn.us/eco/wetlands/nwi_proj.html

Location: This phase of the project focuses on updating the National Wetland Inventory (NWI) maps for a 13-county area in east-central Minnesota surrounding the greater metropolitan region of Minneapolis and St. Paul (figure 1). This phase includes a pilot study to test wetland mapping methods for an as yet unspecified agricultural site in southern Minnesota. In addition, this phase also includes primary imagery acquisition and field validation data collection for southern Minnesota.

Total ENRTF Project Budget:	ENRTF Appropriation	\$ 1,100,000
	Minus Amount Spent:	\$ 1,100,000
	Equal Balance:	\$ 0

Legal Citation: M.L. 2010, Chp. 362, Sec. 2, Subd. 3b and M.L. 2013, Chapter 52, Section 2, Subdivision 17

Appropriation Language:

\$1,100,000 is from the trust fund to the commissioner of natural resources to continue the update of wetland inventory maps for Minnesota. The availability of the appropriation for the following project is extended to June 30, 2014: (3) Laws 2010, chapter 362, section 2, subdivision 3, paragraph (b), Updating Minnesota Wetlands Inventory: Phase 2. This appropriation is available until June 30, 2014 by which time the project must be completed and final products delivered.

II. FINAL PROJECT SUMMARY AND RESULTS:

Updated wetland maps were created for 13 counties in east-central Minnesota (7,150 square miles), encompassing the Twin Cities metropolitan area. Wetlands in Minnesota were originally mapped by the U.S. Fish and Wildlife Service in early 1980's as part of the National Wetlands Inventory (NWI). Although still widely used for land use planning, wetland permit screening and natural resource management, the original maps have

grown increasingly out-of-date due to landscape alterations over the years. The data created for this project marks the first significant update to the NWI in Minnesota.

The new maps are much more accurate, capture more detail and provide more information than the original maps. Besides showing the location, size and type of each wetland, the updated map data include information on the wetland's landscape position and hydrologic characteristics, which can be useful in assessing the benefits provided, such as water quality improvement, flood storage, and fish and wildlife habitat. Updating the NWI is a key component of the State's strategy to monitor and assess wetlands in support of efforts to assure healthy wetlands and clean water for Minnesota. The DNR is planning to complete the NWI update for the entire state by 2020.

Accomplishments for this project phase also include; acquiring high-resolution, spring, leaf-off, digital aerial imagery for 23,900 square miles of southern Minnesota, acquiring field validation data for southern Minnesota, and developing wetland mapping procedures for the agricultural region of Minnesota.

III. PROGRESS SUMMARY AS OF:

July 31, 2014

The DNR Resource Assessment Office has completed the data processing for the LiDAR digital elevation model (DEM) and soils data for the remainder of the state. All tasks for this project have been completed.

January 31, 2014

The remaining imagery for southern Minnesota was acquired, processed, and delivered. The final quality assessment found the imagery data to meet the project requirements. The data was delivered to the State and is posted on the MnGeo web service for public access. With this, all of the tasks in the original work program have been completed.

The remaining budget was redirected to assist with completion of the data processing for the remainder of the state. This additional task is expected to be complete by mid-April 2014.

Amendment Request (9/9/13) – Approved (9/11/13)

A variety of cost savings were achieved for this project such that the project has an overall balance of \$10,175.22 after accounting for the remaining anticipated invoices. We received approval for an amendment to direct the remaining funds into an effort to accelerate the overall project by completing the data pre-processing for the remainder of the state. Previously, this work was done on a phase-by-phase approach. The DNR Resource Assessment Office (RA) will process the LiDAR and soils data to create a number of derivative data sets for the update of the NWI. The LiDAR derived products include slope, topographic position index, and the compound topographic index. The derived soils data include percent hydric soils and the water regime class. This work is already complete for east-central and southern Minnesota. With phase 4 of the project just getting underway, RA will be pre-processing the LiDAR and soils data for northeastern MN. This work program request will partially fund the data preprocessing for the rest of the state. The total additional cost will be \$19,171. We have proposed to

use the remaining \$10,175.22 from this grant and an additional \$8995.78 from savings achieved under the grant for phase 3 of this project.

July 31, 2013:

The remaining imagery for southern Minnesota was acquired and the processing is about 90% complete. Final processing and quality assessment will be performed in FY14. This is the only outstanding deliverable for this project. We anticipate that the project should be 100% complete by October of 2014. Ducks Unlimited completed all contracted tasks and provided a complete set of deliverables. The DNR reviewed these deliverables and conducted the final accuracy assessment on the data. The data have been publicly posted on the DNR Data Deli (<http://deli.dnr.state.mn.us/>) and a copy of the data has also been provided to the US Fish and Wildlife Service for posting on their Wetland Mapper website (<http://www.fws.gov/wetlands/Data/Mapper.html>). The University of Minnesota Remote Sensing and Geospatial Analysis Laboratory (RSGAL) delivered updated wetland map data for all the pilot areas along with a complete wetland mapping protocol document for southern Minnesota. Additionally, the University of Minnesota hosted a wetland mapping methods workshop with the mapping vendor for southern Minnesota and other project stakeholders.

We have received two additional invoices for a total of \$28,456 that are not reflected in the financial report yet. Also, there will be another invoice from MnGeo for the remaining imagery acquisition costs of \$20,016. The expected ending balance is \$10,175. This reflects cost savings achieved on the project. We anticipate submitting an amendment to use the remaining funds to further advance the statewide NWI update.

Amendment Request (1/25/13) – Approved (5/9/13): A weather-related delay in imagery acquisition required a work program amendment. In 2011, imagery was acquired for 35 out of 36 counties during the targeted spring, leaf-off period. Mop-up imagery acquisition operations scheduled for 2012 were scuttled by an early onset of very warm temperatures in March coupled with an extended period of overcast skies. We are requested a one year extension, contingent on Legislative approval to allow the contractor to complete the necessary data processing and deliver the imagery to the state.

January 31, 2013: Ducks Unlimited (DU) has completed the photo-interpretation process for 437 out of 541 quarter quads (81%) in the East Central project area. DU's internal QA/QC process has been completed on 433 out of 541 quarter quads (80%) in the east-central project area. The DNR has conducted a secondary review on 317 quarter quads (59%) for this project area. Comments from the DNR have been incorporated by DU. DU finished the scripting for the plant community and hydrogeomorphic classifications and tested the results and has developed the input layers needed for the HGM classification for the entire East Central project area. RAP has completed draft NWI data for all 50 quarter quads in the Koochiching project area. The primary and secondary QA/QC have been completed for the Koochiching project area. Final revisions are in process.

The University of Minnesota RSGAL hosted a technical workshop on wetland mapping for the southern Minnesota NWI update project. RSGAL researchers continue to study the effects of DEM preprocessing and mapping algorithm choice on the accuracy of

wetlands maps. The results of this study have been compiled and are being prepared for publication in a peer reviewed scientific journal and will be included in final report to LCCMR. RSGAL researchers continued efforts at optimizing mapping methods for wetlands in agriculture dominated portions of Minnesota. Findings from this research will be incorporated into NWI mapping methods.

July 31, 2012

Ducks Unlimited (DU) has completed the image segmentation for 80% of the east-central Minnesota (ECMN) project area. DU has also completed 50% of the initial photo-interpretation and 33% of the ECMN project area has received an internal review for QA/QC. Data that have passed the internal review process are delivered to the DNR for review. The DNR has reviewed updated wetland maps for about 15% of the ECMN. DU is currently developing the computer code to generate additional wetland attributes for the wetland plant community classification and the hydro-geomorphic classification as well as the model code for the wetland probability layer.

Pre-processing of all data for the Koochiching project area was completed and the image segmentation has also been completed for all 50 quarter quads in this area. Initial image interpretation is nearly complete for all quarter quads in this area and an internal review has been completed for 20 quarter quads (40% complete).

The DNR has expanded the level of effort for QA/QC for the NWI data. This includes re-programming funds saved on data acquisition to support additional QA/QC review of the NWI data. This will be accomplished through a combination of additional DNR staff time as well as soliciting feedback from local wetland experts using a web-based review tool developed by the DNR.

Imagery data that was delivered by Surdex for 35 counties in southern MN has passed the quality control assessment. The data have been accepted and posted to the MnGeo web service for public access. Imagery acquisition for last remaining county in the acquisition area was delayed to spring 2013 due to weather issues this past spring.

The University of Minnesota RSGAL researchers have continued their efforts to study of the effects of DEM preprocessing and mapping algorithm choice on the accuracy of wetlands maps for different physiographic regions. In addition, RSGAL is also studying optimal geospatial data types and mapping methods for wetlands in agriculture dominated portions of Minnesota.

Amendment Request (1/25/12) – Approved (1/26/12)

The cost for the imagery acquisition of southern Minnesota came in 4.7% below the estimated cost. We proposed to re-allocate the savings toward updating of wetland maps (result/activity one). We propose to use \$7000 to provide support for some additional staff time to work on web-based system for reviewing draft maps for accuracy. DNR staff will develop a web-based data review application that will allow the project team, technical advisory committee members, and local wetland experts to provide an efficient and standardized way of reviewing and commenting on draft NWI data. The remaining cost savings of \$14,000 will be applied toward expanded quality control of NWI data.

January 31, 2012

Ducks Unlimited (DU) provided a working draft version of the technical procedures document. This document will be revised as adjustments are made to the method. DU conducted the tech transfer workshop with the DNR Resource Assessment Office RA to harmonize the methods and approach used by the groups working on the NWI update. DNR also participated in other coordination meetings in Duluth, St. Cloud, and Bloomington. DU has developed a website for the work they are doing on this project as well as a web-based status map that shows the state of data production.

RA has completed all of the data preprocessing for 11 of the 13 counties in the east-central Minnesota project area and submitted this to DU. The data for the Koochiching NWI update area has been compiled, but is still being preprocessed. DU submitted draft data for nine quarter quads for DNR review. DNR has reviewed and provided comments on the draft data. Errors have all been addressed by DU. RA conducted follow-up field investigations to provide feedback to DU on difficult to interpret wetlands.

UMN continued to develop and test mapping methods for the Swan Lake pilot area near Mankato, MN and they have added another informal pilot area in Big Stone County. UMN has presented their wetland mapping methods to the U.S. Fish and Wildlife Service staff in the Bloomington, MN regional office. They have also submitted two papers based on these methods for publication.

Draft imagery for southern Minnesota was provided by the aerial photography vendor (Surdex, Inc.). A detailed review of the draft imagery was conducted by DNR, MnGeo, and other project partners. All comments were addressed by the vendor and final imagery was delivered for 35 out of 36 counties. Imagery for the last remaining county in this phase of imagery acquisition (Meeker County) will be acquired in spring 2012. UMN completed the field data acquisition and delivered the data to the DNR for 2703 sites (1722 upland and 981 wetland sites).

It should be noted that while the paid expenditures for this project currently total about 32% of the project budget, the task completion status is close to 50%. The reason for this apparent discrepancy is that tasks must be completed by contractors before invoices can be submitted for payment.

July 31, 2011

Data have been gathered for the east-central project area including aerial photos, radar imagery, soils data, DEMs derived from available LiDAR, original NWI data, and other wetland data. Processing of this data is ongoing. Twelve test areas have been selected and the mapping procedures have been tested, refined, and documented based on these test areas. Draft wetland maps for the 12 test areas (about 50-square miles for each test area) are currently being prepared. Field training data was collected for 510 sites this spring and processed along with other sources of wetland training data for input into the classification model.

Methods evaluation work for the agricultural pilot area was initiated. The pilot area is in the vicinity of Swan Lake near Mankato, MN. Optical and radar imagery data for the

pilot areas have been acquired. Processing of this data has begun. Other data have been ordered.

Acquisition of aerial photography was completed for 35 out of 36 counties. Imagery acquisition was conducted this spring. However, due to weather issues, the acquisition for Meeker County will be delayed until spring 2012. Preliminary image processing has been completed for all of the acquired images. Processing of the ortho-rectified imagery is ongoing. The UMN has hired and trained field staff for collecting field validation data. Acquisition of field data for southern MN was initiated this May.

January 31, 2011

A vendor was selected for updating the NWI maps for east-central Minnesota. A contract was signed this fall and a kick-off meeting was held in December. Efforts have begun to compile and pre-process data for the NWI update. A contract was also signed between the UMN and DNR to address the completion of the methods evaluation work and to collect additional field validation data. The process to select a vendor for acquiring spring, leaf-off imagery for southern Minnesota has begun and the DNR and MnGeo have developed an interagency agreement for assistance with managing this project.

IV. OUTLINE OF PROJECT RESULTS:

RESULT/ACTIVITY 1: Updating Wetland Maps

Description: This component of the project is devoted to updating the NWI maps for 13 counties in east-central Minnesota surrounding the greater Twin Cities metropolitan area (figure 1). The primary task of map production will be contracted out using the State’s standard competitive bid process. The Minnesota DNR will provide oversight for this effort including contract management, supplying primary input data, approving mapping procedures, reviewing map products for accuracy, coordinating stakeholder involvement, and distributing updated NWI data to the public.

Summary Budget Information for Result/Activity 1: ENRTF Budget: \$410,742
 Amount Spent: \$410,742
 Balance: \$ 0

Deliverable/Outcome	Completion Date	Budget
1. Updated digital wetland inventory maps for 13 counties in east-central Minnesota	06/30/13	\$410,742

Result Completion Date: *June 30, 2013*

Result Status as of January 2011:

- A Request for Proposals was advertised and a vendor was selected to update the NWI maps for east-central Minnesota. Ducks Unlimited was selected through a competitive bid process. Their proposal was judged to provide the best overall value for the State. They have considerable experience with updating wetland

maps and provided the lowest overall cost. A contract between DNR and DU was developed and signed covering the update of the NWI maps for the 13 county east-central Minnesota project area (scope of work is attached). Some of the pre-processing of data and field-work will be conducted in-house, by the DNR Resource Assessment Office (RA).

- A kick-off meeting was held with Ducks Unlimited and members of the technical advisory panel for the NWI project to provide an overview of the scope of work, identify potential risks, and determine required actions to address these risks.
- DU and RA held a coordination meeting to plan for field logistics and technology transfer.
- DU and RA have begun gathering and formatting aerial imagery and other ancillary GIS data (soils, flood maps, local wetland inventories, DNR Public Water Inventory, and bathymetry). DU has acquired radar imagery from PALSAR.
- DU has begun testing of batch processing operations to move from pilot-scale operations to full production-scale.

Result Status as of July 2011:

- DU and RA have gathered all the data for the east-central project area including aerial photos, radar imagery, soils data, DEMs derived from available LiDAR, original NWI, and other wetland data. These data have been processed for 12 test areas (USGS quads). Some of the data has also been processed for areas outside the 12 test quads.
- Field training data was collected for 510 sites during May 2011 and processed for input into the classification model. Additional training data from other sources has also been processed for input in the classification model.
- Multiple iterations of the segmentation have been performed on the test quads with the aerial photos, digital elevation models, radar and soils data. Segmentation parameters have been optimized to generate segments for the photo-interpretation process.
- Initial runs of the potential wetlands classification have been tested and the process has been finalized. The photo interpreters have been through a two-day training session with the Senior GIS Analyst.
- Detailed methods and procedures have been developed, from data generation to data backup. The process steps from generating the segments and potential wetland classification to performing the photo interpretation and quality control have been determined and tested. Automation scripts have been written for processing the data, attribute editing, and quality control checks.
- The photo interpretation process to produce the draft classification is currently underway for the 12 test quads.

Result Status as of January 2012:

- DU provided a working draft of the technical procedures document. This document will be revised as adjustments are made to the method. A final version will be provided at the end of this project phase.
- DU conducted the tech transfer workshop with RA to harmonize the methods and approach used by the groups working on the NWI update.

- RA has completed all of the data preprocessing for 11 of the 13 counties in the east-central Minnesota project area. Processing continues for Rice and Goodhue counties, as this imagery was just recently delivered from the vendor.
- The data for the Koochiching NWI update area have been compiled, but are still being preprocessed.
- DU submitted draft data for nine quarter quads for DNR review. The production rate for draft NWI is beginning to accelerate.
- RA conducted follow-up field investigations to provide feedback to DU on difficult to interpret wetlands.
- DU has developed a website for the work they are doing on this project (<http://www.ducks.org/Conservation/glaro/glaro-gis-mn-nwi-update>) as well as a web-based status map that shows the status of data production. <http://gis.ducks.org/MNNWI/>
- DNR has reviewed and provided comments on the draft data submitted to date. Errors of omission and commission as well as classification errors have all been addressed by DU. Solutions to a few technical and cartographic issues with the data are still in the development stage, but issues will be resolved shortly.
- Results from the review of the initial draft data were presented to the Technical Advisory Committee to solicit additional feedback.
- DNR and USFWS held a meeting in October targeted at federal and state agencies to provide an overview of the NWI project, the current status, and to seek potential partnering opportunities with these agencies.
- DNR provided an overview and status report to GIS users at the Minnesota GIS/LIS conference in St. Cloud in October.

Result Status as of July 2012:

- DU has completed the initial image segmentation for all quarter quads in the ECMN project area except Rice & Goodhue (80% complete). Image segmentation on the remaining area is awaiting approval of an updated process for ensuring an accurate and efficient edge-matching procedure.
- DU has also completed initial photo-interpretation on 274 quarter quads (50% complete) and conducted an internal QA/QC review on 178 of these (33%). For this reporting period, DU has done the initial photo-interpretation for 244 quarter quads and internal QA/QC for 168 quarter quads.
- Quarter quads that have passed the internal review process are delivered to the DNR for its review.
- Processing has been finalized for 61 quarter quads, except for the addition of the requested enhanced wetland attributes.
- DU is currently developing the computer code to generate additional wetland attributes for the Eggers and Reed wetland plant community classification and the hydro-geomorphic classification.
- DU has also developed the model code for the wetland probability layer.
- Pre-processing of all data for the Koochiching project area was completed and the image segmentation has also been completed for all 50 quarter quads in this area.
- Initial image interpretation is nearly complete for all quarter quads in this area and an internal review has been completed for 20 quarter quads (40% complete).

- The production-level mapping methods developed for this project were presented at the ASPRS annual conference in Sacramento by Aaron Smith (DU/Equinox Analytics)
- The DNR has expanded the level of effort for QA/QC for the NWI data. This includes re-programming funds saved on data acquisition toward a service level agreement with the DNR Resource Assessment (RA) office to provide support for QA/QC review of the NWI data.
- Some of the funds saved from the data acquisition component of this project were re-programmed toward a service level agreement with the DNR Management Information Service (MIS) bureau to develop an online review tool. The purpose of this review tool is to provide a simple and efficient way to gather stakeholder review comments on the NWI from local wetland experts. This web-based mapping application displays the draft NWI data along with aerial imagery and LiDAR data. The application has tools that allow users to submit suggested revisions to the data. This tool has been developed, tested and deployed.
- The budget detail (Attachment A) has been modified to identify the service level agreements with other DNR units as separate line items.
- During this reporting period the DNR developed a Request for Proposal (RFP) for updating the NWI data for southern Minnesota. This RFP has been noticed in the State Register and we will be selecting a contractor to conduct the next phase of the NWI Update project.
- The DNR has completed its review of 78 quarter quads of draft data.

Results Status as of January 2013:

- Ducks Unlimited (DU) has completed the photo-interpretation process for 437 out of 541 quarter quads (81%) in the East Central project area (201 within the last 6 months).
- DU's internal QA/QC process has been completed on 433 out of 541 quarter quads (80%) in the east-central project area (265 within the last six months).
- The DNR has conducted a secondary review on 317 quarter quads (59%) for this project area. Two-hundred and twelve of these were reviewed by the DNR Resource Assessment Program (RAP) and 105 of these were reviewed by DNR Ecological and Water Resources (EWR). Comments from the DNR have been incorporated by DU.
- DU finished the scripting for the plant community and hydrogeomorphic classifications and tested the results and has developed the input layers needed for the HGM classification for the entire East Central project area.
- RAP has completed draft NWI data for all 50 quarter quads in the Koochiching project area.
- RAP completed an internal QA/QC review and EWR conducted a secondary review for the entire Koochiching project area. Final revisions are in process.

Results Status as of July 2013:

Ducks Unlimited completed all contracted tasks and provided a complete set of deliverables. The DNR reviewed these deliverables and conducted the final accuracy assessment on the data. The data have been publicly posted on the DNR Data Deli (<http://deli.dnr.state.mn.us/>) and a copy of the data has also been provided to the US

Fish and Wildlife Service for posting on their Wetland Mapper website (<http://www.fws.gov/wetlands/Data/Mapper.html>). Aaron Smith (Equinox Analytics – Ducks Unlimited) presented at the Society of Wetland Scientists annual meeting in Duluth Minnesota on June 6, 2013 based largely on work that was done for this project.

Results Status as of January 2014:

All of the original tasks under this activity are complete and the data have been posted publicly. The remaining \$10,175 in the budget was re-directed to accelerate the data processing task for the rest of the state. The DNR Resource Assessment Office has developed automated procedures for data processing. The initial results are currently under review. Once the results have been approved, the input GIS data for the rest of the state will be processed. This final additional task is expected to be complete by mid-April of 2014. The DNR and other project participants have also written a manuscript that will be submitted to the journal Wetlands for publication.

Final Report Summary:

Updated wetland maps were created for 13 counties in east-central Minnesota, encompassing the Twin Cities metropolitan area. Wetlands in Minnesota were originally mapped by the U.S. Fish and Wildlife Service in early 1980's as part of the National Wetlands Inventory (NWI). Although still widely used for land use planning, wetland permit screening and natural resource management, the original maps have grown increasingly out-of-date due to landscape alterations over the years. The data created for this project marks the first significant revision to the NWI in Minnesota.

RESULT/ACTIVITY 2: Methods Evaluation

Description: The goal of this component of the project is to complete the methods evaluation work that was started under phase one of the overall project. Specifically, this entails completing a pilot test of wetland mapping techniques for a site located in an agricultural setting in southern Minnesota. The methods evaluation aims to help develop high-accuracy, cost-effective procedures for updating NWI maps for the various major landscapes in Minnesota. The agricultural region pilot study is important because it addresses issues that are unique to wetland mapping in an area that is scheduled for update in the next phase of the overall project. This component of the project will be conducted by the University of Minnesota Remote Sensing and Geospatial Analysis Laboratory.

Summary Budget Information for Result/Activity 2: ENRTF Budget: \$126,040
Amount Spent: \$126,040
Balance: \$ 0

Deliverable/Outcome	Completion Date	Budget
1. Updated wetland maps for pilot area	6/30/12	\$32,000
2. Wetland mapping protocol for southern MN	12/31/12	\$87,040
3. Conduct workshop on wetland mapping	9/30/12	\$7,000

Result Completion Date: June 30, 2012

Result Status as of January 2011:

- A contract between the DNR and the UMN was developed and signed covering the completion of a third pilot test area in the agricultural region of Minnesota (scope of work is attached).
- Preliminary data gathering and evaluation for the third pilot area has begun.

Result Status as of July 2011:

- A pilot area was selected in the vicinity of Swan Lake near Mankato, MN.
- Imagery and other data are being acquired for the pilot areas. Currently, PALSAR and optical data are in hand. Radarsat-2 and lidar-derived DEM data have been ordered.

Result Status as of January 2012:

- UMN continued to develop and test mapping methods for the Swan Lake pilot area near Mankato, MN.
- An additional (informal) pilot area was selected in Big Stone County to look at the effects of a different landscape on wetland mapping methods.
- UMN presented wetland mapping methods to the US Fish and Wildlife Service Regional Office.
- UMN has authored two publications related to the methods development work for this project. One paper has been accepted for publication and the second paper is in review.
 - Corcoran, J., J.F. Knight, B. Brisco, K. Shannon, A. Cull, and K. Murnaghan. Integration of Optical, Topographic, and Radar Data for Wetland Mapping in Northern Minnesota. *Canadian Journal of Remote Sensing*. Accepted for publication.
 - Knight, J.F. and B.P. Tolcser. Remote classification of wetlands using decision trees. *Photogrammetric Engineering and Remote Sensing*. In review.

Result Status as of July 2012:

- RSGAL researchers conducted an extensive study of the effects of digital elevation model (DEM) preprocessing and mapping algorithm choice on the accuracy of wetlands maps in three physiographically different regions of Minnesota. The results of this study are being compiled and will be submitted for publication in a peer reviewed scientific journal and included in reports to LCCMR.
- RSGAL researchers continued studying optimal geospatial data types and mapping methods for wetlands in agriculture dominated portions of Minnesota. Current study sites are in the Swan Lake watershed near Mankato and the Big Stone watershed in western MN. Findings from this research will be incorporated into NWI mapping methods.
- Presentations:
 - Corcoran, J., Knight, J. Incorporating Data from Several Remotely Sensed Platforms to Map Current and Potentially Restorable Wetlands. International Association for Ecology (INTECOL), Orlando, FL, June 6, 2012.

- Corcoran, J., Knight, J. Incorporating Data from Several Remotely Sensed Platforms to Accurately Map Current and Potential Wetlands. American Society for Photogrammetry and Remote Sensing (ASPRS), Sacramento, CA, March 5, 2012.

Results Status as of January 2013:

- The University of Minnesota Remote Sensing and Geospatial Analysis Laboratory (RSGAL) hosted a technical workshop on wetland mapping for the southern Minnesota NWI update project. This full day event was held on the University of Minnesota Campus and was attended by three staff from St. Mary's University Geospatial Services (GSS). GSS was selected under the Phase 3 grant to update the wetland inventory maps for southern Minnesota.
- RSGAL researchers conducted an extensive study of the effects of digital elevation model (DEM) preprocessing and mapping algorithm choice on the accuracy of wetlands maps in three physiographically different regions of Minnesota. The results of this study have been compiled and are being prepared for publication in a peer reviewed scientific journal and included in reports to LCCMR.
- RSGAL researchers continued studying optimal geospatial data types and mapping methods for wetlands in agriculture dominated portions of Minnesota. Current study sites are in the Swan Lake watershed near Mankato and the Big Stone watershed in western MN. Findings from this research will be incorporated into NWI mapping methods.
- Publications
 - Knight, J.F., B. Tolcser, J. Corcoran, and L. Rampi. The effects of data selection and thematic detail on the accuracy of high spatial resolution wetland classifications. *Photogrammetric Engineering and Remote Sensing*. In press.
 - Jiang, Z., Shekhar, S., Mohan, P., Knight, J.F., Corcoran, J. Learning spatial decision tree for geographical classification: a summary of results. ACM SIGSPATIAL GIS 2012. In press.
 - Corcoran, J.M. and Knight, J.F. Influence of Multi-Platform, Multi-Frequency, and Multi-Temporal Remote Sensing Data on the Performance and Accuracy of Decision Tree Classification of Wetlands. *Remote Sensing*. In review.
 - Rampi, L. and Knight, J.F. Using lidar and high resolution imagery for wetland mapping in Minnesota. *Remote Sensing*. To be submitted in spring 2013.
 - Knight, J.F., Kloiber, S.M., Corcoran, J.M., Rampi, L.P. Effects of digital elevation model preprocessing and topographic derivations on wetland mapping accuracy. Journal TBD. To be submitted in summer 2014.
- Presentations
 - Corcoran, J.M.; Knight, J.F.; 2012. The influence of multi-platform, multi-frequency, and multi-temporal remote sensing and field reference data quality on the accuracy of decision tree classification of wetlands. Minnesota GIS/LIS Consortium, St. Cloud, MN, October 7, 2012.

Results Status as of July 2013:

The University of Minnesota delivered updated wetland map data for all the pilot areas along with a complete wetland mapping protocol document for southern Minnesota. Additionally, the University of Minnesota hosted a wetland mapping methods workshop with the mapping vendor for southern Minnesota and other project stakeholders. Lian Rampi (University of Minnesota) presented at the Society of Wetland Scientists annual meeting in Duluth Minnesota on June 6, 2013 based on the methods assessment work performed for this project.

- Publications
 - Knight, J.F., B. Tolcser, J. Corcoran, and L. Rampi. 2013. The effects of data selection and thematic detail on the accuracy of high spatial resolution wetland classifications. *Photogrammetric Engineering and Remote Sensing*, 79(7): 613-623.
 - Corcoran, J.M, J.F. Knight, A.L. Gallant. 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): 3212-3228.
- Presentations
 - Rampi, L.P., Knight J.F. Wetland mapping in Minnesota: an object based approach to integrate lidar and multispectral imagery. Society of Wetland Scientists, Duluth, MN, June 6, 2013.
 - Rampi, L.P., Knight J.F. Wetland mapping in Minnesota: an object based approach to integrate lidar and multispectral imagery. International Lidar Mapping Forum, Denver, CO, February 12, 2013.

Results Status as of January 2014:

All method assessment tasks are complete.

Final Report Summary:

Researchers at the University of Minnesota conducted an extensive study of the effects of digital elevation model (DEM) preprocessing and mapping methods on the accuracy of wetlands maps in three different physiographic regions of Minnesota. This research covered two study sites in agricultural areas including the Minnesota River Headwaters (Big Stone County) and Swan Lake (Nicollet County) as well as a comparison site from northern Minnesota (St. Louis and Carlton Counties). The results of this effort were compiled and submitted for publication in several peer-reviewed scientific journals along with results from the earlier phase of the NWI update project. Three hard copies and one electronic copy of these publications have been submitted with the final report to LCCMR. There have also been numerous presentations at professional conferences.

RESULT/ACTIVITY 3: Data Acquisition

Description: Creating a high quality update of the NWI requires having high quality data. This component will include acquisition of imagery along with field verification data for the next anticipated mapping phase in southern Minnesota. We will acquire high-resolution, spring leaf-off, multi-spectral aerial photography for 36 counties (although specifications could change based on recommendations from Result 2). The imagery

will be used as a base for updating the NWI maps for southern Minnesota. Data acquisition will also include a field-based assessment of wetland type for 400 to 500 sites chosen using a stratified random selection process. The field data will be used to assess the accuracy of the final wetland maps. To maintain the independence of the field data, the field data acquisition will be managed by University of Minnesota, Remote Sensing and Geospatial Analysis Laboratory and not shared with the mapping contractor.

Summary Budget Information for Result/Activity 3: ENRTF Budget: \$563,218
Amount Spent: \$563,218
Balance: \$ 0

Deliverable/Outcome	Completion Date	Budget
1. High-resolution, spring leaf-off, multi-spectral digital aerial imagery	6/30/14	\$475,218
2. Field data acquisition	06/30/13	\$88,000

Result Completion Date: 2014

Result Status as of January 2011:

- A Request for Proposals was developed and advertised. A vendor was selected to acquire high-resolution, leaf-off imagery for southern Minnesota this coming spring.
- An interagency agreement was developed between DNR and MnGeo to establish a partnership to better manage the spring aerial imagery acquisition project for southern MN.
- A contract between the DNR and the UMN was developed and signed covering the acquisition of field validation data for the southern agricultural region of Minnesota. Acquisition of field data will take place this summer.

Result Status as of July 2011:

- Acquisition of aerial photography was completed for 35 out of 36 counties. Imagery acquisition was conducted between April 12, 2011 and May 16, 2011. However, due to weather issues, the acquisition for Meeker County will be delayed until spring 2012.
- Preliminary image processing has been completed for all of the acquired images including aero-triangulation and seam line edits. Processing of the ortho-rectified imagery is ongoing. Delivery of both stereo and ortho-rectified imagery is on schedule for later this summer.
- The UMN has hired and trained field staff for collecting validation data.
- Acquisition of field data for southern MN was initiated in May 2011. However, these efforts were temporarily suspended during the State government shutdown from July 1 to July 21, 2011.

Result Status as of January 2012:

- Draft imagery for 35 out of 36 counties in southern Minnesota was provided by the aerial photography vendor (Surdex, Inc.). Imagery for Meeker County will be acquired in spring 2012.
- A detailed review of the draft imagery was conducted by DNR, MnGeo, and other project partners. All comments were addressed by the vendor (Surdex).
- Final imagery was delivered for 35 counties. The acceptance of the data is pending the results of the horizontal accuracy assessment, which is being conducted by MnDOT.
- UMN completed the field data acquisition and delivered the data to the DNR for 2703 sites (1722 upland and 981 wetland sites). DNR will be reviewing these data.

Results Status as of July 2012:

- Imagery data that was delivered by Surdex for 35 counties in southern MN has passed the quality control assessment. The data have been accepted and posted to the MnGeo web service for public access.
- Imagery acquisition for last remaining county in the acquisition area was delayed to spring 2013 due to weather issues this past spring.
- The DNR and MnGeo have been conducting outreach campaign to reach potential partners for next imagery acquisition phase. A series of informational meeting were held as a part of this effort in Fergus Falls, Bemidji, & Brainerd.

Result Status as of January 2013

- There were no additional actions related to result two (data acquisition) in this reporting period. All field data has already been acquired and the one remaining county of imagery data that needs to be acquired will be acquired this spring.

Results Status as of July 2013:

The remaining imagery for southern Minnesota was acquired and the processing is about 90% complete. Final processing and quality assessment will be performed in FY14. This is the only outstanding deliverable for this project. We anticipate that the project should be 100% complete by October of 2014.

Result Status as of January 2013

The remaining imagery for southern Minnesota was acquired and processing is complete. The final quality assessment found the data to meet the requirements. The data was delivered to the State and is posted on the MnGeo web service for public access 2014. All data acquisition tasks are complete.

Results Status as of July 2013:

All data acquisition tasks are complete.

Results Status as of January 2014:

All data acquisition tasks are complete.

Final Report Summary:

Accomplishments for this project include acquiring high-resolution, spring, leaf-off, digital aerial imagery for 23,900 square miles of southern Minnesota as well as acquiring field validation data for southern Minnesota.

V. TOTAL ENRTF PROJECT BUDGET:

Personnel: \$ 135,000 (DNR Project Manager – 0.65 FTE unclassified employee)

Contracts: \$ 955,000 (Details in Attachment A)

Equipment/Tools/Supplies: \$ 1,500 (batteries and accessories for GPS units, spray paint for accuracy assessment targets, etc.)

Acquisition (Fee Title or Permanent Easements): \$ NA

Travel: \$ 6,000 (\$4,000 for in-state travel for business meetings, field work, and training. \$2,000 for out-state travel for the DNR project manager to attend a professional symposium/workshop regarding current technology for mapping wetlands – American Society of Photogrammetry and Remote Sensing)

Additional Budget Items: \$ 500 (printing field manuals, procedures documents, reports, and maps)

TOTAL ENRTF PROJECT BUDGET: \$1,100,000

Explanation of Capital Expenditures Greater Than \$3,500: None

VI. PROJECT STRATEGY:

A. Project Partners: Joe Knight, Ph.D., of the University of Minnesota, Remote Sensing and Geospatial Analysis Laboratory will receive a total of \$180,000; \$100,000 for Result 2 (methods evaluation) and \$80,000 for Result 3 (field data acquisition).

Other partners providing in-kind services for this project include the Minnesota Pollution Control Agency, the Minnesota Board of Water and Soil Resources, the U.S. Fish and Wildlife Service, and the Minnesota Dept. of Administration's Geographic Information Office.

B. Project Impact and Long-term Strategy: This is the second phase of a multi-phase project to update the National Wetlands Inventory (NWI) for the entire state of Minnesota. The NWI provides critical baseline data that inform many wetland management actions and policies. We anticipate submitting proposals every other year for four additional phases (ie. 2012, 2014, 2016, and 2018). The estimated total budget for the project is \$7.5 million. With this project phase, we will have received \$1.65 million (about 22%) from ENRTF. Upon completion of this project phase, we will have completed 100% of the methods evaluation, 50% of imagery and field validation data acquisition for the state, and 10% of the updated wetland maps for the state.

C. Other Funds Proposed to be Spent During the Project Period: The DNR and its partners listed above will provide approximately \$20,000 of in-kind staff time in support of this project (but not tracked for reporting purposes). In addition, approximately \$154,000 in Department Operations and Division Support charges accruing to this project will be covered by Division general funds or other eligible Division funds.

Based on experience from phase one of this project, we also anticipate being able to find matching funds from local, state, and federal agencies for imagery acquisition. Any savings in the ENRTF budget that result from this will be redirected toward updating NWI maps for additional areas and/or acquisition of additional field data to validate the updated maps.

D. Spending History: The ENRTF provided \$550,000 for the first phase of this project. The first phase included: 1) developing mapping standards designed to ensure that the final product meets the needs of end users; 2) acquiring high-resolution, leaf-off, color infrared aerial imagery for northeastern and east-central Minnesota; and 3) evaluating imagery sources and mapping technologies to identify the most cost-effective, reliable inventory procedures for pilot study sites in northeastern and east-central Minnesota.

Matching funds for imagery acquisitions included: 1) National Geospatial Intelligence Agency (via U.S. Geological Survey) - \$25,000, 2) St. Louis County Planning Department - \$24,999, 3) Minnesota Pollution Control Agency - \$111,373, 4) National Oceanographic Atmospheric Administration (via DNR Coastal Zone Program) - \$24,227, and 5) DNR - \$181,065. We anticipate that the Metropolitan Council will contribute about half of the cost (about \$70,000) for acquiring imagery for the 13-county, east-central Minnesota project area (the Metropolitan Council will cover the costs for the seven-county region that corresponds to their statutory authority).

VII. DISSEMINATION:

Imagery acquired for this project is available to the public through the Minnesota Geospatial Information Office (http://www.mngeo.state.mn.us/chouse/wms/geo_image_server.html). The MnGeo imagery service receives about one million page requests per month for the southern Minnesota imagery. This is the first publicly available leaf-off imagery data for southern Minnesota since 1991.

The updated wetland map data are available through an interactive mapping application on the DNR's website at: <http://www.dnr.state.mn.us/eco/wetlands/map.html>. The data can also be downloaded, free of charge, for use in geographic information system applications through the DNR's data deli at: <http://deli.dnr.state.mn.us/>. The data will eventually be incorporated into the national "Wetland Mapper" application maintained by the U.S. Fish and Wildlife Service.

The wetland mapping procedures and accuracy results for the 13-county updated NWI data are presented and discussed in a manuscript that has been submitted to the journal *Wetlands*, a publication of the Society of Wetland Scientists (SWS). Information from this project was also presented at the SWS annual conference in Duluth, MN in 2013. In addition, a press release was distributed regarding the updated NWI data and the story was published on several online news websites.

Researchers at the University of Minnesota conducted an extensive study of the effects of digital elevation model (DEM) preprocessing and mapping methods on the accuracy of wetlands maps in three different physiographic regions of Minnesota. This research covered two study sites in agricultural areas including the Minnesota River Headwaters (Big Stone County) and Swan Lake (Nicollet County) as well as a comparison site from northern Minnesota (St. Louis and Carlton Counties). The results of this effort were

compiled and submitted for publication in in several peer-reviewed scientific journals along with results from the earlier phase of the NWI update project. Three hard copies and one electronic copy of these publications have been submitted with the final report to LCCMR. There have also been numerous presentations at professional conferences.

Publications

Corcoran, J.M, Knight, J.F., B. Brisco, S. Kaya, A. Cull, and Murhnaghan, K. (2011) The integration of optical, topographic, and radar data for wetland mapping in northern Minnesota. *Canadian Journal of Remote Sensing*, 27(5): 564-582.

Corcoran, J.M., Knight, J.F., and 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.

Knight, J.F., B. Tolcser, J. Corcoran, and Rampi, L. (2013) The effects of data selection and thematic detail on the accuracy of high spatial resolution wetland classifications. *Photogrammetric Engineering and Remote Sensing*, 79(7): 613-623.

Rampi, L.P., Knight, J.F., and Pelletier, K.C. (2014) Wetland mapping in the Upper Midwest United States: An object-based approach integrating lidar and imagery data. *Photogrammetric Engineering and Remote Sensing*. 80(5): 439-449.

Rampi, L.P., Knight, J.F., and Lenhart, C.F. (2014) Comparison of flow direction algorithms in the application of the CTI for mapping wetlands in Minnesota. *Wetlands*, 34(3): 515-525.

VIII. REPORTING REQUIREMENTS: Periodic work program progress reports will be submitted not later than January 2011, July 2011, January 2012, July 2012, January 2013, July 2013, and January 2014. A final work program report and associated products will be submitted between July 31, 2014 and August 31, 2014 as requested by the LCCMR.

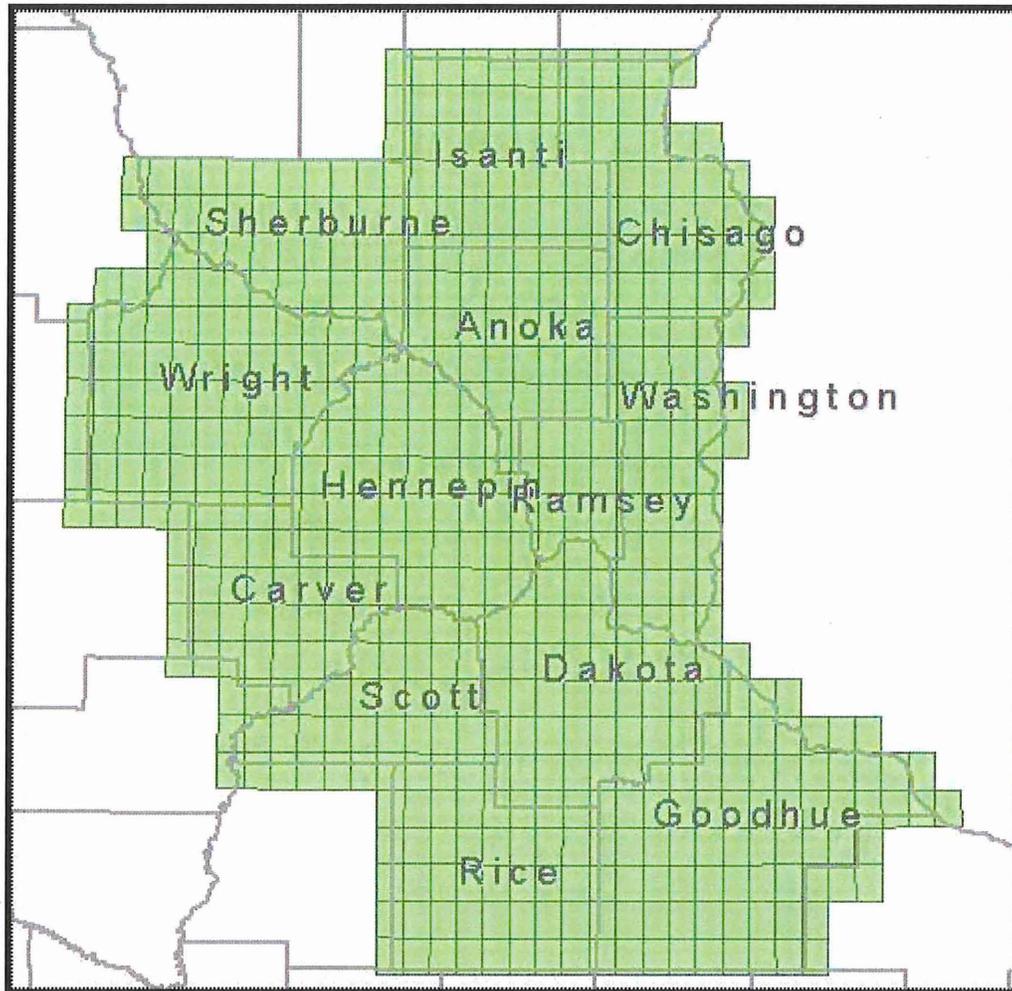


Figure 1: The focus of the first update of NWI will be the 13 counties Metropolitan Area.

Attachment A: Final Budget Detail for 2010 Projects											
Project Title: <i>Updating the National Wetland Inventory for Minnesota: Phase 2 (019-A3)</i>											
Project Manager Name: <i>Steve Kloiber</i>											
Trust Fund Appropriation: \$1,100,000											
1) See list of non-eligible expenses, do not include any of these items in your budget sheet											
2) Remove any budget item lines not applicable											
2010 Trust Fund Budget	Result 1 Budget:	Amount Spent	Balance (7/31/14)	Result 2 Budget:	Amount Spent	Balance (7/31/14)	Result 3 Budget:	Amount Spent	Balance (7/31/14)	TOTAL BUDGET	TOTAL BALANCE
BUDGET ITEM	<i>Updating Wetland Maps</i>			<i>Methods Evaluation</i>			<i>Data Acquisition</i>				
PERSONNEL: wages and benefits (Steve Kloiber 65%FTE - unclassified)	\$52,080	\$52,080	\$0	\$26,040	\$26,040	\$0	\$52,080	\$52,080	\$0	\$130,200	\$0
MIS Direct Support (Craig Perrault/Hal Watson - 87 hrs)	\$5,520	\$5,520	\$0							\$5,520	\$0
Contracts											
Professional/technical (Ducks Unlimited, selected by RFP, Wetland Mapping)	\$288,886	\$288,886	\$0							\$288,886	\$0
Professional/technical (Univ. of MN, Methods Evaluation)				\$100,000	\$100,000	\$0				\$100,000	\$0
Professional/technical (Univ. of MN, Field Data Acq.)							\$80,000	\$80,000	\$0	\$80,000	\$0
Professional/technical (Surdex, selected by RFP, Aerial Imagery)							\$428,873	\$428,873	\$0	\$428,873	\$0
Printing (procedures, reports, & maps)	\$0		\$0							\$0	\$0
Supplies (field supplies, batteries, GPS accessories, spray paint)	\$643	\$643	\$0				\$758	\$758	\$0	\$1,401	\$0
Travel expenses in Minnesota (mileage, per diem, lodging, etc.)	\$1,506	\$1,506	\$0				\$1,506	\$1,506	\$0	\$3,012	\$0
Travel outside Minnesota (conference/training, see note in workplan)	\$0		\$0				\$0		\$0	\$0	\$0
Other (Service Level Agreement with DNR Resource Assessment Office in Grand Rapids, MN for support on wetland mapping including data processing, field recon... and QA/QC)	\$62,107	\$62,107	\$0							\$62,107	\$0
COLUMN TOTAL	\$410,742.10	\$410,742	\$0.00	\$126,040	\$126,040	\$0	\$563,218	\$563,218	\$0	\$1,100,000	\$0

Selected news websites with stories about the updated NWI data

HometownSource.com

brought to you by ECM Publishers, Inc.



Home News Government Columns & Opinion Press Releases ECM Publishers Publications

New wetland map data available

By **Site Editor** on January 6, 2014 at 4:05 pm

St. Paul, MN – The Department of Natural Resources has released updated wetland map data for 13 counties in east-central Minnesota, encompassing the Twin Cities metropolitan area.

The wetlands were originally mapped by the U.S. Fish and Wildlife Service in the late 1970s and early 1980s as part of the National Wetlands Inventory (NWI).

Although still widely used for land use planning, wetland permit screening and natural resource management, the original maps have grown increasingly out-of-date due to landscape alterations over the years. The newly-released map data is the first time the NWI has been updated in Minnesota.

The data are available through an interactive mapping application on the DNR's website at: www.dnr.state.mn.us/eco/wetlands/map.html. The data can also be downloaded, free of charge, for use in geographic information system applications through the DNR's data deli at: <http://deli.dnr.state.mn.us/>.

The new maps reflect the latest technology in remote sensing and mapping including high-resolution aerial imagery and Light Detection and Ranging (LiDAR) data.

"The original NWI maps were quite good considering the imagery and mapping methods of the time, but the new maps are much more accurate, capture more detail and provide more information than the original maps," said Steve Kloiber, the DNR manager of the NWI update project.

Besides showing the location, size and type of each wetland, the updated map data include information on the wetland's landscape position and hydrologic characteristics, which can be useful in assessing the benefits provided, such as water quality improvement, flood storage, and fish and wildlife habitat.

Search...



Shop at Gander Mountain

gandermountain.com

Hunting, Fishing and Outdoor Gear. Free Shipping on \$50 or more!

GIS Solutions

Free Netflow Analyzer

Search for Obituaries

myheritage.com/Obituaries

The top online search engine for obituaries and death records.

Free Obituary Search

Switch to Google Apps

HOME

MNSJ SPORTSMAN'S CLUB

MNSJ MAGAZINE

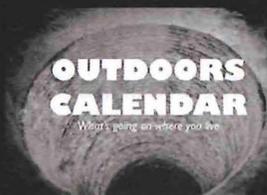
ADVERTISE

MNSJ RADIO

CONTACT US

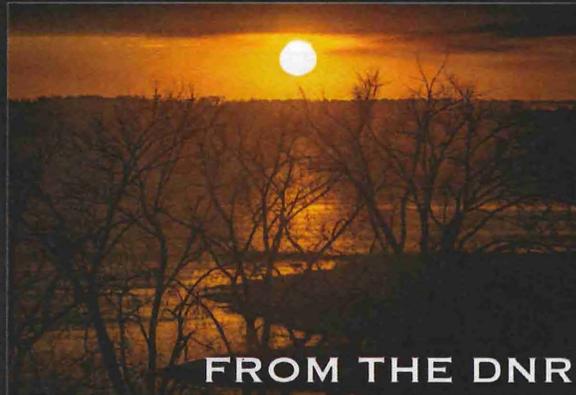
RADIO AFFILIATES LOGIN

SUBSCRIBE



FROM THE DNR: 900 acres of public land in Rice County, National Winter Trails Day, more...

— JANUARY 6, 2014 BY MNSJ STAFF



FROM THE DNR

MINNESOTA DNR NEWS #1

Jan. 6, 2014

All news releases are available in the DNR's website newsroom at www.mndnr.gov/news.
Follow the DNR on Twitter @mndnr.

IN THIS ISSUE

Nearly 900 acres of public land added in Rice County south of Twin Cities

DNR Parks and Trails Division solicits park and trail grant applications for 2014

[New wetland map data available](#)

DNR and REI invite families to discover the fun of snow sports on

National Winter Trails Day

Grassroots allies work together to keep grass on the land

DNR NEWS – FOR IMMEDIATE RELEASE

Nearly 900 acres of public land added in Rice County south of Twin Cities



RECENT POSTS

MINNESOTA AUGUST GOOSE SEASON UPDATE

THE KICKING BEAR CAMP STORY: MNSJ RADIO PODCAST

Fort Snelling State Park: To Reopen

DAKOTA REPORT: North Dakota

Pronghorn Season To Open

WHEN YOUR TRUCK LOOKS GOOD, YOU LOOK GOOD.



Link to the story is highlighted. The full story from the website is shown below.

New wetland map data available

The Department of Natural Resources has released updated wetland map data for 13 counties in east-central Minnesota, encompassing the Twin Cities metropolitan area. The wetlands were originally mapped by the U.S. Fish and Wildlife Service in the late 1970s and early 1980s as part of the National Wetlands Inventory (NWI).

Although still widely used for land use planning, wetland permit screening and natural resource management, the original maps have grown increasingly out-of-date due to landscape alterations over the years. The newly-released map data is the first time the NWI has been updated in Minnesota.

The data are available through an interactive mapping application on the DNR's website at: www.dnr.state.mn.us/eco/wetlands/map.html. The data can also be downloaded, free of charge, for use in geographic information system applications through the DNR's data deli at: <http://deli.dnr.state.mn.us/>.

The new maps reflect the latest technology in remote sensing and mapping including high-resolution aerial imagery and Light Detection and Ranging (LiDAR) data.

"The original NWI maps were quite good considering the imagery and mapping methods of the time, but the new maps are much more accurate, capture more detail and provide more information than the original maps," said Steve Kloiber, the DNR manager of the NWI update project.

Besides showing the location, size and type of each wetland, the updated map data include information on the wetland's landscape position and hydrologic characteristics, which can be useful in assessing the benefits provided, such as water quality improvement, flood storage, and fish and wildlife habitat.

The release of the wetland map data for east-central Minnesota marks completion of the first phase of a statewide update of the NWI. New, high resolution aerial imagery has been acquired for the entire state and wetland mapping is currently underway for the southern third of the state and a portion of northeast Minnesota.

The DNR is planning to complete the entire state by 2020. The NWI update project is being funded by the Minnesota Environment and Natural Resources Trust Fund as recommended by the Legislative-Citizen Commission on Minnesota Resources.

The trust fund is a permanent fund constitutionally established by Minnesotans to assist in the protection, conservation, preservation, and enhancement of the state's air, water, land, fish, wildlife, and other natural resources.

Fishshootouts

[Blog](#)[Beachbooker](#)[Boats](#)[Cabinweb](#)[Tackle](#)[Tournaments](#)

Minnesota: New wetland map data available

January 12, 2014 on 7:32 pm | In Midwest, Minnesota, Mobile APPS, News, USA | No Comments

 Like Sign Up to see what your friends like.

Enterprise File Sharing

 egnyte.com/Enterprise-File-Sharing

Securely Access Files & Collaborate Anywhere, Any Device. Free Trial!



The Department of Natural Resources has released updated wetland map data for 13 counties in east-central Minnesota, encompassing the Twin Cities metropolitan area. The wetlands were originally mapped by the U.S. Fish and Wildlife Service in the late 1970s and early 1980s as part of the National Wetlands Inventory (NWI).

Although still widely used for land use planning, wetland permit screening and natural resource management, the original maps have grown increasingly out-of-date due to landscape alterations over the years. The newly-released map data is the first time the NWI has been updated in Minnesota.

The data are available through an interactive mapping application on the DNR's website at: www.dnr.state.mn.us/eco/wetlands/map.html. The data can also be downloaded, free of charge,

[Outdoorplay.com](#) -
Free Kayak Shipping
on order

Select Language ▾

Powered by  Google Translate



CATEGORIES:

Categories

- [Abalone \(rss\)](#)
- [Aguariums \(rss\)](#)
- [Alligator Gar \(rss\)](#)
- [Apparet \(rss\)](#)
- [Bail \(rss\)](#)
- [Barracuda \(rss\)](#)
- [Bass \(rss\)](#)
- [Billfish \(rss\)](#)
- [Bluegill \(rss\)](#)
- [Boats \(rss\)](#)
- [Charters \(rss\)](#)
- [Kayaking Canoeing \(rss\)](#)
- [Rafting \(rss\)](#)
- [Bonefish \(rss\)](#)
- [Bonito \(rss\)](#)
- [Books Magazines \(rss\)](#)
- [Bowfishing \(rss\)](#)
- [Bream \(rss\)](#)
- [Bullhead \(rss\)](#)
- [Burbot \(rss\)](#)
- [Burrimundi \(rss\)](#)
- [Camping \(rss\)](#)
- [Carp \(rss\)](#)
- [Catfish \(rss\)](#)



Minnesota Environment and Natural Resources Trust Fund shared a link.

January 7

ENRTF project at DNR updating the National Wetland Inventory for Minnesota has released new wetland map data.



New wetland map data available

hometownsource.com

St. Paul, MN – The Department of Natural Resources has released updated wetland map data for 13 counties in east-central Minnesota, encompassing the Twin Cities metropolitan area.

Like · Comment · Share

1 Share

From LCCMR's Facebook page

A semi-automated, multi-source data fusion update of a wetland inventory for east-central Minnesota, USA

Steven M. Kloiber¹, Robb D. Macleod², Aaron J. Smith³, Joseph F. Knight⁴, and Brian J. Huberty⁵

- 1) Minnesota Department of Natural Resources (corresponding author)
500 Lafayette Road North
St. Paul, MN 55155
Phone: 651-259-5164 / FAX: 651-296-1811
E-mail: steve.kloiber@state.mn.us
- 2) Ducks Unlimited Inc.
1220 Eisenhower Place
Ann Arbor, MI 48108
- 3) Equinox Analytics Inc.
PO Box 6941
Columbia, SC 29204
- 4) Department of Forest Resources
University of Minnesota
1530 Cleveland Ave N
Saint Paul, MN 55108
- 5) U.S. Fish & Wildlife Service
5600 American Blvd West; Suite 990
Bloomington, MN 55437

DRAFT

1 **ABSTRACT**

2 Comprehensive wetland inventories are an essential tool for wetland management, but developing and
3 maintaining an inventory is expensive and technically challenging. Funding for these efforts has also
4 been problematic. Here we describe a large-area application of a semi-automated process used to
5 update a wetland inventory for east-central Minnesota. The original inventory for this area was the
6 product of a labor-intensive, manual photo-interpretation process. The present application incorporated
7 high resolution, multi-spectral imagery from multiple seasons; high resolution elevation data derived
8 from lidar; satellite radar imagery; and other GIS data. Map production combined image segmentation
9 and random forest classification along with aerial photo-interpretation. More than 1000 validation data
10 points were acquired using both independent photo-interpretation as well as field reconnaissance.
11 Overall accuracy for wetland identification was 90% compared to field data and 93% compared to
12 photo-interpretation data. Overall accuracy for wetland type was 72% and 78% compared to field and
13 photo-interpretation data, respectively. By automating the most time consuming part of the image
14 interpretations, initial delineation of boundaries and identification of broad wetland classes, we were
15 able to allow the image interpreters to focus their efforts on the more difficult components, such as the
16 assignment of detailed wetland classes and modifiers.

17 **Keywords:** wetlands inventory, wetland mapping, accuracy assessment, remote sensing

18 INTRODUCTION

19 Wetland inventory maps are essential tools for wetland management, protection, and restoration
20 planning. They provide information for assessing the effectiveness of wetland policies and management
21 actions. These maps are used at all levels of government, as well as by private industry and non-profit
22 organizations for wetland regulation and management, land use and conservation planning,
23 environmental impact assessment, and natural resource inventories. Wetland inventories are used to
24 assess impacts of regulatory policy (Gwin et al. 1999), assess habitat distribution and quality (Austin et
25 al. 2000; Hepinstall et al. 1996; Marchand and Litvaitis 2004; Knutson et al. 1999), evaluate carbon
26 storage potential and climate change impacts (Euliss et al. 1999; Burkett and Kusler 2000), and measure
27 and predict waterfowl and amphibian population distribution (Yerkes, et al. 2007; Munger et al. 1998;
28 Knutson et al. 1999).

29 There are several notable efforts across the globe to conduct national and regional comprehensive
30 wetland inventories. The Canadian Wetland Inventory (CWI) is developing a comprehensive wetland
31 inventory based on remote sensing data from Landsat and Radarsat platforms (Li and Chen 2005;
32 Fournier et al 2007). The CWI maps wetlands down to a minimum mapping unit of 1 ha using a five class
33 system. In 1974, the U.S. Fish and Wildlife Service began an effort to implement the National Wetlands
34 Inventory (NWI) for the United States (Cowardin et al. 1979). The NWI is based on manual aerial photo-
35 interpretation with a target map unit of 0.2 ha and a detailed hierarchical classification scheme involving
36 wetland systems, classes, subclasses, water regimes, and special modifiers (Dahl 2009). The
37 Mediterranean wetland initiative promotes standardized methods for wetland inventory and monitoring
38 across multiple countries in the Mediterranean region (Costa et al 2001). Wetland classification and
39 mapping recommendations for this initiative closely follow the NWI. More recently, wetlands across
40 China have been mapped using Landsat data into three broad classes with 15 subtypes generally based
41 on landscape and landform characteristics (Gong et al. 2010). Despite these efforts, a review of the

42 status of wetland inventories concluded that there still are significant gaps in our knowledge about the
43 extent and condition of global wetland resources. Finlayson and Spiers (1999) found that outside of a
44 few of the more developed countries and regions, wetland inventories were generally incomplete or
45 non-existent.

46 Even regions with comprehensive wetland inventories require periodic updates. For example, in
47 Minnesota, most of the NWI is 25 to 30 years old. Many changes in wetland extent and type have
48 occurred since the original inventory was completed. Agricultural expansion and urban development
49 have contributed to wetland loss. Conversely, various wetland conservation policies and programs have
50 resulted in the restoration of some previously drained wetlands and the creation of new wetlands.
51 Furthermore, limitations in the technology, methodology and source data for the original NWI resulted
52 in an under representation of certain types of wetlands. In northeastern Minnesota, wetlands were
53 originally mapped using 1:80,000 scale panchromatic imagery. The resulting wetland maps in this area
54 tend to be very conservative, missing many forested and drier emergent wetlands (LMIC 2007).

55 Updating the wetland inventory for such areas enhances the ability of conservation organizations to
56 make better management decisions. There is a significant ongoing need to develop and update wetland
57 inventories.

58 Maintaining wetland inventories can be expensive and technically challenging given the complexity of
59 wetland features and user expectations for a high degree of accuracy. Federally funded updates to the
60 NWI are required to conform to the federal wetland mapping standard (FGDC 2009). This standard calls
61 for $\geq 98\%$ producer's accuracy for all wetland features larger than 0.2 ha and a wetland class-level
62 accuracy of $\geq 85\%$. Unfortunately, funding for mapping in the NWI program has declined over the past 20
63 years (Tiner 2009) and has been almost entirely eliminated as of 2014 (NSGIC 2014).

54 Historically, the NWI has been primarily the product of manual aerial photo-interpretation (Tiner 1990).
65 Much of the original delineation and classification was done using hardcopy stereo imagery with mylar
66 overlays. In the last decade, NWI mapping efforts have largely transitioned to heads-up, on-screen
67 digitizing and classification from digital orthorectified imagery (Drazkowski et al 2004; Dahl et al. 2009).
68 Despite the efficiency gains achieved by migrating to an on-screen digitizing process, the process is still
69 labor-intensive.

70 Automated classification of wetlands from remote sensing data has had varied results. Ozesmi and
71 Bauer (2002) compare the results of automated wetland classification using satellite imagery to wetland
72 mapping from manual photo-interpretation. In their review, they note that the limitations of satellite
73 imagery, specifically resolution limitations when compared to aerial photography as well as limitations
74 related to spectral confusion between classes, led the NWI program to choose a method based on
75 photo-interpretation. However, given the advancements in the fields of remote sensing and image
76 analysis since the NWI was originally designed, the use of automated mapping and classification
77 techniques warrants reconsideration.

78 Collecting, managing, and analyzing large quantities of high spatial resolution digital imagery has
79 improved significantly over the past two or three decades. Airborne imagery acquisition systems like the
80 Zeiss/Intergraph Digital Mapping Camera (Z/I DMC) and the Vexcel Ultracam are commonly used to
81 acquire four-band multispectral imagery at less than 1-meter resolution. In addition, high-resolution,
82 multispectral imagery is also available through various satellite systems such as Worldview-2, Quickbird
83 and IKONOS. The costs for data storage required for the large quantities of high-resolution imagery data
84 have dropped significantly and advances in automated image analysis techniques have improved the
85 efficiency with which these data can be processed.

86 Radar imagery shows potential to provide new information such as water level changes in wetlands, soil
87 saturation and vegetation structure (Corcoran et al. 2011; Bourgeau-Chavez et al. 2013). In the near
88 term, the sources of satellite radar imagery are somewhat limited. Yet, Radarsat imagery is being used
89 operationally as part of the Canadian Wetland Inventory (Brisco et al. 2008).

90 Recent widespread adoption of scanning topographic lidar systems also provides a new source of highly
91 relevant digital information for wetland mapping. The distribution and occurrence of wetlands is heavily
92 influenced by topography. For example, Beven and Kirkby (1979) described a topographic index to
93 predict spatial patterns of soil saturation based on the ratio of the upslope catchment area to the
94 tangent of the local slope. Numerous researchers have used this topographic index, alternately known
95 as the compound topographic index (CTI) or the wetness index, to predict the occurrence of wetlands
96 (Hogg and Todd 2007; Murphy et al. 2007; Rampi et al. 2014b). As such, topographic analysis of lidar
97 data is an important emerging technology for wetland mapping.

98 Image segmentation is a process that groups adjacent image pixels into larger image objects based on
99 criteria specified by the image analyst. The goal of segmentation is to simplify the image into a smaller
100 number of potentially meaningful objects which can then be classified using various attributes
101 describing these objects (i.e. brightness, texture, size, and shape). This technique simultaneously
102 reduces data volume while incorporating spatial contextual information in the classification process.

103 Image segmentation has been shown to be a potentially valuable technique for improving image
104 classification accuracy for mapping land cover (Myint et al. 2011) and wetlands (Frohn et al. 2009).

105 Classification algorithms like random forest (Breiman 2001) have greatly improved our ability to
106 effectively integrate data from multiple sources into an automated classification procedure.

107 Incorporating data from multiple sensor systems as well as ancillary GIS data can potentially improve

108 wetland classification accuracy (Corcoran et al. 2011, Knight et al. 2013, Corcoran et al. 2013, Rampi et
109 al. 2014a).

110 Here we describe a large area application of a semi-automated classification process used to update the
111 NWI. The objective of this effort was to determine whether automated techniques such as image
112 segmentation, digital terrain analysis, and random forest classification could be combined with multiple
113 high-resolution remote sensing and GIS data sets and traditional photo-interpretation to efficiently
114 produce an accurate and spatially detailed wetland inventory map.

115 **METHODS**

116 **Study Area**

117 The study area is 18,520 square kilometers, centered on the 13-county metropolitan area of
118 Minneapolis and Saint Paul, Minnesota (Figure 1). The study area is situated primarily in the Eastern
119 Broadleaf Forest Ecological Province (DNR 2013) and the climate is typical of its continental position
120 with hot summers and cold winters. Typical annual precipitation ranges from about 76 to 81 centimeters
121 (Minnesota Climatology Working Group 2012). Land use in the study area varies from a dense urban
122 core with a mix of commercial and high density residential area, to lower density suburban and exurban
123 communities, and rural agricultural and forests.

124 **Input Data**

125 The primary imagery used for the NWI update was spring, leaf-off, digital aerial imagery with four
126 spectral bands (red, green, blue, and near infrared) in 541 orthorectified USGS quarter quadrangle tiles.
127 The imagery was acquired using a Z/I DMC camera in early April of 2010 and late April to early May of
128 2011. Imagery for 60% of the project area was acquired at a spatial resolution of 30 cm, while imagery
129 for the other 40% was acquired at 50 cm resolution. The imagery has a horizontal root mean square

130 error (radial) of 78 cm (MnGeo 2010). For the image segmentation process, the 30cm images were
131 resampled to 50cm resolution using a bilinear interpolation algorithm.

132 Thirteen single-date scenes of PALSAR L-band radar were acquired to cover the project area to aid in the
133 identification of forested wetlands. The scenes available were a combination of single and dual
134 polarization during a leaf-off seasonal window. The Alaska Satellite Facility MapReady Remote Sensing
135 Tool Kit (ASF 2011) was used for terrain correction and geo-referencing. Additional geo-referencing was
136 performed in ArcGIS using control points selected from the aerial imagery. A radar processing extension
137 in Opticks was used to reduce speckle in the data (Opticks 2011). Radar imagery was classified using a
138 10-class maximum-likelihood ISODATA clustering routine implemented in ERDAS Imagine software
139 (ERDAS 2008). The classes associated with "wet forest" training sites were identified and the
140 classification was applied to all clusters within the radar image.

141 Digital elevation models (DEMs) were derived from lidar data for approximately 60% of project area,
142 while DEMs for the remainder were 10-meter resolution DEMs obtained from the National Elevation
143 Dataset. The typical lidar point spacing was about 1 point per square meter. The Minnesota DNR
144 processed the bare earth points into a digital elevation model using 3D Analyst for ArcGIS by importing
145 the points into a terrain data set and then interpolating a 1-meter DEM that was subsequently
146 resampled to a 3-meter DEM. This lidar DEM has a vertical root mean square of 18 cm.

147 ArcGIS Spatial Analyst (ESRI 2011) was used to calculate slope, curvature, plan curvature, profile
148 curvature, topographic position index (TPI) and compound topographic index (CTI). TPI was calculated by
149 subtracting the mean elevation for a given pixel from the mean elevation of its neighborhood (Guisan et
150 al. 1999). We used an annulus neighborhood with radii of 15 and 20 meters. The CTI (Moore 1991) was
151 calculated using a sinkless version of the DEM. A slope grid and upstream catchment area grid were

52 calculated using the D-Infinity flow directions tool from TauDEM (Tarboton 2003). CTI was then
153 computed from slope and contributing drainage area using a custom python script.

154 The Natural Resources Conservation Service (NRCS) digital Soil Survey Geographic (SSURGO) layers were
155 compiled for the project area (NRCS 2010). Two derived raster products were produced from SSURGO
156 data; (1) the soil water regime class, and (2) the percentage of hydric soil. The variables used to derive
157 these products included drainage class, flood frequency for April, pond frequency for April, and pond
158 frequency for August.

159 The layers described above were formatted for input to an Object Based Image Analysis (OBIA) process
160 using the Cognition Network Language (CNL) implemented within eCognition software (Trimble 2010).
161 Images were clipped to the boundary of the relevant quarter quad tile and stacked with ERDAS Imagine
162 software (ERDAS 2008) into a single multi-layer file subsequently referred to here as the layer-stack.

163 **Training Data**

164 Reference field data were collected to serve as training data for the random forest classification and to
165 guide the interpreters during the image interpretation process. A set of 12 representative sub-areas
166 were selected for field visits to provide representative training data for the wetland types found
167 throughout the project area. The sub-areas were selected to be spatially distributed and to represent
168 the range of landscape types in the project area. Within these sub-areas, individual wetland sites were
169 selected for field visits using a stratified-random sampling approach with strata proportioned according
170 to the frequency of wetland classes. Rarely occurring wetland types were always flagged for field visits.

171 A total of 510 field sites were visited. The training data were augmented by including 1967 sites selected
172 from field data provided by field biologists at the Metropolitan Mosquito Control District as well as 873
173 sites image-interpreted by Ducks Unlimited.

174 All training data were classified according to the Cowardin classification system (Cowardin et al. 1979),
175 which is a hierarchical system developed to standardize the classification of wetlands and deepwater
176 habitats of the United States. Additional details of the classification system including the definition of
177 each system, subsystem, class, and subclass can be found in Cowardin et al. (1979) and Dahl et al.
178 (2009).

179 **Automated Components**

180 The object-based image analysis (OBIA) rule set consisted of several steps to separate wetlands from
181 other land cover types. The process began with a multi-resolution segmentation algorithm (Baatz and
182 Schape2000) that created image objects (groups of spectrally similar pixels). Parameters for the initial
183 segmentation were; scale factor = 6, shape = 0.5, compactness = 0.9, RGB weight = 1, and near infrared
184 weight = 2. A relatively small scale parameter was chosen to ensure that small wetlands would be
185 represented in the lowest level of the image object hierarchy. A three-tier hierarchy consisting of
186 spatially nested sub-objects, mid-level objects, and super objects provided a flexible framework for
187 iteratively integrating information from different image and topographic data sources. The rule set was
188 designed to draw boundaries for real world features of interest (e.g., stream beds) by iteratively
189 aggregating sub-objects at a temporary mid-level according to rules defining specific features of interest
190 for each major sequence of the larger rule set. Once useful boundaries for a particular sequence were
191 identified (using temporary classification thresholds and modification of the object boundaries at the
192 mid-level), the feature boundary information was conveyed to the super-level for inclusion in the final
193 output. Each modified mid-level was then destroyed and the unmodified sub-objects were re-used to
194 initialize a new version of the mid-level to repeat the process of selective aggregation and classification
195 for the next feature of interest.

196 The first major process sequence was designed to identify wooded-wetlands using the PALSAR radar
197 data. Sub-objects were aggregated at a temporary mid-level according to boundaries created from the

198 previously classified PALSAR data. A mask layer with the boundaries of the PALSAR wetland clusters was
199 incorporated into the layer stack data. The boundaries created by the 20m resolution PALSAR-derived
200 wooded wetland mask were not cartographically compatible with boundaries for other features derived
201 from the 0.5m resolution image data. This difference was reconciled in the eCognition rule set via a
202 custom-built iterative pixel-based object merging and reshaping algorithm applied to the mid-level in
203 the object hierarchy.

204 The second major process sequence in the rule set was designed to isolate open water stream features
205 and stream-bed topographic features. A preliminary linear stream vector layer was generated using Arc
206 Hydro terrain modeling software (Maidment 2002) to identify likely flow pathways using the lidar
207 derived DEM data. This linear flow path layer was used to seed a region growing sequence that
208 identified spectrally dark sub-objects contiguous to the modeled stream lines. These objects were
209 merged at the mid-level and the boundaries were smoothed to form the stream polygons, which were
210 then stored at the super-object level. A spectral difference segmentation algorithm (Definiens Imaging,
211 2009) was then used on the DEM (threshold value of 0.05m) to generate temporary elevation contours.
212 The contour objects containing nested stream-sub-objects were then identified and classified as
213 potential riparian areas, which were more likely to contain wetlands.

214 The third major process sequence in the rule set separated forested areas from non-forested areas and
215 selectively generated contour lines in forest polygons. Forested areas were identified by aggregating
216 sub-objects at a temporary mid-level according to image spectral characteristics ($0.017 < NDVI < 0.28$
217 and $RGB \text{ brightness} < 150$) and textural characteristics (average mean difference to neighbors of sub-
218 objects > 0.95 in the NIR band). Small candidate forest objects were then merged into stand sized
219 forested objects. Based on prior experience, the photo interpretation team requested that elevation
220 data be added to forested areas. A spectral difference algorithm which merged together objects with

221 similar elevation values was applied to the sub-objects of the forest stand objects. An elevation
222 threshold value of 0.33m was used to create objects that approximate 0.33m contour intervals.

223 The final major process sequence in the rule set was designed to create a background layer of general-
224 purpose image objects, which are delivered to the photo-interpretation team for editing in order to
225 create the final wetland map. A multi-resolution segmentation algorithm (parameters: scale factor =
226 400, shape = 0.1, compactness = 0.9, RGB weight = 1 and NIR weight = 2) was used in all areas not
227 classified in the previous sequences to delineate strongly visible boundaries in the spring leaf-off
228 imagery. This finalized set of image objects was then smoothed and exported in a vector shape-file
229 format for transfer to the photo-interpretation team.

230 Each image object has numerous associated attributes derived from the imagery, DEM, and other
231 ancillary data sets. These attributes, along with the training data, were used to create a classification
232 model using the randomForest package in R (R Development Core Team 2011; Breiman 2001). All image
233 objects were also assigned a unique identification number so that the classification model results could
234 be linked back to the image objects.

235 **Manual Components**

236 A 750-meter square grid system (enabling the interpreter to completely view an image section on a
237 monitor at 1:3,000) overlaid on each image was used to systematically guide image-interpretation
238 efforts and ensure complete interpretation of each image. Interpreters viewed the classified image
239 segmentation data superimposed over the spring imagery to identify and categorize wetlands.

240 Additional ancillary data were used during the interpretation process when needed, including; summer
241 imagery from 2008-2010, SSURGO soils derived products, the DEM, and DEM derived products. The
242 interpreters could use the segmentation derived boundary without modification, manually edit the
243 polygon boundary, or discard the segmentation based boundary to manually digitize a new boundary.

44 Adjacent wetland polygons of the same class were merged. All automated wetland classification values
245 were either confirmed or manually reclassified by a human interpreter. As with the field data, all
246 mapped wetland polygons were classified according to the Cowardin classification system (Table 1).

247 **Validation Data**

248 Two sets of independent validation data were created using field checks and independent image-
249 interpretation, respectively. The validation data were not made available to the image analysts. These
250 data were reserved to make a post-processing accuracy assessment of the updated wetland inventory
251 maps.

252 We created a set of 951 validation points through field checks and another set of 901 validation points
253 through independent image-interpretation. All points were initially selected using a stratified-random
254 sampling process with the strata defined by a recently developed land cover dataset from the
255 Minnesota wetland status and trends monitoring program (Kloiber et al. 2012). The stratification was
256 designed to place 75% of the selected points in wetlands and 25% in uplands. We used this sampling
257 scheme in an attempt to ensure that all wetland classes were well represented in the validation data.

258 Field validation points were evaluated by crews making ground-level assessments of wetland class
259 between May and September of 2010. Geographic coordinates were acquired at each observation site
260 using a Trimble Juno GPS data logger and the data were differentially corrected to improve positional
261 accuracy. Image-interpretation validation points were classified using image-interpretation of high-
262 resolution, digital stereo imagery, lidar-derived digital elevation models, and other ancillary data. Digital
263 stereo imagery was viewed using a stereo-photogrammetry workstation equipped with StereoAnalyst
264 software for ArcGIS (ERDAS 2010) and a Planar SD1710 stereo-mirror monitor.

265 The mapped wetland class was associated with the validation reference class using a spatial join process
266 in ArcGIS. Distances to the wetland feature and class boundaries were computed. To address potential

267 confusion between classification accuracy and positional accuracy, image-interpreted points that fell
268 within the 95% confidence interval for the positional accuracy of the imagery (1.53 meters) of a wetland
269 feature or class boundary were excluded from analysis. Field points that fell within the combined 95%
270 confidence interval for the positional accuracy of the imagery and the GPS (5.64 meters) of a wetland
271 feature or class boundary were also excluded.

272 The data were compared at two levels: agreement for a simple two-category system of wetland-upland
273 features, and agreement for the wetland class-level. The producer's accuracy, the user's accuracy, and
274 the overall accuracy were calculated (Congalton and Green 2008). The producer's accuracy is equal to
275 the complement to the omission error rate for the map, whereas the user's accuracy is equal to the
276 complement to the commission error rate. Mixed classes occur occasionally in the mapped data due to
277 spatial scale limitations. Wetland features that consist of highly interspersed classes are impractical to
278 separate and classify at the map scale. However, mixed classes did not occur in the validation data. For
279 the purposes of the accuracy assessment, if the field class matched either of the classes in a mixed class
280 map unit, it was counted as a match.

281 **RESULTS**

282 **Intermediate Automated Classification Results**

283 Initial image segmentation efforts resulted in many small image objects, requiring significant time spent
284 merging, classifying, and editing features (Figure 2). However, feedback from the photo-interpreters was
285 incorporated into a refined image segmentation rule set to provide image objects which more closely
286 represented the wetland features of interest. Initially, the typical number of image objects per quarter
287 quad tile was about 96,000; after refining the segmentation rules the per-tile average object count was
288 about 4,300. The refined segmentation rules aggregated image objects resulting in an increase in the

289 mean object size of 430 m² to 1,600 m². The minimum object area stayed roughly the same, while the
290 maximum object area went from 8,900 m² to 57,000 m².

291 The subsequent random forest classification had an overall bootstrapped accuracy of 92% for separating
292 wetlands from uplands and an overall bootstrapped accuracy of 69% for wetland class assignment.

293 These values should be treated with some degree of caution, as the bootstrapped accuracy results are
294 not directly comparable to the final accuracy assessment using the independent validation data.

295 Nonetheless, these results do support the notion that the automated classification component
296 significantly reduces the work load of the manual photo-interpreter by providing a reasonably accurate
297 intermediate product.

298 **Final Product Accuracy Assessment**

299 There were 743 field validation data points after excluding points within the positional uncertainty of a
300 mapped wetland boundary. The overall field accuracy for discriminating between wetland and upland
301 was 90%. The wetland producer's accuracy was 90% and the user's accuracy was 96% (Table 2).

302 The overall accuracy at the wetland class-level was 72% (Table 3) when compared to the field validation
303 data. Many of the discrepancies between the field class and the mapped class were the result of
304 confusion between the limnetic (L1) and littoral (L2) systems as well as confusion between the aquatic
305 bed (AB) and unconsolidated bottom (UB) classes.

306 There were 891 validation points in the image-interpreted dataset after excluding points within the
307 positional uncertainty of the imagery of a mapped wetland boundary. The overall image-interpretation
308 accuracy for discriminating between wetland and upland was 93% (Table 4). The wetland producer's
309 accuracy was 93% and the user's accuracy was 98%.

310 The overall accuracy at the wetland class-level was 78% (Table 5) when compared to the image-
311 interpretation validation data. As with the assessment using field data, many of the classification

312 discrepancies were associated with confusion between the limnetic and littoral subsystems as well as
313 confusion between the aquatic bed and unconsolidated bottom classes.

314 **Comparison to Original NWI**

315 The original NWI data for the 13-county project area has 125,586 wetland class features with a total
316 surface area of 2,958 square kilometers. Whereas, the updated NWI data for the same area includes
317 195,983 wetland class features with a total surface area of 3,104 square kilometers; an increase of 56%
318 for the number of wetland class features and an increase of 4.9% in wetland area. The increase in the
319 number of individual wetland class features suggests that the updated NWI was better able to
320 distinguish between wetland habitat classes within wetland complexes, identifying more wetland
321 polygons with less cross-class aggregation. However, an increase of total wetland area of 4.9% over a
322 period where urban development is widely believed to have resulted in wetland loss suggests that the
323 updated wetland inventory also mapped many wetlands that were missed in the original inventory. A
324 visual comparison of the results also supports this conclusion as well as clearly showing a more precise
325 boundary placement (Figure 3).

326 Using our validation data, we found that present-day feature-level accuracy of the original NWI is 76%
327 based on the image-interpreted validation data and 75% based on the field validation data (Table 6). The
328 updated wetland inventory described here has significantly better accuracy for upland-wetland
329 discrimination for present-day users. Likewise, the class-level accuracy for the updated NWI is also
330 better than the original NWI for present-day users. The class-level accuracy increased by 19% based on
331 the field validation data while it increased by 26% based on the image-interpreted validation data. To be
332 fair, we recognize that the original NWI has a much lower accuracy at the present time in large part due
333 to its age as well as from differences in the technical mapping approach.

34 **DISCUSSION**

335 **Automation Efforts**

336 Past efforts using automated classification of remote sensing data for the NWI have largely focused on
337 the use of relatively coarse resolution, optical satellite imagery data (Tiner 1990; FGDC 1999; Ozesmi
338 and Bauer 2002). Mapping and classifying wetlands to the Cowardin classification system used in the
339 NWI is inherently difficult due to the number of classes, sub-classes and modifiers and the temporal
340 variability associated with wetlands. Therefore, we opted not to attempt to fully automate the
341 classification process; instead we designed the automation strategy around making the human image
342 interpretation process more efficient. By automating the most time-consuming part of the image
343 interpretations, initial delineation of boundaries and identifying broad wetland classes, we were able to
344 allow the image interpreters to focus more of their efforts on the most difficult components of the
345 process, such as the assignment of detailed wetland classes and modifiers.

346 A significant task during this project was adapting automation techniques developed in a research
347 setting (Corcoran et al. 2011, Knight et al. 2013, Corcoran et al. 2013, Rampi et al. 2014a) for use in
348 production over a large area. The effort allocated to building, testing and refining the automation steps
349 required an up-front investment, but the labor saved during the image interpretation process resulted in
350 a net gain in efficiency. Rampi et al. (2014a) used a similar automated method for a simple four-class
351 map without subsequent manual photo-interpretation, achieving overall accuracies for wetlands in the
352 range of 96-98 percent. These results support our assertion that the initial wetland mapping steps can
353 be partially automated, while leaving the more detailed classification steps to human photo-
354 interpreters. This strategy provides improvements in overall efficiency while still maintaining high
355 standards for spatial resolution, classification detail, and accuracy.

356 **Accuracy Assessment**

357 The federal wetland mapping standard provides recommendations on map accuracy goals but little
358 specific guidance on how to conduct wetland mapping accuracy assessments. There are many design
359 decisions involved in developing an accuracy assessment method for a remote sensing wetland
360 inventory that can influence the results. We used two different validation data sets with different
361 methods of acquisition, one using field data and another using image-interpreted data. Simply changing
362 the data acquisition method resulted in a difference in the overall accuracy of 3% at the feature level
363 and 6% at the class-level. Changes in a number of other variables such as the distribution across the
364 sampling strata or the threshold used for screening out the effects of position uncertainty would affect
365 the calculation of final map accuracy values. Comparing accuracy results from one project to the next
366 will be difficult without some additional standardization for the accuracy assessment method.

367 The federal wetland mapping standard does not address errors of commission. The standard states that
368 98% of all wetlands “visible on an image” and larger than 0.2 ha shall be mapped (FGDC 2009). Based on
369 this, the producer’s accuracy for this project fell 5% short of the requirement. However, the federal
370 wetland mapping standard only specifies a threshold for errors of omission and not errors of
371 commission. A user’s accuracy of 98% carries no weight with respect to the federal wetland mapping
372 standard, but clearly it is an important consideration for the end users of the data. Without specific
373 quantification of commission errors, it is possible to bias a mapping project toward meeting the federal
374 standards by intentionally over-classifying upland features as wetlands. The federal standard also calls
375 for 85% attribute accuracy for wetland classes, but it is not clear whether this is intended to be a
376 standard for the overall class accuracy or the user’s or producer’s accuracy on individual wetland
377 classes.

378 There is an important relationship between class accuracy, the number of classes mapped, and how
379 distinct these classes are. In the present case, the overall class accuracy for this project is 78%, but some
380 of the observed classification error is certainly due to confusion between highly similar or temporally

31 variable wetland classes. For example, the distinction between the limnetic and littoral systems is
382 primarily based on water depth. The portion of a lacustrine system deeper than 2 meters is defined as
383 limnetic; whereas the portion shallower than 2 meters is defined as littoral (Cowardin et al. 1979).
384 Accurate classification of limnetic and littoral areas is very difficult without bathymetric survey data
385 (Irish and Lillycrop 1999; Dost and Mannaerts 2008). Not only are the optical imagery, near-infrared
386 lidar, and radar data used in this mapping effort limited in their ability to assess water depth, but also,
387 the field validation data were acquired from shore. As a result, it is difficult to determine whether the
388 error lies within the field data or the map data. In another example, the distinction between aquatic bed
389 and unconsolidated bottom wetland classes is defined by the presence or absence of rooted aquatic
390 vegetation. The confusion between these classes likely arises in large part due to the dynamic nature of
391 aquatic vegetation. Aquatic vegetation may be present in one part of the wetland in a given year (or
392 season within a year) and then appear in a different part of the same wetland in another year. Given the
393 expense and difficulty associated with separating out some of the wetland classes in the Cowardin
394 system, if a high level of accuracy for individual wetland classes is desired, it would be preferable to
395 simplify the classification by aggregating some classes.

396 This mapping effort exceeded many of the input data requirements of the federal wetland mapping
397 standard. The base imagery exceeded both the spectral and spatial resolution requirements as well as
398 the positional accuracy requirement. The input data requirements were also exceeded by including
399 datasets like lidar, radar, and multi-temporal imagery. Given the unusually high quality and richness of
400 the source data used in this project, the results raise the question whether it is practically feasible to
401 achieve the federal wetland mapping standard in large scale wetland mapping projects.

402 In addition to the above observations about issues with the interpretation and application of the federal
403 wetland mapping standard, another key result from this work was to quantify the overall improvement
404 in accuracy resulting from the update of the wetland inventory. Our results showed that when

405 compared to current field data we achieved a 15% increase in wetland-upland discrimination and a 19%
406 increase in wetland class accuracy. We have previously noted that this was not meant to be an
407 assessment of the accuracy of the original NWI at the time of its creation. It seems likely that the original
408 NWI had a higher accuracy at the time it was created. However, it is also important to note that in the
409 absence of an updated wetland inventory, people will continue to use the original NWI to assess current
410 conditions. Continuing to use inaccurate and outdated data results is likely to result in unnecessary
411 effort or inadequate wetland protection. The updated NWI provides a better source of information from
412 which to base present day natural resource management decisions.

413 In conclusion, we believe these results show that it is possible to produce high quality wetland
414 inventories using a semi-automated process that will meet many, if not all, of the needs stated in the
415 beginning of this paper. With the limited funding for these types of mapping efforts, additional work is
416 needed to continue to increase the efficiency of wetland mapping, while at the same time producing
417 results that meet the needs of the resource managers. Also, there is a need to refine and standardize
418 wetland mapping accuracy assessment methods. Furthermore, detailed accuracy assessment results,
419 such as presented here, provide important information to users who seek to understand the potential
420 limitations of remotely sensed wetland inventory data.

421 **ACKNOWLEDGEMENTS**

422 Funding for this project was provided by the Minnesota Environment and Natural Resources Trust Fund
423 as recommended by the Legislative-Citizen Commission on Minnesota Resources. The Trust Fund is a
424 permanent fund constitutionally established by the citizens of Minnesota to assist in the *protection,*
425 *conservation, preservation, and enhancement of the state's air, water, land, fish, wildlife, and other*
426 *natural resources.* Special thanks to Molly Martin of the Minnesota Department of Natural Resources for
427 technical and field assistance.

28 REFERENCES

- 429 Alaska Satellite Facility (2011). MapReady Remote Sensing Tool Kit Version 2.3.17 [computer
430 software]. Available from <https://www.asf.alaska.edu/data-tools/mapready/>
- 431 Austin, J. E., Buhl, T. K., Guntenspergen, G. R., Norling, W., and Sklebar, H. T. (2001). Duck populations as
432 indicators of landscape condition in the prairie pothole region. *Environmental Monitoring and*
433 *Assessment*, 69(1), 29-48.
- 434 Baatz, M., and Schäpe, A. (2000). Multiresolution segmentation: An optimization approach for high
435 quality multi-scale image segmentation. *Angewandte Geographische Informationsverarbeitung XII*
436 (J. Strobl and T. Blaschke, editors), Wichmann, Heidelberg, pp. 12–23.
- 437 Ball Aerospace & Technologies Corp (2011). Opticks Version 4.7.1. Available from <http://opticks.org>
- 38 Bergeson, M.T. (2011). Wetlands Data Verification Toolset: Installation Instructions and User
439 Information. U.S. Fish and Wildlife Service, National Standards and Support Team. Retrieved from
440 [http://www.fws.gov/wetlands/Data/tools/Wetlands-Data-Verification-Toolset-Installation-Instructions-](http://www.fws.gov/wetlands/Data/tools/Wetlands-Data-Verification-Toolset-Installation-Instructions-and-User-Information.pdf)
441 [and-User-Information.pdf](http://www.fws.gov/wetlands/Data/tools/Wetlands-Data-Verification-Toolset-Installation-Instructions-and-User-Information.pdf)
- 442 Beven, K. J., and Kirkby, M. J. (1979). A physically based, variable contributing area model of basin
443 hydrology. *Hydrological Sciences Journal*, 24(1), 43-69.
- 444 Bourgeau-Chavez, L.L., Kowalski, K.P., Carlson Mazur, M.L., Scarbrough, K.A., Powell, R.B., Brooks, C.N.,
445 Huberty, B., Jenkins, L.K., Banda, E.C., Galbraith, D. M., Laubach, Z.M. and Riordan, K. (2013). Mapping
446 Invasive *Phragmites australis* in the Coastal Great Lakes with ALOS PALSAR Satellite Imagery for Decision
447 Support. *Journal of Great Lakes Research*, 39(1), 65-77.
- 448 Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.

449 Brisco, B., Touzi, R., Van der Sanden, J., Charbonneau, F., Pultz, T., and D'lorio, M. (2007). Water
450 resource applications with Radarsat-2. *International Journal of Digital Earth*, 1(1): 130-147.

451 Burkett, V., and Kusler, J. (2000). Climate Change: Potential Impacts and Interactions in Wetlands of the
452 United States. *Journal of the American Water Resources Association*. 36(2) 313-320.

453 Congalton, R. G., and Green, K. (2008). Assessing the accuracy of remotely sensed data: principles and
454 practices. CRC press.

455 Corcoran, J.M, Knight, J.F., B. Brisco, S. Kaya, A. Cull, K. Murhnaghan (2011) The integration of optical,
456 topographic, and radar data for wetland mapping in northern Minnesota. *Canadian Journal of Remote*
457 *Sensing*, 27(5): 564-582.

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

461 Costa, L. T., Farinha, J. C., Vives, P. T., Hecker, N., & Silva, E. P. (2001). Regional wetland inventory
462 approaches: The Mediterranean example. In *Wetland Inventory, Assessment and Monitoring: Practical*
463 *Techniques and Identification of Major Issues*. Finlayson CM, Davidson NC & Stevenson NJ (eds).
464 Proceedings of Workshop 4, 2nd International Conference on Wetlands and Development, Dakar,
465 Senegal, 8–14 November 1998, Supervising Scientist Report 161, Supervising Scientist, Darwin.

466 Cowardin, L.M., Carter, V., Golet, F.C. and LaRoe, E.T. (1979). Classification of Wetlands and Deepwater
467 Habitats of the United States. U.S. Fish and Wildlife Service Report No. FWS/OBS/-79/31. Washington,
468 D.C.

469 Dahl, T.E., Dick, J., Swords, J. and Wilen, B.O. (2009). Data Collection Requirements and Procedures for
470 Mapping Wetland, Deepwater and Related Habitats of the United States. Division of Habitat and

71 Resource Conservation, National Standards and Support Team, Madison, WI. U.S. Fish and Wildlife
472 Service 85 pp.

473 Definiens, A. G. (2009). Definiens eCognition Developer 8 User Guide. Definiens AG, Munchen, Germany.

474 Drazkowski, B., May, M., and Herrera, D.T. (2004). Comparison of 1983 and 1997 Southern Michigan
475 National Wetland Inventory Data. Michigan Department of Environmental Quality, Geological and Land
476 Management Division. 15 pp.

477 Dost, R. J. J., & Mannaerts, C. M. M. (2008). Generation of lake bathymetry using sonar, satellite imagery
478 and GIS. In *ESRI 2008: Proceedings of the 2008 ESRI International User Conference*

479 ERDAS (2008). ERDAS Imagine Version 9.2 [computer software]. Intergraph Corporation. Norcross, GA.

480 ERDAS (2010). StereoAnalyst for ArcGIS Version 10.0 [computer software]. Intergraph Corporation.
481 Norcross, GA.

482 ESRI (2011). ArcGIS Desktop: Release 10 [computer software]. Environmental Systems Research
483 Institute. Redlands, CA.

484 Euliss Jr, N. H., Gleason, R. A., Olness, A., McDougal, R. L., Murkin, H. R., Robarts, R. D., Bourbonniere,
485 R.A., and Warner, B. G. (2006). North American prairie wetlands are important nonforested land-based
486 carbon storage sites. *Science of the Total Environment*, 361(1), 179-188.

487 FGDC (2009). Wetlands mapping standard. US Geological Survey Federal Geographic Data Committee,
488 Reston, Document Number FGDC-STD-015-2009.

489 Finlayson, C.M. and Spiers, A.G. (1999). Global Review of Wetland Resources and Priorities for Wetland
490 Inventory. Supervising Scientist Report 144. Wetlands International Publication 53, Supervising Scientist,
491 Canberra, Australia.

492 Frohn, R. C., Reif, M., Lane, C., and Autrey, B. (2009). Satellite remote sensing of isolated wetlands using
493 object-oriented classification of Landsat-7 data. *Wetlands*, 29(3), 931-941.

494 Fournier, R. A., Grenier, M., Lavoie, A., & Hélie, R. (2007). Towards a strategy to implement the Canadian
495 Wetland Inventory using satellite remote sensing. *Canadian Journal of Remote Sensing*, 33(S1), S1-S16.

496 Gong, P., Niu, Z.G., Cheng, X., Zhao, K.Y., Zhou, D.M., Guo, J.H., Liang, L., Wang, X.F., Li, D.D., Huang, H.B.,
497 Wang, Y., Wang, K., Li, W.N., Wang, X.W., Ying, Q., Yang, Z.Z., Ye, Y.F., Li, Z., Zhuang, D.F., Chi, Y.B., Zhou,
498 H.Z. And Yan, J. (2010). China's wetland change (1990–2000) determined by remote sensing. *Science*
499 *China Earth Sciences*, 53(7), 1036-1042.

500 Guisan, A., Weiss, S. B., and Weiss, A. D. (1999). GLM versus CCA spatial modeling of plant species
501 distribution. *Plant Ecology*, 143(1), 107-122.

502 Gwin, S. E., Kentula, M. E., and Shaffer, P. W. (1999). Evaluating the effects of wetland regulation
503 through hydrogeomorphic classification and landscape profiles. *Wetlands*, 19(3), 477-489.

504 Hepinstall, J. A., Queen, L. P., and Jordan, P. A. (1996). Application of a modified habitat suitability index
505 model for moose. *Photogrammetric engineering and remote sensing*, 62(11), 1281-1286.

506 Hogg, A. R., and Todd, K. W. (2007). Automated discrimination of upland and wetland using terrain
507 derivatives. *Canadian Journal of Remote Sensing*, 33(S1), S68-S83.

508 Irish, J. L., & Lillycrop, W. J. (1999). Scanning laser mapping of the coastal zone: the SHOALS system.
509 *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2), 123-129.

510 Kloiber, S.M. (2010). Technical Procedures for the Minnesota Wetland Status and Trends Program:
511 Wetland Quantity Assessment. Minnesota Department of Natural Resources, Division of Ecological and
512 Water Resources Report, November 2012. 41 pp. Retrieved from

13 Knight, J.F., B. Tolcser, J. Corcoran, and L. Rampi. (2013) The effects of data selection and thematic detail
514 on the accuracy of high spatial resolution wetland classifications. *Photogrammetric Engineering and*
515 *Remote Sensing*, 79(7): 613-623.

516 Knutson, M. G., Sauer, J. R., Olsen, D. A., Mossman, M. J., Hemesath, L. M., and Lannoo, M. J. (1999).
517 Effects of landscape composition and wetland fragmentation on frog and toad abundance and species
518 richness in Iowa and Wisconsin, USA. *Conservation Biology*, 13(6), 1437-1446.

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

521 LMIC. (2007). (Now called Minnesota Geographic Information Office) Metadata for the National Wetlands
522 Inventory, Minnesota. Retrieved from <http://www.mngeo.state.mn.us/chouse/metadata/nwi.html>

523 Macleod, R.D., Paige, R.S., and Smith, A.J. 2013. Updating the National Wetland Inventory in East-Central
524 Minnesota: Technical Documentation. Minnesota Department of Natural Resources, Division of
525 Ecological and Water Resources. St. Paul, MN. Retrieved from
526 http://files.dnr.state.mn.us/eco/wetlands/nwi_ecmn_technical_documentation.pdf

527 Maidment, D. R. (Ed.). (2002). *Arc Hydro: GIS for water resources* (Vol. 1). ESRI, Inc. Redlands, CA.

528 Marchand, M. N., and Litvaitis, J. A. (2004). Effects of habitat features and landscape composition on the
529 population structure of a common aquatic turtle in a region undergoing rapid development.
530 *Conservation Biology*, 18(3), 758-767.

531 Minnesota Climatology Working Group. 2012. 1981-2010 Normal Precipitation Maps. Retrieved from
532 http://www.climate.umn.edu/doc/historical/81-10_precip_norm.htm

533 MnGeo. 2010. Map Accuracy Report: East-Central Minnesota Ortho Project. Minnesota Geographic
534 Information Office, St. Paul, MN. Retrieved from

535 http://www.mngeo.state.mn.us/chouse/airphoto/2010_DOQ_Accuracy_Report_MnDOT_ecmn2010_03
536 Jan2012.pdf

537 Moore, I.D., Grayson, R.B. and Ladson, A.R. (1991). Digital Terrain Modelling: A Review of Hydrological,
538 Geomorphological, and Biological Applications. *Hydrological Processes*, 5:3-30.

539 Munger, J.C., Gerber, M., Madrid, K., Carroll, M., Petersen, W., and Heberger, L. (1998). U.S. National
540 Wetland Inventory Classifications as Predictors of the Occurrence of Columbia Spotted Frogs (*Rana*
541 *Luteiventris*) and Pacific Treefrogs (*Hyla regilla*). *Conservation Biology*, 12(2), 320-330.

542 Murphy, P. N., Ogilvie, J., Connor, K., and Arp, P. A. (2007). Mapping wetlands: a comparison of two
543 different approaches for New Brunswick, Canada. *Wetlands*, 27(4), 846-854.

544 Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., and Weng, Q. (2011). Per-pixel vs. object-based
545 classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of*
546 *Environment*, 115(5), 1145-1161.

547 NRCS (2010). Natural Resources Conservation Service, Soil Survey Staff, United States Department of
548 Agriculture. Soil Survey Geographic (SSURGO) Database for [Anoka, Carver, Chisago, Dakota, Goodhue,
549 Hennepin, Isanti, Ramsey, Rice, Scott, Washington, and Wright Counties, Minnesota]. Retrieved from
550 at <http://soildatamart.nrcs.usda.gov/> on 01/12/2010.

551 NSGIC (2014). Letter from President Kenneth Miller to the Secretary of Interior, Sally Jewell, April 16,
552 2014. National States Geographic Information Council.

553 Ozesmi, S. L., and Bauer, M. E. (2002). Satellite remote sensing of wetlands. *Wetlands Ecology and*
554 *Management*, 10(5), 381-402.

55 R Development Core Team. (2011). R: A language and environment for statistical computing, reference
556 index version 2.12. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0,
557 URL <http://www.R-project.org>.

558 Rampi, L.P., Knight, J.F., and Pelletier, K.C. 2014a. Wetland mapping in the Upper Midwest United
559 States: An object-based approach integrating lidar and imagery data. *Photogrammetric Engineering and*
560 *Remote Sensing*. 80(5), 439-449.

561 Rampi, L.P., Knight, J.F., and Lenhart, C.F., 2014b. Comparison of flow direction algorithms in the
562 application of the CTI for mapping wetlands in Minnesota. *Wetlands*, Feb. 2014.

563 Tarboton, D. G. (2003). Terrain analysis using digital elevation models in hydrology. In 23rd ESRI
564 international users conference, San Diego, California (Vol. 14).

565 Tiner, R. W. (1990). Use of high-altitude aerial photography for inventorying forested wetlands in the
66 United States. *Forest Ecology and Management*, 33, 593-604.

567 Tiner, R. W. (2009). Status report for the National Wetlands Inventory Program 2009. US Fish and
568 Wildlife Service, Branch of Resource and Mapping Support, Arlington, VA.

569 Trimble (2010). *eCognition Developer 8.64.0 User Guide*; Trimble: Munich, Germany. 27pp.

570 Yerkes, T., Paige, R., Macleod, R., Armstrong, L., Soulliere, G., and Gatti, R. (2007). Predicted distribution
571 and characteristics of wetlands used by mallard pairs in five Great Lakes States. *The American Midland*
572 *Naturalist*. 157(2), pp. 356-364.

TABLES

Table 1

Class Code	Class Description
L1UB	Lacustrine Limnetic Unconsolidated Bottom
L2AB	Lacustrine Littoral Aquatic Bed
L2EM	Lacustrine Littoral Emergent
L2UB	Lacustrine Littoral Unconsolidated Bottom
L2US	Lacustrine Littoral Unconsolidated Shore
PAB	Palustrine Aquatic Bed
PEM	Palustrine Emergent
PFO	Palustrine Forested
PSS	Palustrine Scrub-Shrub
PUB	Palustrine Unconsolidated Bottom
R2AB	Riverine Lower Perennial Aquatic Bed
R2UB	Riverine Lower Perennial Unconsolidated Bottom
R2US	Riverine Lower Perennial Unconsolidated Shore
UPL	Upland

Table 2

Reference Determination	Map Determination		
	Upland	Wetland	Total
Upland	201	18	219
Wetland	54	470	524
Total	255	488	743

Overall Accuracy	90%
Wetland Producer's Accuracy	90%
Wetland User's Accuracy	96%

Table 3

Reference Class	Map Class												Total
	L1UB	L2AB	L2EM	L2UB	PAB	PEM	PFO	PSS	PUB	R2AB	R2UB	UPL	
L1UB	1												1
L2AB	2	14		2	2							1	21
L2EM													0
L2UB	2			21								1	24
PAB		7		2	24	3			27				5
PEM		1			3	130	1	3	6			1	37
PFO						2	22	6					24
PSS							8	6	18				13
PUB				1	3				27				3
R2AB													2
R2UB												12	3
UPL						6	7		1				223
Total	5	22	2	24	32	149	36	27	61	0	17	308	683

Table 4

Reference Determination	Map Determination		
	Upland	Wetland	Total
Upland	208	12	220
Wetland	47	624	671
Total	255	636	891

Overall Accuracy	93%
Wetland Producer's Accuracy	93%
Wetland User's Accuracy	98%

Table 5

Reference Class	Map Class														Total
	L1UB	L2AB	L2EM	L2UB	L2US	PAB	PEM	PFO	PSS	PUB	R2AB	R2UB	R2US	UPL	
L1UB	39			5								8			52
L2AB	2	26	9	3		1	4								45
L2EM															0
L2UB	5	3	3	31								3			45
L2US					1										1
PAB						21	5			11	1	1			39
PEM						2	99	2	1	1				18	123
PFO							1	30	3					19	53
PSS							13	2	20			1		7	43
PUB		1		1		22	7	1	1	142				5	180
R2AB															0
R2UB							2	2					58		62
R2US								1	1			6	6		14
UPL							5	5				1		208	219
Total	46	30	12	40	1	48	137	41	25	154	1	78	6	257	876

Table 6

	Original NWI	Updated NWI
Feature Accuracy		
Field	75%	90%
Image-interpreted	76%	93%
Class Accuracy		
Field	53%	72%
Image-interpreted	52%	78%

DRAFT

TABLE CAPTIONS

Table 1: Wetland class codes and associated descriptions from Cowardin et al. (1979) applicable to the study area.

Table 2: Accuracy comparison for wetland-upland discrimination using field validation data. Class agreement between the two datasets is indicated by the shaded cells in the table

Table 3: Accuracy comparison between the field validation class and the mapped wetland class in the updated NWI data. Class agreement between the two datasets is indicated by the shaded cells in the table.

Table 4: Accuracy comparison for wetland-upland discrimination using photo-interpreted validation data. Class agreement between the two datasets is indicated by the shaded cells in the table.

Table 5: Accuracy comparison between the image-interpreted validation class and the mapped wetland class in the updated NWI data. Class agreement between the two datasets is indicated by the shaded cells in the table.

Table 6: Comparison of present-day accuracy of the original NWI to the accuracy of the updated NWI.

FIGURE CAPTIONS

Figure 1: The project area includes thirteen counties in east-central Minnesota, USA.

Figure 2: Illustration of the image classification process showing (a) the infrared band from the spring imagery, (b) the lidar hillshade DEM, (c) initial image objects, (d) refined multi-resolution objects, and (e) the final wetland inventory map.

Figure 3: A comparison of the original NWI wetland boundaries (dashed black line) to the updated wetland boundaries (white line) shown on top of a lidar hillshade layer.

Figure 4 (electronic supplemental material - online only): A comparison of the original NWI wetland boundaries (green) to the updated wetland boundaries (blue) shown on top of a false color-infrared aerial image.

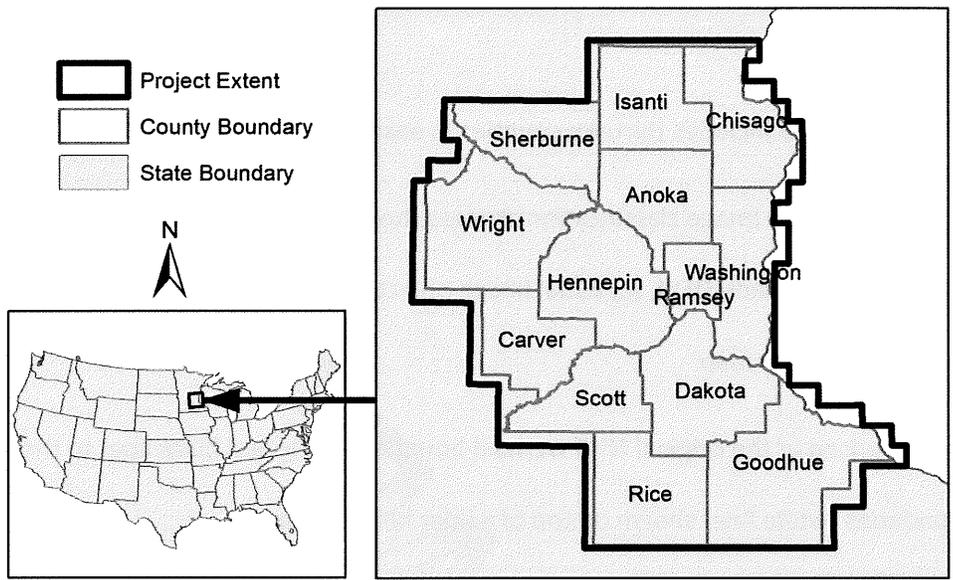


Figure 1: The project area includes thirteen counties in east-central Minnesota, USA.

DRAFT

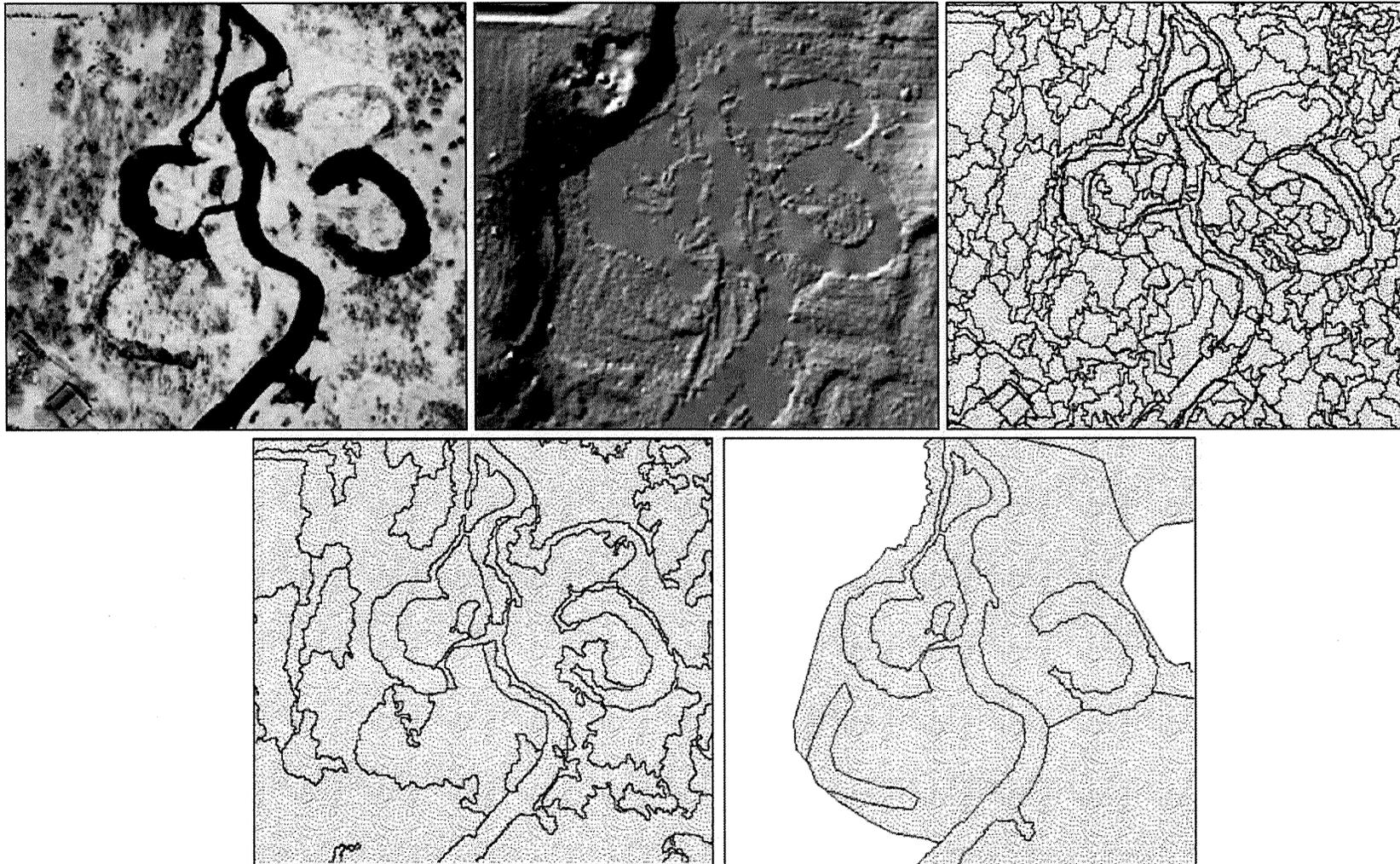


Figure 2: Illustration of the image classification process showing (a) the infrared band from the spring imagery, (b) the lidar hillshade DEM, (c) initial image objects, (d) refined multi-resolution objects, and (e) the final wetland inventory map.

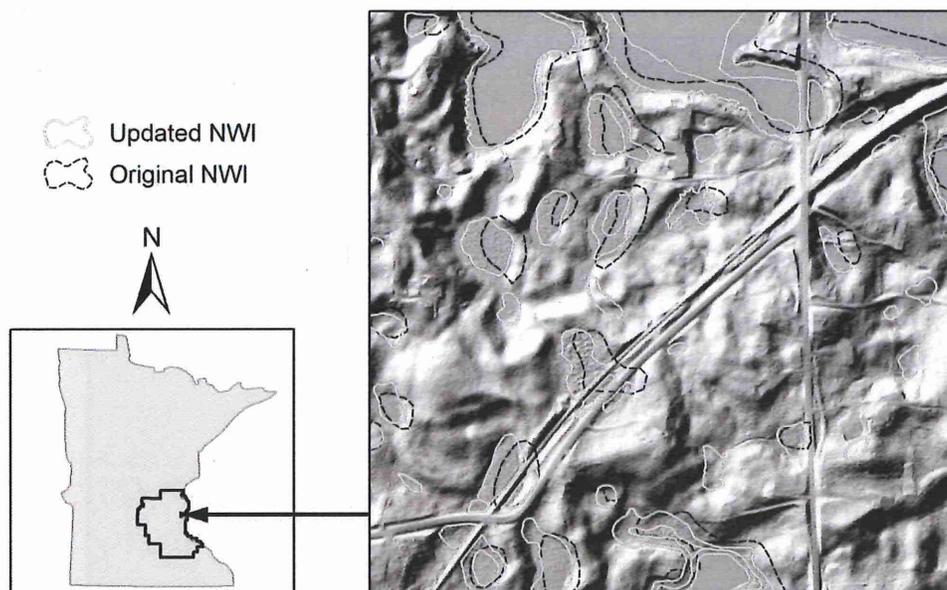


Figure 3: A comparison of the original NWI wetland boundaries (dashed black line) to the updated wetland boundaries (white line) shown on top of a lidar hillshade layer.

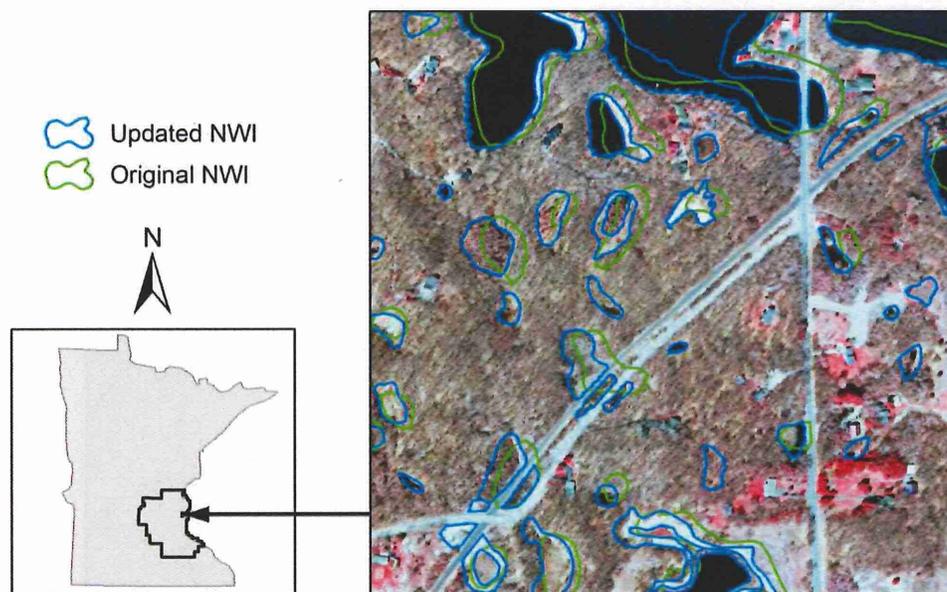


Figure 4 (electronic supplemental material - online only): A comparison of the original NWI wetland boundaries (green) to the updated wetland boundaries (blue) shown on top of a false color-infrared aerial image.

The integration of optical, topographic, and radar data for wetland mapping in northern Minnesota

Jennifer Corcoran, Joseph Knight, Brian Brisco, Shannon Kaya, Andrew Cull, and Kevin Murnaghan

Abstract. Accurate and current wetland maps are critical tools for water resources management, however, many existing wetland maps were created by manual interpretation of one aerial image for each area of interest. As such, these maps do not inherently contain information about the intra- and interannual hydrologic cycles of wetlands, which is important for effective wetland mapping. In this paper, several sources of remotely sensed data will be integrated and evaluated for their suitability to map wetlands in a forested region of northern Minnesota. These data include: aerial photographs from two different times of a growing season, National Elevation Dataset and topographical derivatives such as slope and curvature, and multitemporal satellite-based synthetic aperture radar (SAR) imagery and polarimetric decompositions. We identified the variables that are most important to accurately classify wetland from upland areas and discriminate between wetland types for a forested region of northern Minnesota using the decision-tree classifier randomForest. The classifier was able to differentiate wetland from upland and water with 75% accuracy using optical, topographic, and SAR data combined, compared with 72% using optical and topographical data alone. Classifying wetland type proved to be more challenging; however, the results were significantly improved over the original National Wetland Inventory classification of only 49% compared with 63% using optical, topographic, and SAR data combined. This paper illustrates that integration of remotely sensed data from multiple sensor platforms and over multiple periods during a growing season improved wetland mapping and wetland type classification in northern Minnesota.

Résumé. Les cartes précises et à jour des milieux humides sont des outils essentiels pour la gestion des ressources en eau; toutefois, de nombreuses cartes des milieux humides existantes furent créées à l'aide de l'interprétation manuelle d'une seule image aérienne pour chaque zone d'intérêt. Comme tel, ces cartes ne contiennent pas d'information inhérente sur les cycles hydrologiques intra- et interannuels des milieux humides qui constitue une information essentielle pour la cartographie efficace des milieux humides. Dans cet article, diverses sources de données de télédétection seront intégrées et évaluées pour leur capacité à cartographier les milieux humides dans une région boisée située dans le nord du Minnesota. Ces données incluent : des photographies aériennes acquises à deux périodes différentes de la saison de croissance, un ensemble de données « National Elevation Dataset » et des dérivées topographiques comme la pente et la courbure, des images satellite multi-temporelles radar à synthèse d'ouverture (RSO) ainsi que des décompositions polarimétriques. On identifie les variables les plus importantes pour la classification précise des milieux humides par rapport aux zones de hautes terres et pour la détermination des types de milieux humides pour une zone boisée dans le nord du Minnesota à l'aide du classifieur randomForest basé sur un arbre de décision. Le classifieur a permis de différencier les milieux humides des hautes terres et de l'eau avec une précision de 75 % en utilisant une combinaison de données optiques, topographiques et radar comparativement à 72 % en utilisant des données optiques et topographiques uniquement. La classification des types de milieux humides s'est avérée plus difficile à réaliser; cependant, les résultats étaient significativement meilleurs par rapport à la classification originale du « National Wetland Inventory » qui était de seulement 49 % comparativement à 63 % en utilisant une combinaison de données optiques, topographiques et radar. Globalement, on montre dans cet article que l'intégration des données de télédétection multi-capteurs et sur des périodes multiples durant la saison de croissance peut améliorer la cartographie des milieux humides ainsi que la classification des types de milieux humides dans le nord du Minnesota.

Introduction

Wetlands are valuable ecosystems in many ways. For example, wetlands provide filtration of wastewater (Vymazal,

2005), groundwater recharge (van der Kamp and Hayashi, 1998; Acharya and Barbier, 2000), and water retention to reduce damages caused by flooding (Mitsch and Gosselink, 2000). Accurate wetland maps are important

Received 1 April 2011. Accepted 5 December 2011. Published on the Web at <http://pubs.casi.ca/journal/cjrs> on 16 March 2012.

Jennifer Corcoran¹ and Joseph Knight. University of Minnesota, Department of Forest Resources, 115 Green Hall, 1530 Cleveland Avenue N, St. Paul, MN 55108.

Brian Brisco, Shannon Kaya, Andrew Cull, and Kevin Murnaghan. Canada Centre for Remote Sensing, Natural Resources Canada, 588 Booth Street, Ottawa, ON K1A 0Y7, Canada.

¹Corresponding author (e-mail: murph636@umn.edu).

for conservation and restoration efforts, and they are crucial for developing emergency response plans for natural disasters. For instance, in the Wild Rice River Watershed of the Red River Basin, which covers portions of the United States and Canada, agencies from two U.S. states, one Canadian province, and both national governments respond to the frequent and significant flood events on the Red River (Hearne, 2007). Though these agencies have different laws and methods for responding to extreme flooding, all require the most accurate and current water resource maps, and the techniques to create them, to assess and manage flood events.

A wetland map is a two-dimensional representation of a four-dimensional phenomenon (including space and time). Wetland boundaries are dynamic and fluctuate both inter- and intra-annually depending on many factors including rainfall, evaporation, ground water flow, and land use manipulation. Wetland inventories are important tools for managing and protecting wetlands. Therefore, the accuracy of wetland mapping methods is critically important for a broad range of water resource management concerns. Among these concerns are regulatory purposes such as permitting, mitigation, and monitoring compliance; monitoring of changes in wetland extent or function due to natural and anthropogenic causes; and selection of areas that have the most suitable hydrologic and vegetative characteristics for wetland restoration or conservation (Deschamps et al., 2002; Brooks et al., 2006; Hearne, 2007). Given the dynamic nature of wetlands, having access to a synoptic wetland inventory map is an important first step in making sound water resource management decisions.

The U.S. National Wetlands Inventory (NWI) from the U.S. Fish and Wildlife Services was not designed to show exact wetland boundaries, but rather to provide generalized boundaries and approximate locations in a snapshot of time (United States Fish and Wildlife Service, 2009). The U.S. Environmental Protection Agency called for achieving a net increase in wetland acres by 2011. Similarly, the State of Minnesota enacted the Wetland Conservation Act with a goal of “no net loss” in wetlands statewide. Both of the aforementioned goals require the continuous creation of robust maps of current wetlands and practical techniques for monitoring land use change impacts on wetlands over large geographical areas. Once presented with a reliable wetland inventory, water resource managers charged with accomplishing these regulatory goals can design adaptive management approaches to prioritize areas for conservation and restoration.

Traditional methods of mapping wetlands have relied on aerial photograph interpretation or classification of optical satellite imagery. However, such maps are typically based on single-date optical imagery, are often several years old, may not be representative of the current state of the environment, and do not take into account the dynamic nature of wetlands. One wetland type in particular that is problematic to map is forested wetlands. Separating forested wetlands from forested uplands with optical imagery is challenging

because the imagery, even if collected during leaf-off conditions, may not reveal the underlying hydrology of site. The collection of optical imagery can also be hindered by cloud cover, thus potentially missing the critical post-snow, leaf-off period for wetland inventory. Many wetlands are only flooded or saturated ephemerally, so those wetlands may not have been mapped in the original NWI (Dahl, 1990). Optical imagery may reveal these wetlands, if the timing of the imagery collection is perfect, but it is difficult to predict when that time is and to complete data acquisition during that time.

The addition of other remotely sensed data, such as radio detection and ranging (radar) data, can offer unique information about surface features beyond the radiometric response measured with optical data. This additional information can help to identify inundation (Hess et al., 2003; Frappart et al., 2005; Lu and Kwoun, 2008; Lane and D’Amico, 2010) and classify wetland areas (Touzi, 2006; Ban et al., 2010) based on surface structure and hydrologic features that may not otherwise be differentiable with aerial photography alone.

In certain areas and during periods of frequent cloud cover, optical wavelengths have an obvious disadvantage in that data cannot be acquired. Long-wave radar signals, on the other hand, are not sensitive to the atmosphere, do not require daylight hours for acquisition, and thereby increase the possibility for frequent data collection (Townsend, 2001; Parmuchi et al., 2002). In addition, polarimetric information from synthetic aperture radar (SAR) allows for the discrimination between different scattering mechanisms contributing to the overall backscatter in an image (Townsend, 2001; Parmuchi et al., 2002; Brisco et al., 2008). Polarimetric scattering signatures can be interpreted to identify landscape variables associated with the primary surface-scattering mechanism identified for each area through products known as polarimetric decompositions (Touzi et al., 2009).

Incorporating data from multiple sensor platforms and over multiple seasons will increase the likelihood of differentiating between a broader range of wetland types (Ramsey et al., 1995; Ozesmi and Bauer, 2002; Töyrä et al., 2002; Li and Chen, 2005; Castañeda and Ducrot, 2009; Ramsey et al., 2009; Bwangoy et al., 2010). By acquiring fully polarimetric SAR data from multiple dates over a season, the relative backscatter response from varying hydrologic periods and both leaf-on and leaf-off conditions can help determine the seasonality of wetlands and thus classify wetland types with higher accuracy. However, given the integration of such a large number of data inputs, it is important to determine the optimal set of data to reduce redundancy and increase the accuracy and efficiency of implementing mapping wetlands over large spatial scales, a goal that can be accomplished by decision-tree classification.

This study investigated how the accuracy of wetland mapping can be improved by integrating several sources of remotely sensed data, including: leaf-on and leaf-off high resolution aerial orthophotos, National Elevation Dataset

(NED) and topographic derivatives, and fully polarimetric SAR imagery. We address the following hypotheses: (i) seasonal fully polarimetric SAR imagery provides important information about surface scattering mechanisms, allowing more accurate distinction of wetland type; and (ii) the integration of optical, topographic, and SAR data using a decision-tree classifier provides a more accurate method for wetland mapping and classifying specific wetland types.

Methodology

Study site

This research focused on improving wetland classification accuracy, in particular the classification of forested wetlands in northern Minnesota. Minnesota is rich with geological history, containing volcanic and sedimentary rocks from millions of years ago. Much of the state has been carved by several glacial advances and retreats over the millennia, leaving glacial deposits, lakes, and rivers in their wake (MN DNR, 2011). Northeastern Minnesota, otherwise known as the Arrowhead, is a region currently dominated by hardwood and conifer forests, as well as woody and herbaceous wetlands (USDA-NASS, 2011). This region is sparsely populated, with the exception of a few larger cities near Lake Superior, namely Cloquet and Duluth, with populations of 12 124 and 86 265 in 2010, respectively (AdminMN, 2011). The chosen study site centered on Cloquet, Minn. is generally representative of the land cover characteristic of the Arrowhead region (Figure 1). The elevation in this study site ranges from about 330–450 m above sea level (mean of 392 m) and the slope of the landscape is on average less than 1.7°.

Classification schemes

Two classification schemes were used in this paper, including a simple upland/water/wetland determination and a modified version of the Cowardin classification (Cowardin, 1979). The modified Cowardin classification scheme involved reclassifying the following classes: flooded and intermittent lakes, unconsolidated bottom water bodies, and rivers merged into one “water” class; aquatic bed and emergent wetlands merged into “emergent wetlands”; “forested wetlands” and “scrub/shrub wetlands” remained the same; and all nonwetland areas were initially separated into “agriculture”, “forest”, “grassland”, “rural”, and “urban” classes for training the decision-tree classifier, then later merged into one “upland” class. The simple upland/water/wetland determinant classification was based on appropriate consolidation of the aforementioned upland and Cowardin wetland classes.

Field data

Multiple sets of field data were used in this research, including field point data collected in the summers of 2009 and 2010 and the MN DNR Wetland Status and Trends Monitoring Program (WSTMP) polygons from 2006–2008 (Table 1). The WSTMP polygons were created by randomly distributing 4990 one-square-mile primary sampling units (PSUs) statewide, divided into three panels. One panel was photographed with spring leaf-off (or early leaf on-set) high-resolution aerial photography each year and the PSUs were digitized using a Cowardin classification scheme by trained photo interpreters. The initial digitized polygons were then reviewed by a second team of senior photo interpreters and a subset of the PSUs was field-verified and used to evaluate

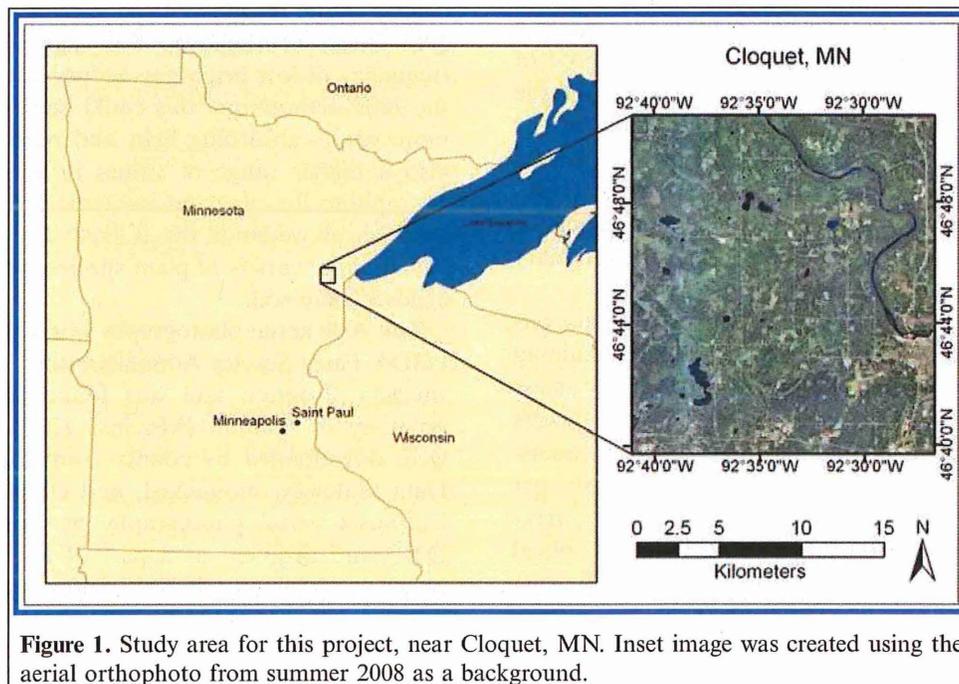


Figure 1. Study area for this project, near Cloquet, MN. Inset image was created using the aerial orthophoto from summer 2008 as a background.

For personal use only.

the accuracy of each panel. A general 30% rule was followed while digitizing the WSTMP polygons, in which if a land class appeared to occupy more than 30% of the polygon, then it was designated as that class. An exception to the 30% rule was where more than one class of vegetation existed; in this case, the taller plant class took precedence (MN DNR, 2010). The centroids of WSTMP polygons were used in addition to the field point data collected in 2009 and 2010.

The field data collection protocol for 2009 and 2010 involved the following: locating and physically visiting ground reference points with a GPS unit; recording the position with a minimum of 50 GPS fixes; identifying the dominant wetland type using the Cowardin classification scheme (Cowardin, 1979); taking representative photographs using the built-in camera on the GPS unit; and recording the point ID, description, and spatial coordinates in a field notebook for back-up purposes. In addition to the 2009, 2010, and WSTMP training data, additional points were added using manual photo interpretation to ensure a suitable distribution of points per class.

Decision-tree classification

A decision-tree classification approach provides an efficient means of establishing relationships between dependent and independent variables using training data, such that predictions can be repeatedly and robustly made of unclassified datasets (Hogg and Todd, 2007). Several decision-tree classification software programs are available, each having strengths and weaknesses regarding usability, accuracy, and performance (Ruefenacht et al., 2008). The decision-tree classifier randomForest was used in this research and was run using the R Statistical Package module within Python. RandomForest was chosen for this research because of the robustness of the results, ease and speed of use, and the ability to produce confidence maps of the classification results. Programming code was provided by the U.S. Department of Agriculture (USDA) Forest Service Remote Sensing Applications Center (Ruefenacht et al., 2008) to generate an output classification and associated confidence map, while R was used to compute summary statistics and figures about the classification results.

A stratified random sample of 75% of the field data was used to train the decision-tree classifier, while the remaining 25% were used as a reference dataset to independently assess the accuracy. For a summary of the number of points available per land class for training and accuracy assessment, see **Table 1**. Two decision trees were built per classification scheme, the first using optical and topographical data alone, the second using a combination of optical and topographic data as well as all available SAR data (including two dates of backscatter from four polarizations and four dates of three different polarimetric decompositions).

Table 1. Summary of decision-tree classifier training points and independent tests points for accuracy assessment of the results.

Land Cover Classification	No. of points		
	Training	Test	Total
Upland–wetland determinant			
Upland	464	152	616
Water	69	23	92
Wetland	421	140	561
Total	954	315	1269
Modified Cowardin class			
Water	69	23	92
Emergent wetland	97	37	134
Forested wetland	156	48	204
Scrub/shrub wetland	168	55	223
Upland	464	152	616
Total	954	315	1269

Datasets

Optical and topographical

Two periods of high-resolution aerial orthophotos were used in this research, including: 2008 mid-summer (full canopy, 1 m resolution) and 2009 spring (early leaf onset, 50 cm resolution) imagery (**Figure 2**). Both sets of imagery were acquired with color and near infrared bands. **Figure 3** shows the response signature, or frequency diagram, of the brightness values in each optical band at two different times for upland and wetland classified reference field sites. The decision-tree classifier will attempt to capitalize on the spectral differences between the land class types in each optical band. Rooted in the wetland response signatures are the response signatures of each wetland class, shown in **Figure 4**. There were noticeable differences between emergent wetlands and forested or scrub/shrub wetlands, particularly in the responses from early leaf-onset period of the 2009 aerial orthophotos. For example, there was a high frequency of low brightness values in emergent wetlands in the 2009 orthophoto, this could be due to wetter soils and more plants absorbing light and reflecting less. There was also a higher range of values in each band of the 2009 orthophoto for emergent wetlands compared to forested or scrub/shrub wetlands, this is likely due to emergent wetlands having more variety of plant species, wetness, and patches of exposed bare soil.

The 2008 aerial photography was acquired as a part of the USDA Farm Service Administration National Agricultural Imagery Program and was found to have a horizontal accuracy of 2.66 m (MnGeo, 2011). The 2008 imagery were downloaded by county from the USDA Geospatial Data Gateway, mosaicked, and clipped. The spring early leaf-onset aerial photography was provided by the MN DNR and acquired as a part of a collaboratively funded program between several state and federal agencies, including the MN DNR, MN Pollution Control Agency, and U.S. Geological Survey (USGS). All bands from each date of aerial orthophotos were used in the decision-tree

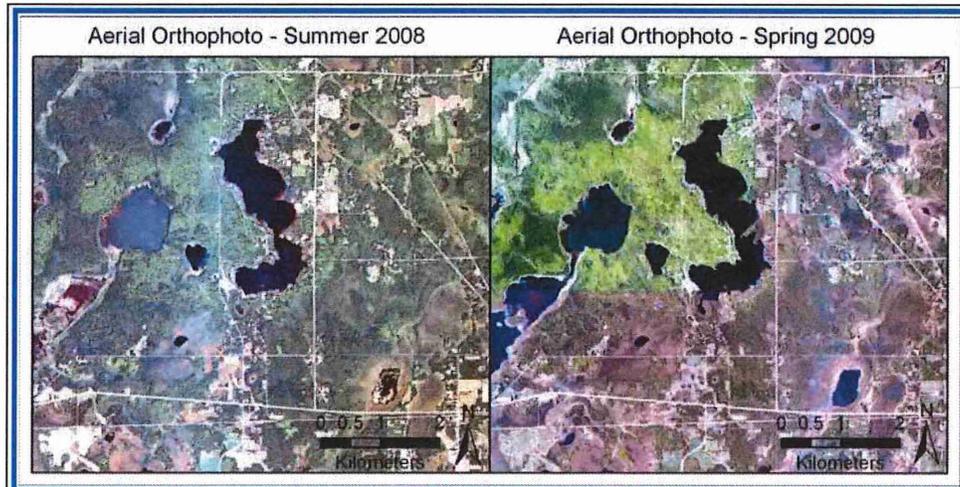


Figure 2. Two different time periods of aerial orthophoto imagery were used in the decision-tree classifier: full canopy in mid-summer 2008 and early leaf-onset in spring 2009. This subset area illustrates how the imagery for 2009 was collected at different times during the spring.

classification. In addition, the red and near infrared bands were used to calculate normalized difference vegetation index (NDVI) (Campbell, 2007). Because of a requirement of the decision-tree algorithm utilized in this research, the

imagery was degraded to mimic the minimum resolution (10 m) of all concurrent input datasets (Figure 5).

Wetlands tend to be located in low-lying or depressional areas on the landscape. Therefore, the USGS NED was

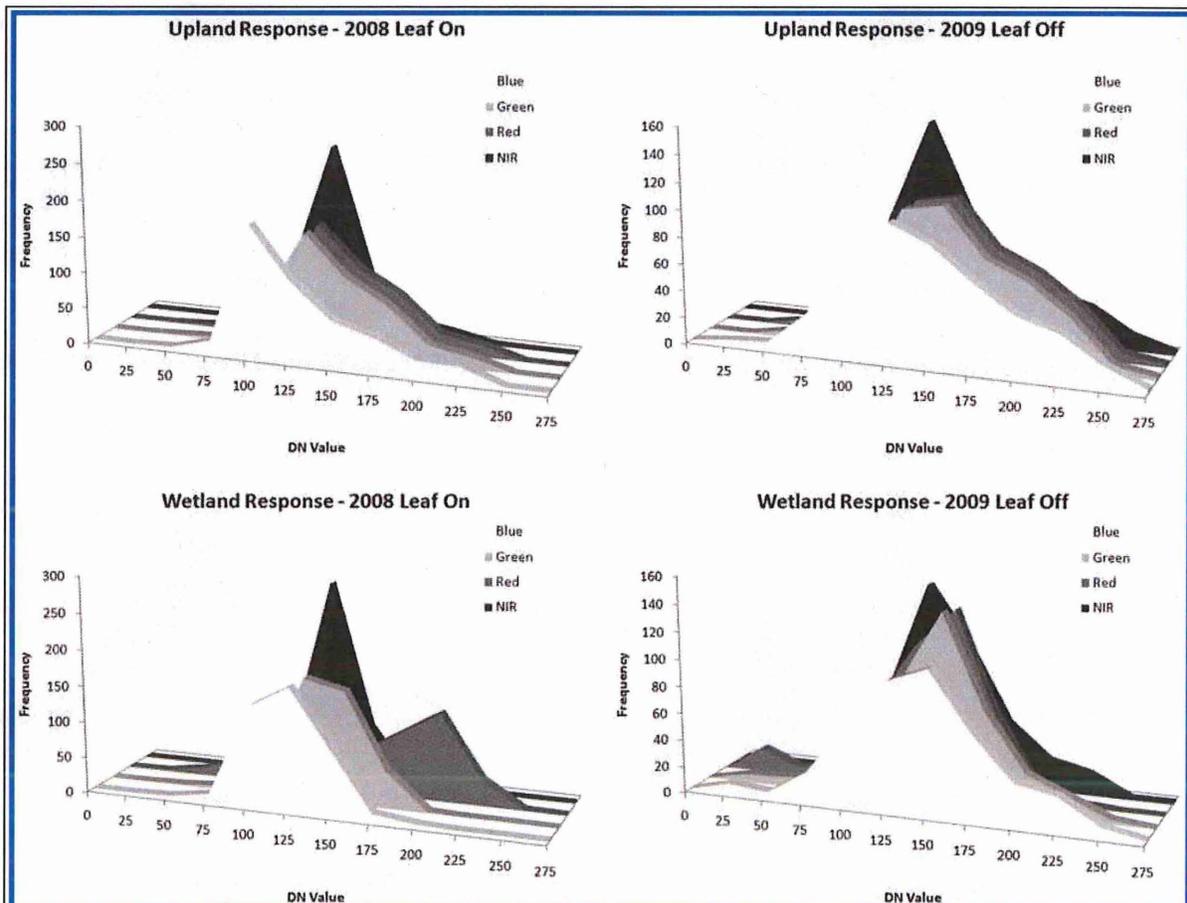


Figure 3. Spectral response signatures for each of the optical bands extracted from field reference data points of upland and wetland class categories and for each source of aerial orthophoto, summer 2008 and 2009 early leaf-onset.

For personal use only.

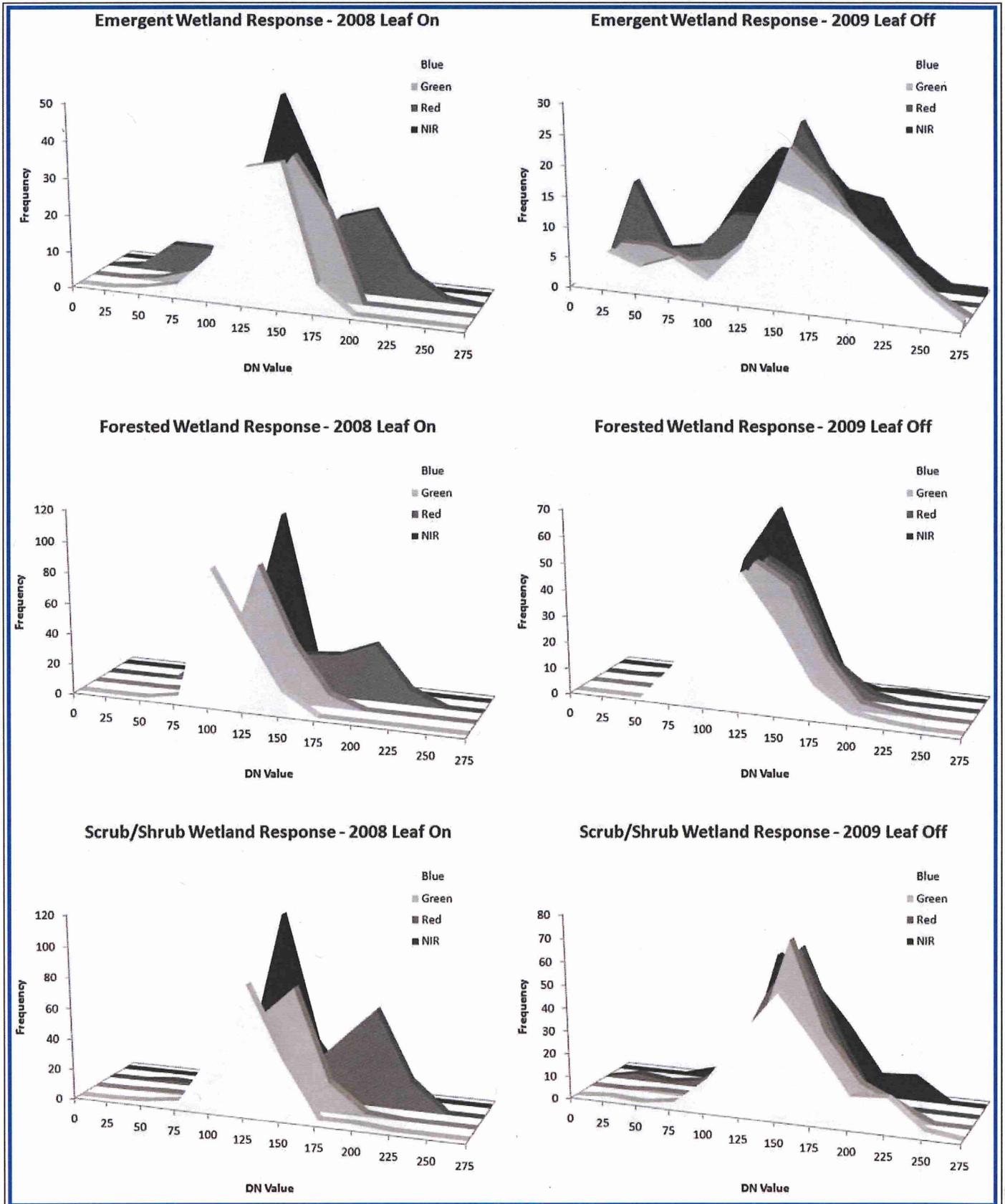


Figure 4. Spectral response signatures for each of the optical bands extracted from field reference data points of emergent, forested, and scrub/shrub wetland class types and for each orthophoto source: summer 2008 and 2009 early leaf-onset.

For personal use only.

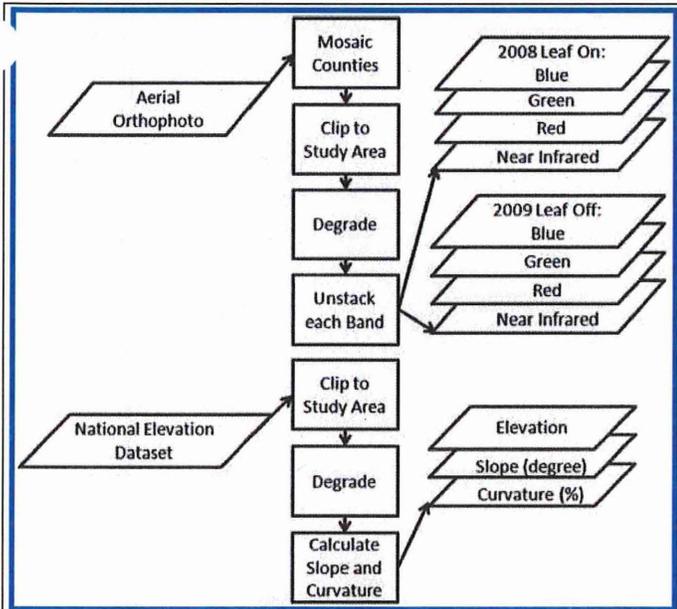


Figure 5. Optical and topographic input data, preprocessing methods, and output data for decision-tree classification.

obtained for the study area. The NED is available at 1/3 Arc Sec, or 10 m resolution, and was similarly downloaded by county from the USDA Geospatial Data Gateway, mosaicked, clipped, and degraded to mimic the resolution of the other input datasets. The mean horizontal accuracy of the NED was evaluated by the USGS using 13 305 geodetic control points nationwide and was found to be 2.44 m. The mean vertical accuracy was found to be 1.64 m based on 9187 unique pairs of geodetic reference points (USGS, 2011). The NED data was the coarsest resolution utilized in this research. Because of the requirements of the decision-tree algorithm used in this research, all other datasets were resampled to the same 10 m spatial resolution.

Slope and curvature were generated from the NED by running tools in the Environmental Systems Research Institute (ESRI) ArcGIS software. The assumption behind using these derived products from elevation data was that they contained additional information that describes physical characteristics about the drainage of water in a basin, where the slope of a landscape can affect the rate of flow of water and the curvature influences the convergence and divergence of that flow (Moore et al., 1991).

Radar

Two fully polarimetric C-band (5.6 cm wavelength) RADARSAT-2 look complex SAR images were obtained through the Canadian Space Agency's Science and Operational Applications Research (SOAR) Program. The dates of these images were 15 June and 19 September 2009. Fully polarimetric SAR imagery is collected with varying transmitted and received signal polarizations (horizontal–horizontal, HH; vertical–vertical, VV; horizontal–vertical,

HV; and vertical–horizontal, VH). In addition, two dates of polarimetric decomposition products were obtained through the Canada Center for Remote Sensing for 9 July and 26 August 2009. All SAR images for this research were acquired in fine quad-beam mode with near and far incidence angles of 26.9 and 28.7, respectively (FQ8). The constant beta look-up table was applied for calibration to avoid over saturation of the data (Kaya, 2010). A 7 × 7 boxcar filter was applied to each image to reduce speckle noise and increase the number of looks needed for polarimetric decomposition (Figure 6), and the images were resampled to 10 m spatial resolution. The digital number (DN) values, representing amplitude, were converted to sigma naught (σ^0) or backscattering coefficient in units of decibels for quantitative analysis (Parmuchi et al., 2002).

Response signatures of the SAR backscatter values for each polarization of two different periods in time for upland and wetland classified reference field sites are shown in Figure 7. As previously described for the optical response signatures, Figure 8 shows how each wetland class is represented in the SAR response signature of wetlands as a whole. There are only slight differences between the backscatter response values of each wetland type for the two periods of the season. For example, there was no change in the peak HH response of emergent wetlands; however, the range of backscatter values in September shifted down by 5 decibels compared with June, possibly indicating a change in physical characteristics of the vegetation present later in the season. The HV response had a similar shift in the range of backscatter response values, but the peak response was 5 decibels lower for emergent wetlands in September. Looking at the response signatures of scrub/shrub wetlands, there was no discernible difference between June and September backscatter values in each of the polarizations in terms of the peak backscatter or range of values. The range of HH backscatter values for forested wetlands was the same for June and September; however, the peak values similarly shifted 5 decibels lower in September compared with June. The VV response for forested wetlands had a 5 decibel shift downward in the range of backscatter values from June to September. Though these shifts in peak backscatter value

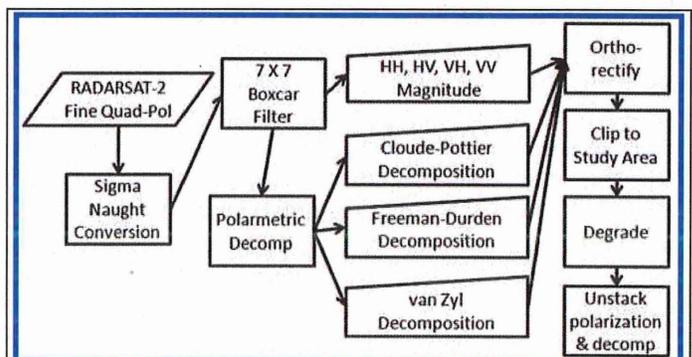
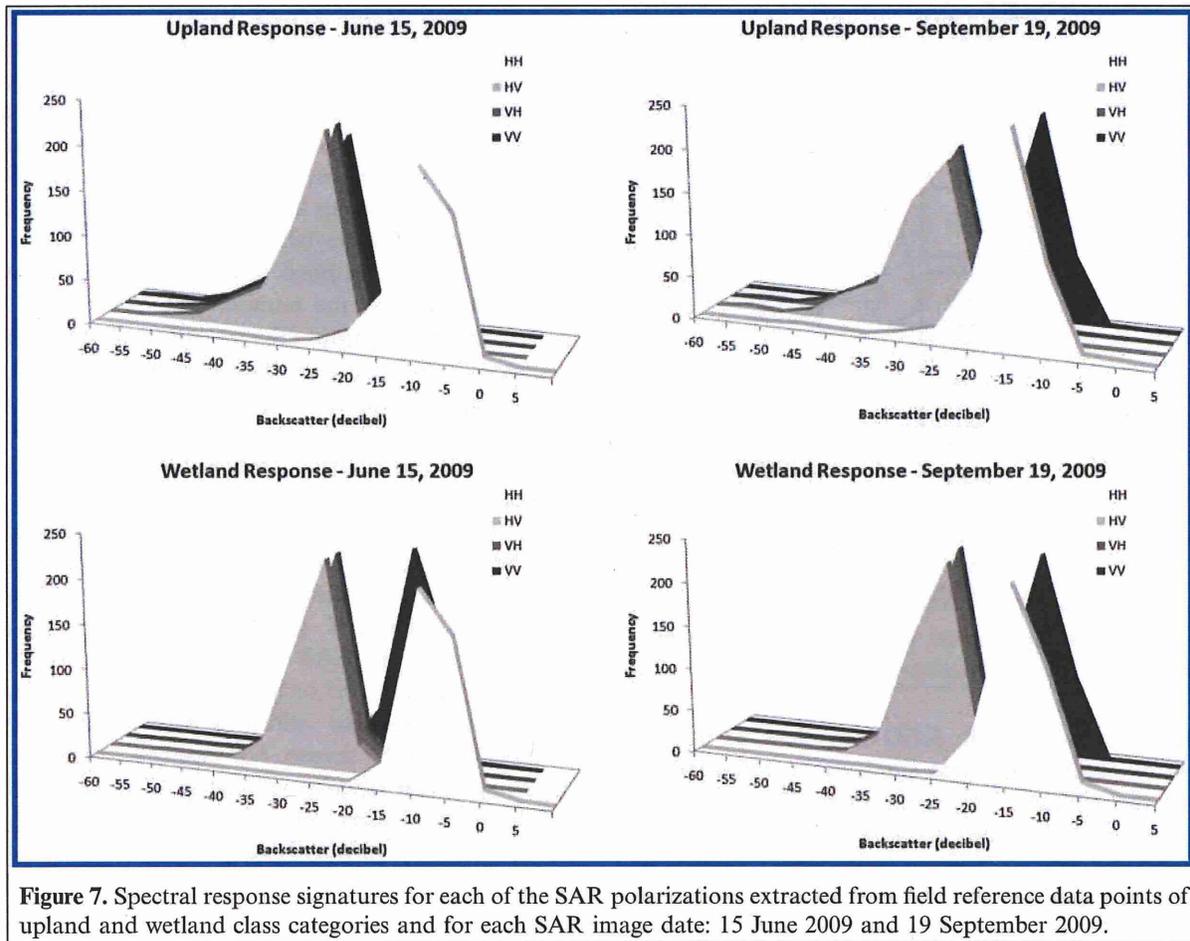


Figure 6. Synthetic aperture radar (SAR) input data, preprocessing methods, and output data for decision-tree classification.

For personal use only.



and the range and variability of backscatter values per image date, land class type, and polarization were subtle, the decision-tree classification method was expected to use this information to improve the results of the wetland classification.

Polarimetric decompositions were used to assess the importance of radar polarimetry on the accuracy of wetland mapping. Combinations and differences between the transmitted and received signal polarizations detected vegetation differences quite well (Baghdadi et al., 2001; Henderson and Lewis, 2008; Slatton et al., 2008). Many supervised and unsupervised algorithms have been developed to exploit multiple polarization data to distinguish physical features on the ground in a radar scene. This research used three of the most frequently used unsupervised polarimetric decompositions in the literature, including the Van Zyl, Freeman–Durden, and Cloude–Pottier decompositions and their related parameters. Each of the polarimetric decompositions was performed prior to orthorectification in an attempt to reduce resampling error, particularly in thematic decompositions.

The van Zyl polarimetric decomposition is an unsupervised thematic classification based on the phase and backscatter response of scattering targets on the ground (van Zyl, 1989). Each pixel is categorized as a single, odd, or diffuse

scatterer based on the number of phase shifts that occurred per pixel between co-polarized (HH and VV) scattering waves, where every scattering event is expected to add a 180° phase shift. The van Zyl decomposition product therefore is a single thematic layer per SAR image date.

The Freeman–Durden polarimetric decomposition is similar to van Zyl's in that it is a technique for identifying physically-based scattering mechanisms on the ground. However, the Freeman–Durden decomposition effectively breaks down the total backscatter for each pixel into relative portions of three scattering mechanisms: surface scatter, double bounce, and canopy scatter (or volume scatter). Each pixel then has a relative weight for each scattering mechanism, instead of a single category (Freeman and Durden, 1998). The Freeman–Durden decomposition product is therefore three layers of data per image date.

The third polarimetric decomposition utilized in this paper was presented by Cloude and Pottier (1997). In this decomposition, the parameters of entropy, alpha angle, and anisotropy are calculated from the eigenvalues and eigenvectors of the coherency matrix. Cloude and Pottier showed that these parameters represent different scattering mechanisms, directly relating to the affect that the physical structure of the target has on the received backscatter.

For personal use only.

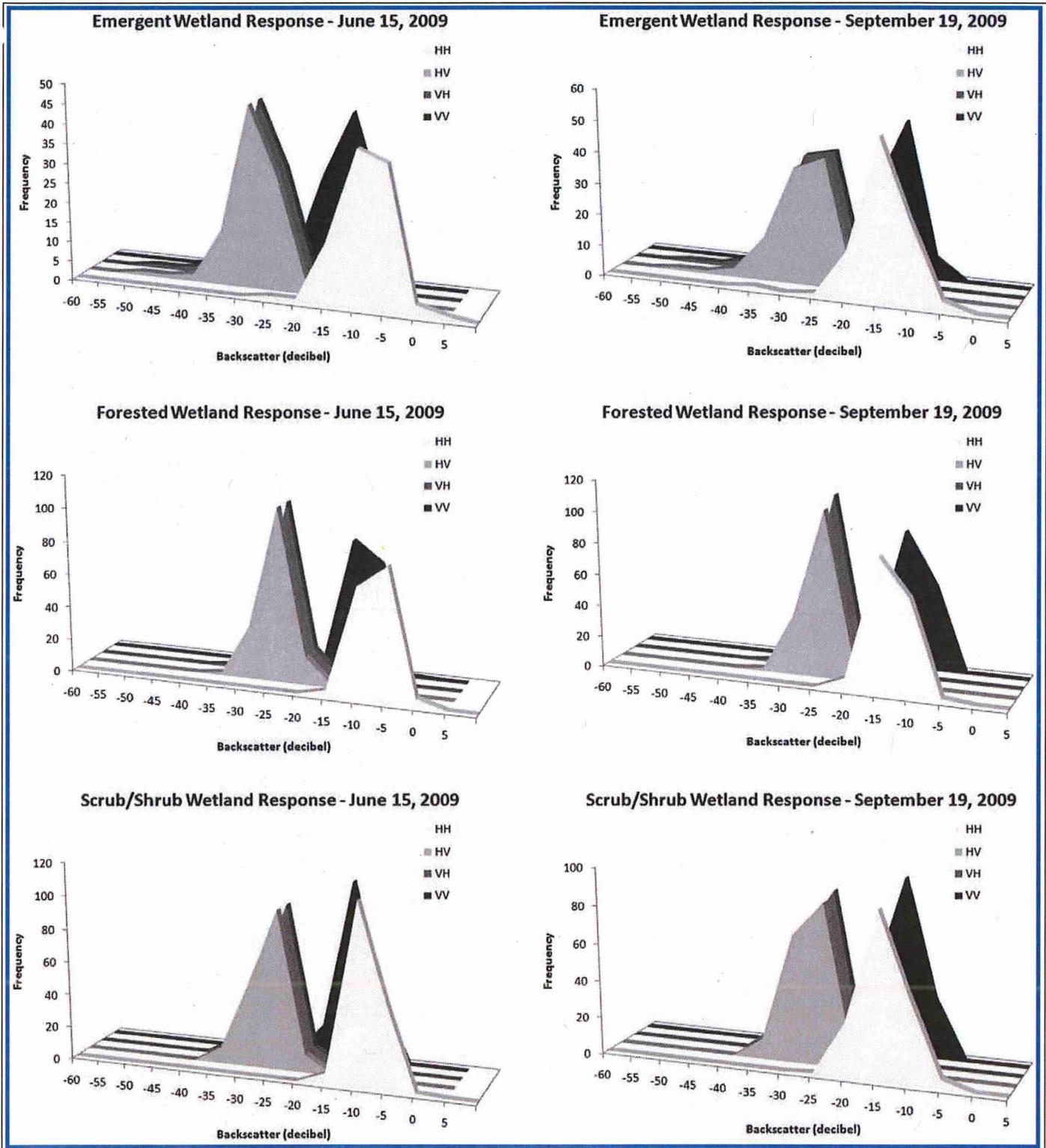


Figure 8. Spectral response signatures for each of the SAR polarizations extracted from field reference data points of emergent, forested, and scrub/shrub wetland class types and for each SAR image date: 15 June 2009 and 19 September 2009.

Entropy is defined as the randomness of scattering, alpha angle is indicative of the dominant scattering mechanism, and anisotropy is a parameter that indicates whether there are multiple scattering mechanisms occurring. Cloude and

Pottier (1997) also developed an unsupervised classification scheme based on regions of the entropy, alpha, anisotropy space. This research utilizes the parameters entropy, alpha angle, and anisotropy as separate layers in the classifier, in

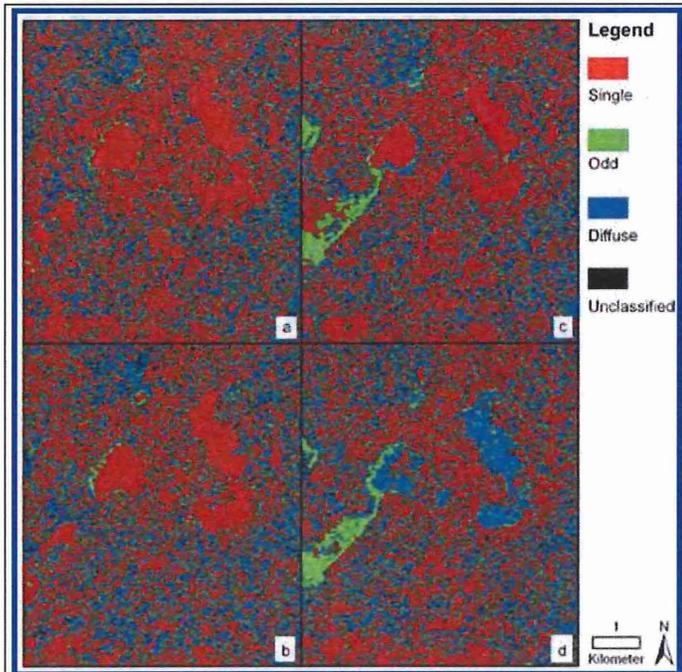


Figure 9. Subset area showing the van Zyl SAR polarimetric decomposition results for all four dates used in this research: (a) 15 June 2009, (b) 9 July 2009, (c) 26 August 2009, and (d) 19 September 2009.

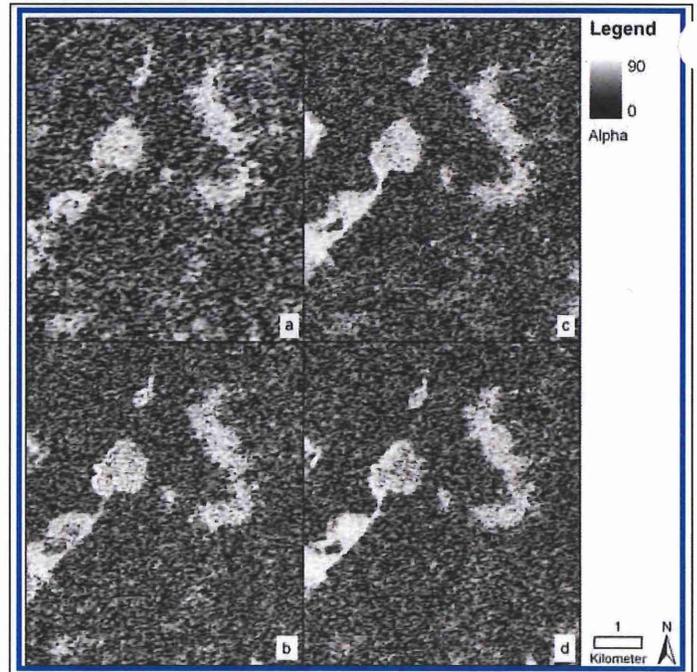


Figure 11. Subset area showing the Cloude–Pottier SAR polarimetric decomposition parameter alpha for all four dates used in this research: (a) 15 June 2009, (b) 9 July 2009, (c) 26 August 2009, and (d) 19 September 2009.

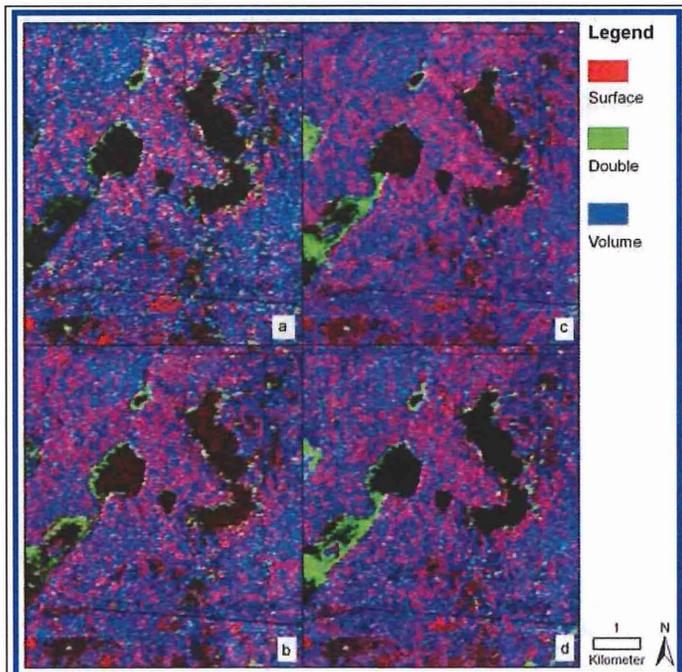


Figure 10. Subset area showing the Freeman–Durden SAR polarimetric decomposition results for all four dates used in this research: (a) 15 June 2009, (b) 9 July 2009, (c) 26 August 2009, and (d) 19 September 2009.

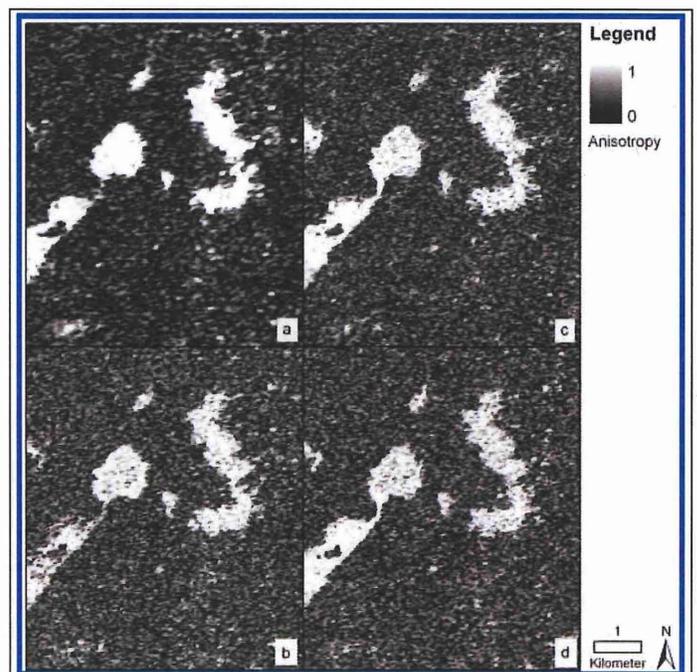


Figure 12. Subset area showing the Cloude–Pottier SAR polarimetric decomposition parameter anisotropy for all four dates used in this research: (a) 15 June 2009, (b) 9 July 2009, (c) 26 August 2009, and (d) 19 September 2009.

For personal use only.

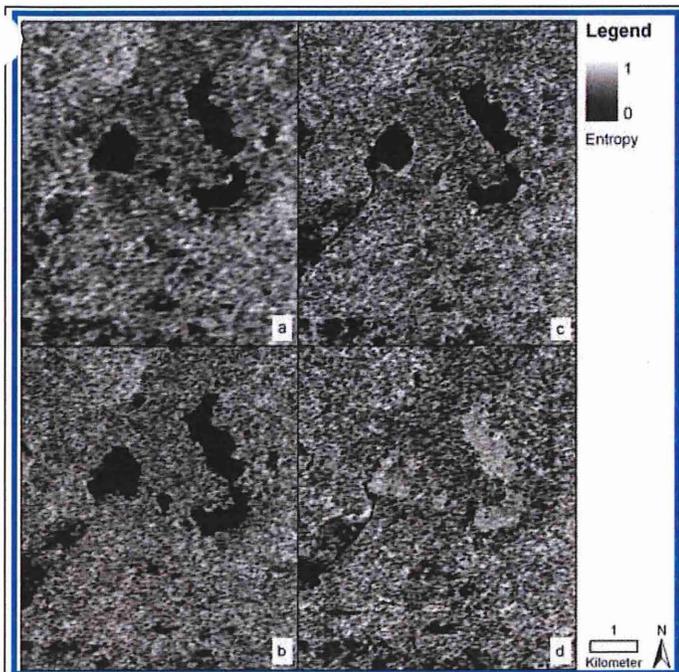


Figure 13. Subset area showing the Cloude–Pottier SAR polarimetric decomposition parameter entropy for all four dates used in this research: (a) 15 June 2009, (b) 9 July 2009, (c) 26 August 2009, and (d) 19 September 2009.

distributed tie points and orthorectified using the 2008 aerial orthophotos and NED (Figure 3). The boxcar filter, calculation of σ^0 , polarimetric decomposition processing, and orthorectification of SAR images were completed using PCI Geomatica and the image clip and spatial resolution resampling procedures were done using ERDAS Imagine software.

Accuracy assessment

To assess the accuracy of the decision tree classification results, the aforementioned independent reference dataset was utilized (25% of the field point data). Error matrices were produced for both the upland/water/wetland determinant and modified Cowardin classifications, as well as both input dataset combinations of optical/topographic and optical/topographic/SAR input data. For each classification, user's and producer's accuracies were calculated, along with errors of omission and commission, overall accuracy, and the kappa statistic (\hat{k}) (Congalton and Green 1999). A significance test of both error matrix \hat{k} values was used to compare the input dataset combinations. In addition to the above analyses, the overall accuracy of the original NWI data was assessed using the same independent reference points. A similar \hat{k} significance test was performed between the optical/topographic/SAR input dataset and the NWI for each classification scheme.

A classification tree is created using training data to determine, branch-by-branch, the best dichotomous split to reduce intraclass variability and the resulting ruleset is applied to the whole set of input data. RandomForest has the capacity to grow multiple decision trees and the end result is a classification tree, which received the best vote of confidence by cross-validation. The outputs of the random-forest classification described in this paper include: (i) a

addition to the thematic Cloude–Pottier classification product, totalling four layers of data per image date.

To finish preparing SAR data for the decision-tree classifier, the polarimetric decompositions and associated data layers, plus the backscattering coefficient for each polarization, were stacked for each SAR image date. The stacked image dates were then co-registered using 30 evenly

Table 2. Error matrices and associated accuracy results from the upland/water/wetland determinant classification using optical and topographical data only and for using optical, topographic, and SAR imagery combined.

Classified data	Reference data			Row total	User accuracy (%)	Commission error (%)
	Upland	Water	Wetland			
Optical and Topographical input data only						
Upland	121	2	43	166	73	27
Water	0	13	3	16	81	19
Wetland	30	8	94	132	71	29
Column total	151	23	140	315		
Producers accuracy (%)	80	57	67			
Omission error (%)	20	43	33			
Optical, Topographical, and SAR data						
Upland	119	6	35	160	74	26
Water	1	13	2	16	81	19
Wetland	32	4	103	139	74	26
Column total	152	23	140	315		
Producers accuracy (%)	78	57	74			
Omission error (%)	22	43	26			

For personal use only.

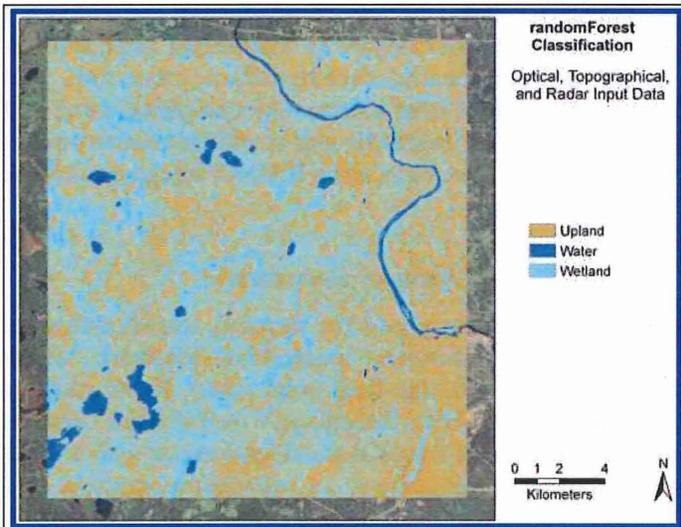


Figure 14. Upland/water/wetland determinant classification result from a randomForest decision-tree classification using a combination of optical, topographical, and SAR imagery as inputs data in the classifier.

measure of confidence per cell, created by cross-validating observed versus predicted classes; (ii) the gini index, which is used to evaluate the most significant input layers; and (iii) the mean decrease in accuracy per input layer.

The gini index corresponds to the structure of a decision tree, such that every time a split is determined by an input variable, there is a resulting decrease in the gini index for that variable (Breiman, 2001). The mean decrease in accuracy is determined by the resulting accuracy from several out-of-bag samples, in which a variable is randomly included or excluded and the resulting change in the overall accuracy for all trees is averaged (Breiman, 2001). The partial dependence of specific values of a variable is determined by comparing the error rate from an out-of-bag sample using a random selection of values to the error rate of the same out-of-bag sample using all values of that variable. The result is a graphical description, or profile, of the effect that a variable's values have on the class probability, after accounting for the effects of the other variables. The y-axis of a partial dependence plot is the predicted function and log of the fraction of votes (logits) for the classification (Breiman, 2001). In this study, the most important variables in the classification were determined by both the gini index and the mean decrease in accuracy. A selection of the top input variables were evaluated for the partial dependence of its values.

Results and discussion

Polarimetric decompositions

The results for the van Zyl polarimetric decomposition are shown in **Figure 9**. There were a few notable trends.

For the frequency of pixels classified as odd and diffuse scattering increased, but surface scattering decreased over time. The increase of odd and diffuse scattering was particularly noticeable around the water bodies in the central part of the area of interest. Looking back at the spring and summer aerial orthophotos in **Figure 2**, there was a noticeable change of emergent vegetation around the water bodies between the summer and spring images. Single scattering decreased generally over time, but stayed the same in certain areas, notably the south part of **Figure 8** where there is urban development.

The Freeman–Durden polarimetric decomposition results are shown in **Figure 10**. Beyond the trend in increased double-bounce scattering around the water bodies, similar to the findings in the van Zyl decomposition, there was a trend toward a higher fraction of volume scattering per pixel over time (indicated by the increased frequency of bluer pixels). In **Figure 2**, the areas that have a concentration of volume scattering pixels appear to be forested wetlands and the areas in that have a concentration of mixed double/volume scattering pixels (indicated by the magenta color) tend to be upland forest.

The results from the Cloude–Pottier decomposition are shown in **Figures 11–13**. The parameter alpha, indicative of the dominant scattering mechanism, appears to be fairly noisy in all four dates (**Figure 11**). However, a closer look at the map reveals that there was a slight trend toward more definition of surface features, where the central water bodies were outlined by high alpha values. Overall, June had the lowest range and mean alpha angles and August had the highest, likely due to differences in maturation and density of vegetation in these two time periods.

The result of the Cloude–Pottier decomposition parameter anisotropy, indicative of the presence of multiple scattering mechanisms, or surface roughness, is shown in **Figure 12**. High anisotropy values indicated that the scattering was predominantly singular scattering mechanisms, while low values indicated multiple scattering types. The water bodies, as expected, had anisotropy values indicative of a specular scattering mechanism, however, it was difficult to discern differences in other land cover types.

The last Cloude–Pottier parameter examined was entropy, which indicated the relative randomness of scattering on the ground, as shown in **Figure 13**. Entropy and anisotropy had an inverse relationship, where areas that had multiple scattering mechanisms (low anisotropy values) had a high degree of randomness (high entropy values). Looking at the water bodies in the central part of the figure, it makes sense that the entropy values were low while the anisotropy values were high. However, later in September (**Figure 13b**), the entropy values increased in the same areas that saw an increase in the van Zyl diffuse scattering mechanisms. This was likely due to the wind causing small changes of the water surface and not likely indicative of a change in vegetation.

For personal use only.

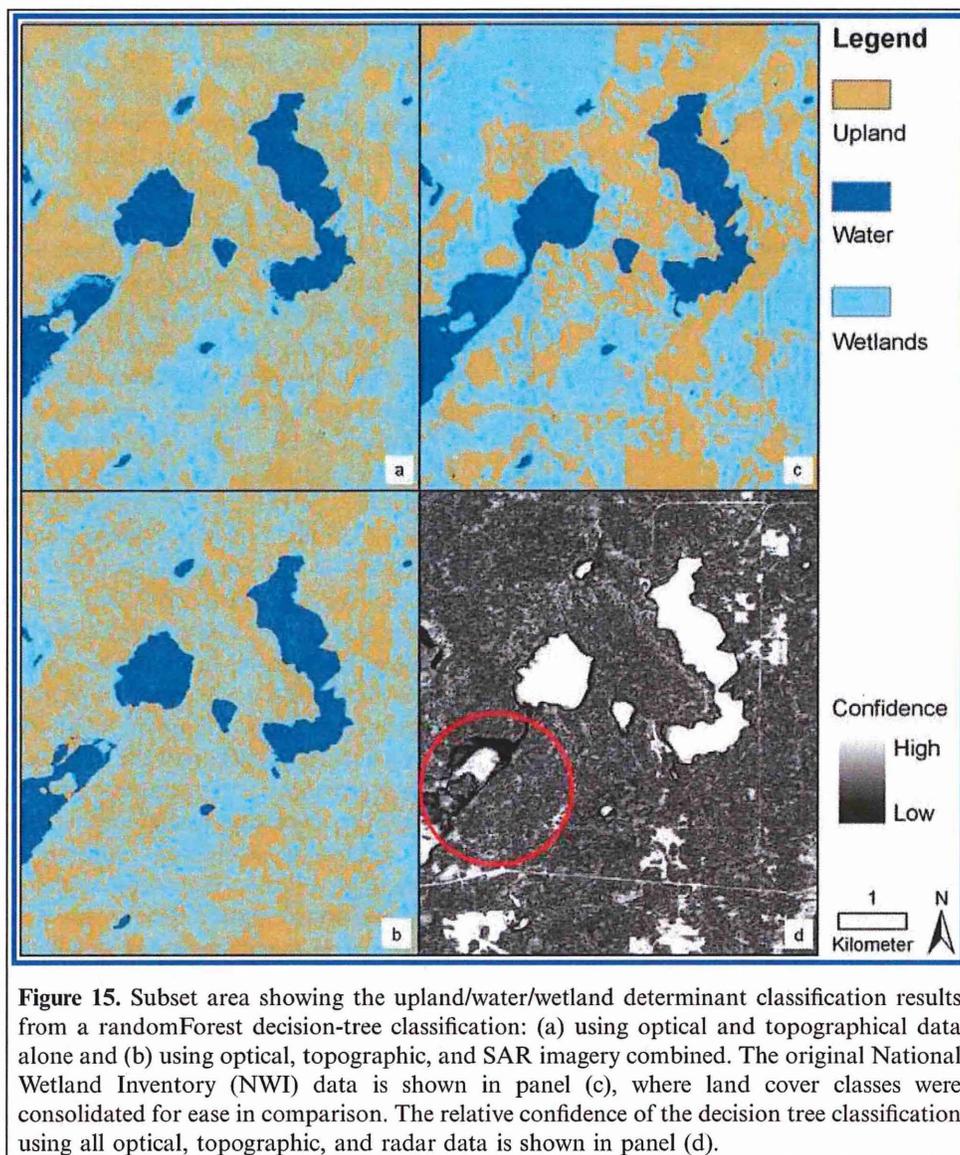


Figure 15. Subset area showing the upland/water/wetland determinant classification results from a randomForest decision-tree classification: (a) using optical and topographical data alone and (b) using optical, topographic, and SAR imagery combined. The original National Wetland Inventory (NWI) data is shown in panel (c), where land cover classes were consolidated for ease in comparison. The relative confidence of the decision tree classification using all optical, topographic, and radar data is shown in panel (d).

Upland/water/wetland determinant classification

The first classification analysis presented is the result from the upland/water/wetland determinant classification. The error matrix in **Table 2** illustrates that without the inclusion of SAR data, the decision-tree classifier confused upland areas with wetland areas 27% of the time (commission error) and by including all available SAR data there was a slight improvement to 26% commission error. The largest improvement in terms of differentiating between upland/water/wetland was in the omission error of wetlands, meaning SAR data helped to ensure that more reference points were correctly classified as wetlands. The output classification map for the entire study area can be seen in **Figure 14**.

A subset area was chosen to illustrate an area with diverse land cover and seasonality (**Figure 2**). **Figure 15** shows that differentiation between the upland/water/wetland classes

was fairly similar between the two decision-tree tests (optical/topographical only versus optical/topographic/SAR included); however, both classifications were quite different from that of the original NWI. As it is based on older imagery, the NWI may not reflect land use change of wetlands being converted to other upland classes. As a result, the NWI appeared to significantly overestimate wetland area compared with this classification based on current imagery. It is also important to point out areas with relatively low confidence in the classification result. For example, the confidence of the shoreline of a southwest lake (circled in **Figure 15d**) was particularly low. This was likely due to the apparent seasonality observed in the aerial photos of **Figure 2** and a possible lack of temporal coverage to accommodate the seasonality.

The mean decrease in accuracy and gini index plots from randomForest were assessed in **Figure 16**, which shows the most effective input datasets for the upland/water/

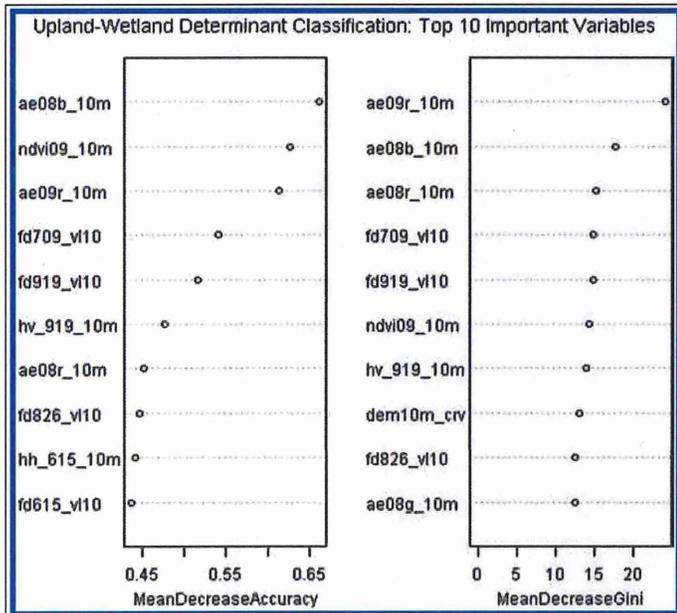


Figure 16. Mean decrease accuracy and gini index plots for the upland/water/wetland determinant classification using optical, topographic, and SAR imagery combined.

and the Freeman–Durden volume scattering parameter from the mid-summer 7 July 2009 image (“fd709_v110”). It was interesting to note the relative values for each dataset that were most sensitive to the accuracy of the classification (Figure 17). These partial dependence plots indicated the values with higher probability of significance for improving the accuracy of the classification. For example, the DN values of the red band that are responsible for improving the classification accuracy (around 45–65) were similar for both the early leaf-onset and leaf-on mid-summer aerial photos. Conversely, the DN values less than 75 or greater than 150 of the blue band in the leaf-on aerial photo were the most probable in improving the accuracy of the classification.

Though SAR input data were not among the most important variables for decision-tree classification, several layers were in the top ten, including: the volume scattering channel of the Freeman–Durden polarimetric decomposition from all four image dates utilized in this study, the HH channel from 15 June 2009, and the HV channel from 19 September 2009. These results illustrated that including relative backscatter as well as polarimetric information about scattering mechanisms typically observed by vegetative canopies helped improve upland/water/wetland classification accuracy. Although the Freeman–Durden parameters were important in the accuracy of the classification, it was difficult to assess the partial dependence of specific values of one scattering mechanism without knowing the relative percentage of the other two scattering mechanisms. The least important input datasets were found to be the thematic polarimetric decompositions of Cloude–Pottier and, most particularly, van Zyl. These findings were likely due to the limiting and unrealistic nature of assigning a single scattering mechanism to each pixel on the ground.

Modified Cowardin land cover classification

Table 3 shows the error matrix for the modified Cowardin classification. With optical and topographic data alone, the decision-tree classifier confused scrub/shrub areas 51% of the time (commission error) and mainly misclassified these areas as upland. Unfortunately, there was no improvement in commission error of scrub/shrub wetlands by including SAR data, mainly due to the additional confusion with forested wetlands. The addition of SAR improved the accuracy of classifying water but there was little difference in the omission or commission errors of the upland and wetland classes. The output of this classification for the entire study area is shown in Figure 18.

In the same subset area as discussed previously, Figure 19 shows how the modified Cowardin classification results are again visually similar between the two decision-tree tests (optical/topographical only versus optical/topographic/SAR included), but vary greatly from that of the original NWI. It was clear that the original NWI estimated a much higher coverage of forested wetlands and little scrub/shrub in comparison. Pointing out the same

wetland determinant classification. Among these, the most important input variables include: the blue and red bands of 2008 leaf-on aerial orthophoto (“ae08b_10m” and “ae08r_10m”), the red band and NDVI of 2009 early leaf-onset aerial orthophoto (“ae09r_10m” and “ndvi09_10m”),

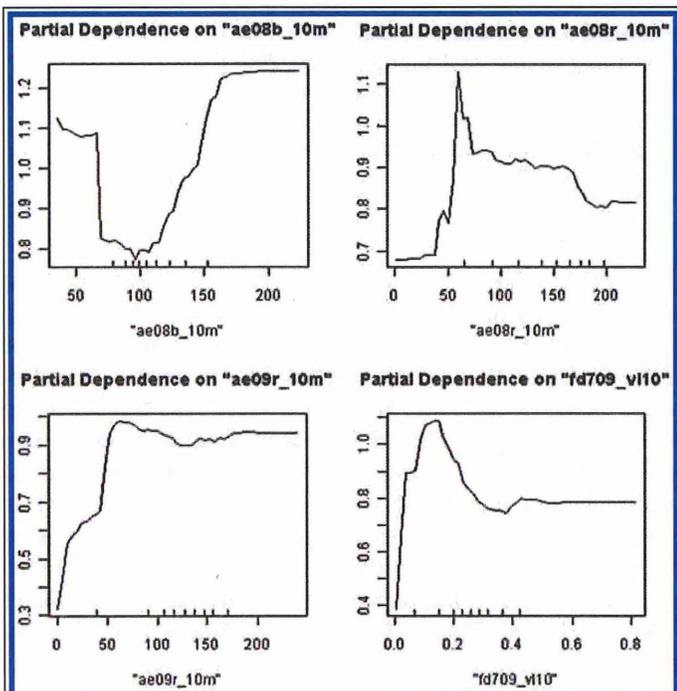


Figure 17. Value partial dependence plots for a selection of the most important input variables for the upland/water/wetland determinant classification using optical, topographic, and SAR imagery combined.

For personal use only.

Table 3. Error matrices and associated accuracy results from the modified Cowardin land cover classification using optical and topographical data only and for using optical, topographic, and SAR imagery combined.

Classified data	Reference data					Row total	User accuracy (%)	Commission error (%)
	Water	Emergent wetlands	Forested wetlands	Scrub/shrub wetlands	Upland			
Optical and topographical input data only								
Water	13	3	0	0	1	17	76	24
Emergent wetlands	6	15	2	3	3	29	52	48
Forested wetlands	0	0	17	8	8	33	52	48
Scrub/shrub Wetlands	0	6	6	31	20	63	49	51
Upland	4	13	23	13	120	173	69	31
Column Total	23	37	48	55	152	315		
Producers accuracy (%)	57	41	35	56	79			
Omission error (%)	43	59	65	44	21			
Optical, Topographical, and SAR Data Included								
Water	16	2	0	0	1	19	84	16
Emergent wetlands	3	15	1	3	5	27	56	44
Forested wetlands	0	0	19	10	12	41	46	53
Scrub/shrub wetlands	1	11	9	31	15	67	46	54
Upland	3	9	19	11	119	161	74	26
Column total	23	37	48	55	152	315		
Producers accuracy (%)	70	41	40	56	78			
Omission error (%)	30	59	60	44	22			

southwestern lake as described previously, the confidence in classification of wetland type increased (Figure 19d) and there were few areas within this subset area that had very low confidence.

The mean decrease in accuracy and gini index plots from randomForest for the modified Cowardin classification are shown in Figure 20. The most important input datasets for this classification include: the blue and red bands of the 2008

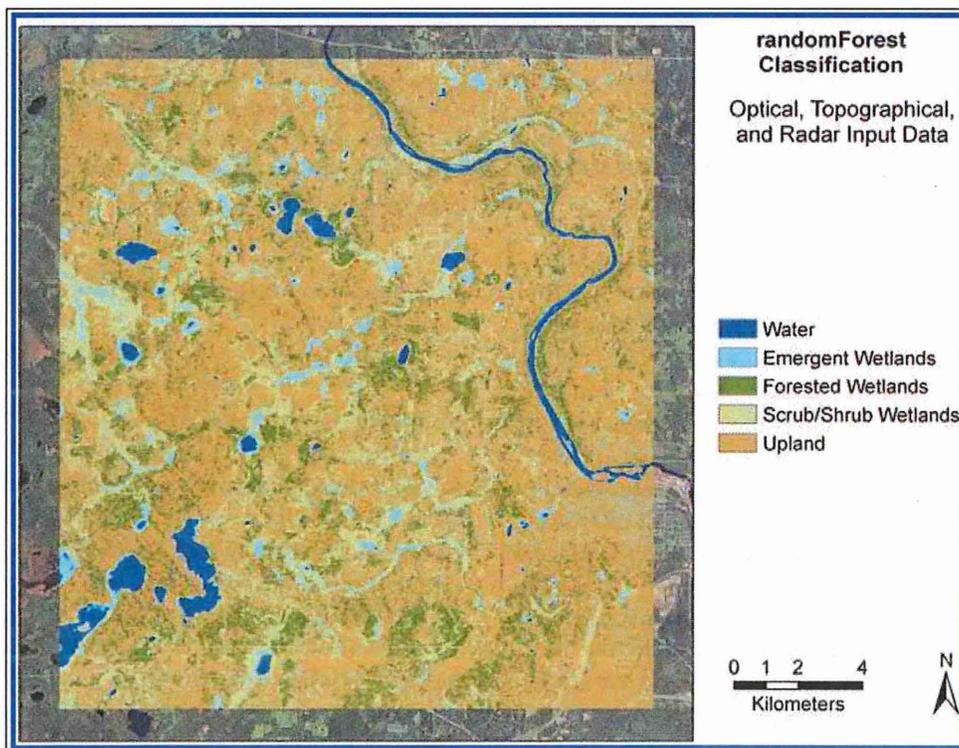


Figure 18. A modified Cowardin land cover classification result from a randomForest decision tree classification using a combination of optical, topographical, and SAR imagery as inputs data in the classifier.

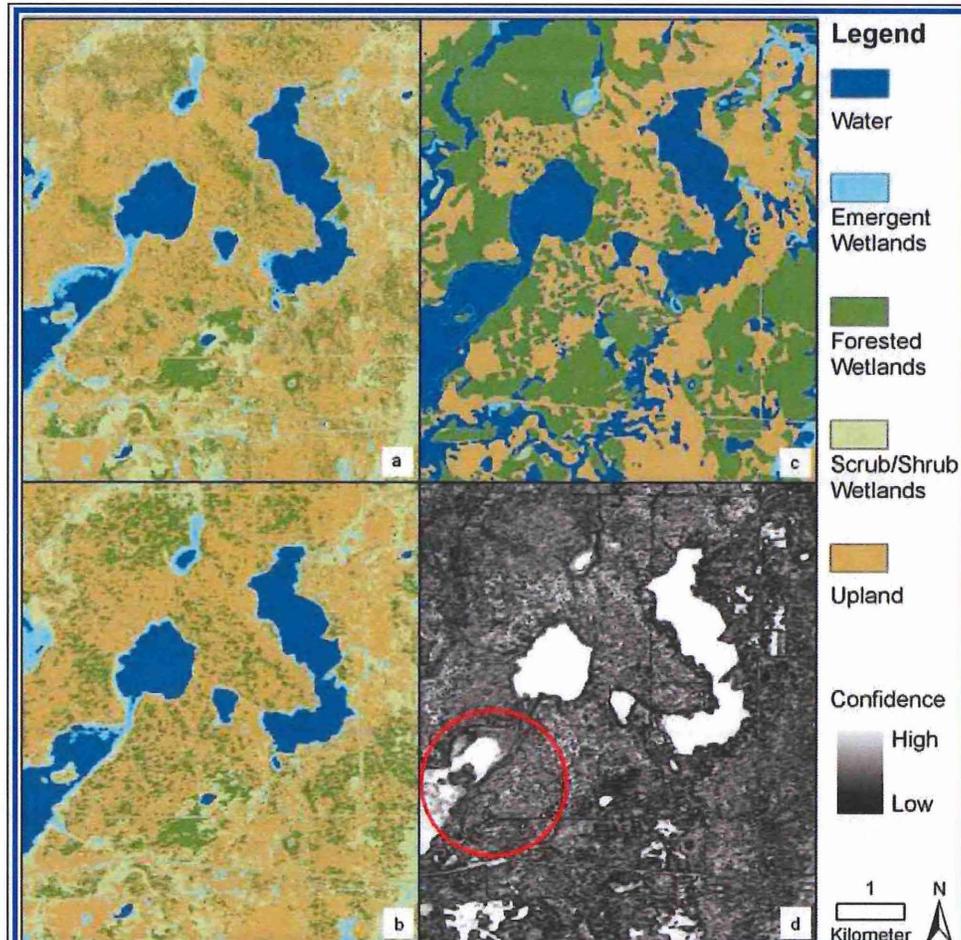


Figure 19. Subset area showing the upland/water/wetland determinant classification results from a randomForest decision tree classification: (a) using optical and topographical data alone and (b) using optical, topographic, and radar imagery combined. The original National Wetland Inventory (NWI) data is shown in panel (c), where land cover classes were consolidated for ease in comparison. The relative confidence of the decision tree classification using all optical, topographic, and radar data is shown in panel (d).

leaf-on aerial orthophoto (“ae08b_10m” and “ae08r_10m”), the red band and NDVI of 2009 early leaf-onset aerial orthophoto (“ae09r_10m” and “ndvi09_10m”), and elevation (“dem_10m”). Except for the apparent dependence on elevation, the important variables for the modified Cowardin classification were not very different from the variables important for upland/water/wetland determinant classification. However, it was most interesting that the sensitive DN values for each of these datasets were very different (**Figure 21**). Here, DN values of the leaf-on blue band that were least important for the upland/water/wetland determinant classification were among the most important values for the Cowardin classification. A similar pattern was found with the important values in the early leaf-onset red band, where the DN values below 45 were much more important in the Cowardin classification than for the upland/water/wetland determinant classification. These findings strengthened the case for the inclusion of

seasonal aerial orthophotos for improving wetland mapping accuracy.

Similar to the previous findings, the volume scattering channel of the Freeman–Durden polarimetric decomposition from most image dates and the HV channel from June were found to be important for classifying wetland types. Again, these results illustrated that including relative backscatter as well as polarimetric information can improve wetland mapping accuracy over having the broad thematic definitions of surface scattering mechanisms.

Table 4 shows a comparison of all error matrices for each classification scheme. The addition of SAR data in the upland/water/wetland determinant classification improved the accuracy by 3% over having optical and topographical data alone and by 5% over the original NWI. Each of the upland/water/wetland classification error matrices were found to have statistically significant z statistics, but the differences between the optical/topographic and optical/

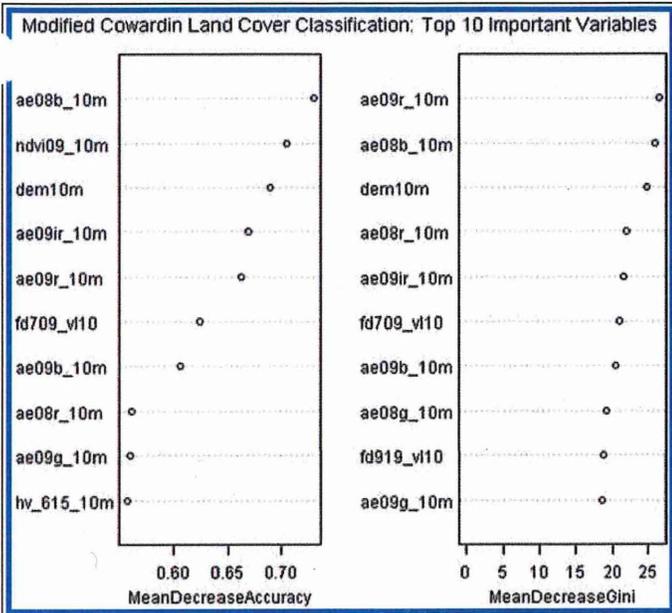


Figure 20. Mean decrease accuracy and gini index plots for the modified Cowardin land cover classification using optical, topographic, and SAR imagery combined.

original NWI and optical/topographic/SAR matrices showed that this method significantly improved the classification wetland types, increasing the accuracy by 14%.

Summary and conclusion

The research presented here showed that the integration of multitemporal, multisensor, and multifrequency remotely sensed data improved the accuracy of a decision-tree classification of wetlands in a forested region of northern Minnesota. Forested wetlands are typically very challenging to map, due to the obstruction of tree canopy cover. The incorporation of radar backscatter and polarimetric data was shown to improve the commission error of forested wetlands and improved the overall accuracy of all wetland types. The original NWI has many disadvantages and this study offered a method to improve the classification accuracy of wetlands using a robust, free, decision-tree algorithm.

The potential to apply the methodology presented in this paper over another study site is limited mainly by input imagery and training data availability. Currently, SAR data is not freely available and can be expensive if fully polarimetric imagery is desired over large geographic areas. However, certain landscapes may not benefit from the addition of SAR data, particularly if there is very little tree canopy or if the weather is always clear.

The prevailing advantage of longer wavelength radar signals is thought to be their circumvention and deeper penetration of dense canopy cover. Though SAR data improved the accuracy of differentiating between upland and wetland areas, the performance of the decision-tree classifier was not significantly different than without SAR for this study site. Changes in surface structure directly affected backscatter brightness and classified scattering mechanisms. However, the temporal variability in these land classes are apparently not significant enough that SAR data contributed considerable improvement to the accuracy of the land cover classifications shown in this paper. Further research is therefore needed in other SAR sensor platforms with longer wavelengths, such as the Advanced Land Observing Satellite Phased Array type L-band Synthetic Aperture Radar (23 cm wavelength) data and in alternative optical platforms with more spectral bands available, such as the Landsat Thematic Mapper. Incorporation of optical, topographic, C-band, and L-band data may increase the accuracy of classifying forested wetlands.

The results of this study included a wetland classification map and the relative confidence of each pixel. The methods presented here provide a valuable tool for automated mapping of wetland areas and provide an effective aid for facilitating the manual mapping of more challenging wetland class types using aerial photos and a human interpreter.

topographic/SAR matrices and between the optical/topographic/SAR and NWI matrices were not statistically significant. For the modified Cowardin classification scheme, the addition of SAR data only improved the accuracy 1% over having optical and topographical data alone. However, a comparison of the *z* statistics from the

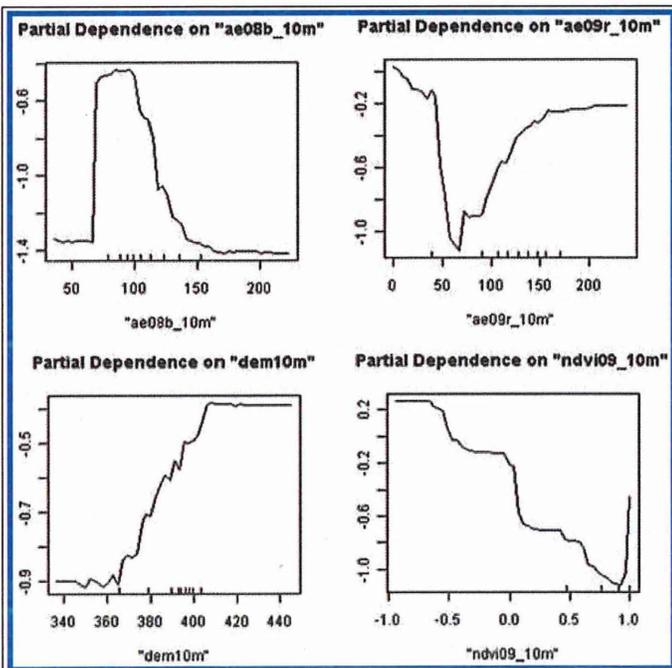


Figure 21. Value partial dependence plots for a selection of the most important input variables for the modified Cowardin land cover classification using optical, topographic, and SAR imagery combined.

For personal use only.

Table 4. Summary of the independent accuracy assessment results from both decision-tree classification schemes (upland/water/wetland determinant and modified Cowardin class) and for both input dataset combinations of optical/topographic input data only and optical/topographic/SAR input data. Using the same independent reference data points, an accuracy assessment was also done on the original National Wetland Inventory (NWI).

Land cover classification	Total accuracy	Kappa-hat statistic	Kappa-hat variance	Z statistic (significant)	Compare Z statistic
Upland–wetland determinant					
Optical and topographical input data only	72	0.51	0.0021	11.1*	0.6
Optical, topographical, and SAR data included	75	0.54	0.0020	12.2*	
Original National Wetland Inventory (NWI)	70	0.47	0.0021	10.3*	1.1
Modified Cowardin class					
Optical and topographical input data only	62	0.44	0.0015	11.1*	0.53
Optical, topographical, and SAR data included	63	0.47	0.0015	12.2*	
Original National Wetland Inventory (NWI)	49	0.31	0.0010	9.70*	3.2*

Acknowledgements

This research was carried out as part of a Canadian Space Agency's Science and Operational Applications Research (SOAR) Program project and was supported by the Minnesota Environment and Natural Resources Trust Fund and the Minnesota Department of Natural Resources. Special thanks are due to Brian Huberty of the U.S. Fish and Wildlife Service and Steve Kloiber of the Minnesota Department of Natural Resources, who initiated collaboration with the Canada Centre for Remote Sensing and secured project funding. The authors gratefully acknowledge the helpful comments of this paper's two reviewers.

References

- Acharya, G., and Barbier, E. 2000. Valuing groundwater recharge through agricultural production in the hadejia-nguru wetlands in northern Nigeria. *Agricultural Economics*, Vol. 22, No. 3, pp. 247–259. doi: 10.1111/j.1574-0862.2000.tb00073.x.
- Baghdadi, N., Bernier, M., Gauthier, R., and Neeson, I. 2001. Evaluation of C-band SAR data for wetlands mapping. *International Journal of Remote Sensing*, Vol. 22, No. 1, pp. 71–88. doi: 10.1080/014311601750038857.
- Ban, Y., Hu, H., and Rangel, I. 2010. Fusion of Quickbird MS and RADARSAT SAR data for urban land-cover mapping: Object-based and knowledge-based approach. *International Journal of Remote Sensing*, Vol. 31, No. 6, pp. 1391–1410. doi: 10.1080/01431160903475415.
- Breiman, L. 2001. Random forests. *Machine Learning*, Vol. 45, No. 1, pp. 5–32. doi: 10.1023/A:1010933404324.
- Brisco B., Touzi R., van der Sanden, J., Charbonneau, F., Pultz, T., and D'Iorio, M. 2008. Water resource applications with RADARSAT-2: A preview. *International Journal of Digital Earth*, Vol. 1, No. 1, pp. 130–147. doi: 10.1080/17538940701782577.
- Brooks, R., Wardrop, D., and Cole, C. 2006. Inventorying and monitoring wetland condition and restoration potential on a watershed basis with examples from Spring Creek Watershed, Pennsylvania, USA. *Environmental Management*, Vol. 38, pp. 673–687. doi: 10.1007/s00267-004-0389-y.
- Bwagony, J., Hansen, M., Roy, D., Grandi, G., and Justice, C. 2010. Wetland mapping in the Congo basin using optical and radar remotely sensed data and derived topographical indices. *Remote Sensing of Environment*, Vol. 114, No. 1, pp. 73–86. doi: 10.1016/j.rse.2009.08.004.
- Campbell, J. 2007. *Introduction to Remote Sensing*. Guilford Press, New York, p. 466.
- Castañeda, C., and Ducrot, D. 2009. Land cover mapping of wetland areas in an agricultural landscape using SAR and Landsat imagery. *Journal of Environmental Management*, Vol. 90, No. 7, pp. 2270–2277. doi: 10.1016/j.jenvman.2007.06.030.
- Cloude, S., and Pottier, E. 1997. An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 35, No. 1, pp. 68–78. doi: 10.1109/36.551935.
- Congalton, R., and Green, K. 1999. *Assessing the accuracy of remotely sensed data: principles and practices*. Lewis Publishers, Boca Raton, FL.
- Cowardin L., Carter, F., and LaRoe, E. 1979. *Classification of wetlands and deepwater habitats of the United States*. U.S. Department of the Interior, Fish and Wildlife Service, Northern Prairie Publication, Jamestown, ND. Report No. 0421.
- Dahl, T. 1990. *Wetlands losses in the United States 1780's to 1980's*. U.S. Department of the Interior, Fish and Wildlife Service, Office of Biological Services, Washington D.C.
- Deschamps, A., Greenlee, D., Pultz, T., and Saper, R. 2002. Geospatial data integration for applications in flood prediction and management in the Red River Basin. *IEEE International, Geoscience and Remote Sensing*, Vol. 6, pp. 3338–3340.
- Frappart, F., Seyler, F., Martinez, J., León, J., and Cazenave, A. 2005. Floodplain water storage in the Negro River Basin estimated from microwave remote sensing of inundation area and water levels. *Remote Sensing of Environment*, Vol. 99, No. 4, pp. 387–399. doi: 10.1016/j.rse.2005.08.016.
- Freeman, A., and Durden, S. 1998. A three-component scattering model for polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 36, No. 3, pp. 963–973. doi: 10.1109/36.673687.
- Hearne, R. 2007. Evolving water management institutions in the Red River Basin. *Environmental Management*, Vol. 40, No. 6, 842 p. doi: 10.1007/s00267-007-9026-x.

- Henderson, F., and Lewis, A. 2008. Radar detection of wetland ecosystems: A review. *International Journal of Remote Sensing*, Vol. 29, No. 20, pp. 5809–5835. doi: 10.1080/01431160801958405.
- Hess, L., Melack, J., Novo, E., Barbosa, C., and Gastil, M. 2003. Dual-season mapping of wetland inundation and vegetation for the central Amazon Basin. *Remote Sensing of Environment*, Vol. 87, No. 4, pp. 404–428. doi: 10.1016/j.rse.2003.04.001.
- Hogg, A., and Todd, K. 2007. Automated discrimination of upland and wetland using terrain derivatives. *Canadian Journal of Remote Sensing*, Vol. 33, pp. S68–S83. doi: 10.5589/m07-049.
- Kaya, S. 2010. Personal communication, June 30, Ottawa, ON. Environment Canada, Canada Center for Remote Sensing.
- Lu, Z., and Kwoun, O. 2008. RADARSAT-1 and ERS InSAR analysis over southeastern coastal Louisiana: Implications for mapping water-level changes beneath swamp forests. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 46, No. 8, pp. 2167–2184. doi: 10.1109/TGRS.2008.917271.
- Minnesota Department of Administration (AdminMN), Office of Geographic and Demographic Analysis State Demographic Center. 2011. 2010 Census: Minnesota City Profiles. Available from <http://www.demography.state.mn.us/CityProfiles2010/index.html> [Accessed 22 November 2011].
- Minnesota Department of Natural Resources (MN DNR). 2010 Wetlands Status and Trends. Available from http://www.dnr.state.mn.us/eco/wetlands/wstm_prog.html [Accessed 20 August 2010].
- Minnesota Department of Natural Resources (MN DNR). Natural History – Minnesota's Geology. Available from <http://www.dnr.state.mn.us/snas/naturallhistory.html> [Accessed 31 March 2011].
- Minnesota Geospatial Information Office (MnGeo), Geographic Data Clearinghouse. Minnesota Aerial Photography: 2008. Available from http://www.mngeo.state.mn.us/chouse/airphoto/program_details.html [Accessed 31 March 2011].
- Minnesota Geospatial Information Office (MnGeo), Geographic Data Clearinghouse. Spring Aerial Imagery Program for Minnesota: 2009-2015. Available from <http://www.mngeo.state.mn.us/chouse/airphoto/spring2009-2015.html> [Accessed 31 March 2011].
- Mitsch, W., and Gosselink, J. 2000. The value of wetlands: Importance of scale and landscape setting. *Ecological Economics*, Vol. 35, No. 1, pp. 25–33. doi: 10.1016/S0921-8009(00)00165-8.
- Moore, I., Grayson, R., and Landson, A. 1991. Digital terrain modeling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, Vol. 5, pp. 3–30. doi: 10.1002/hyp.3360050103.
- Ozesmi, S., and Bauer, M. 2002. Satellite remote sensing of wetlands. *Wetlands Ecology and Management*, Vol. 10, No. 5, pp. 381–402. doi: 10.1023/A:1020908432489.
- Parmuchi, M., Karszenbaum, H., and Kandus, P. 2002. Mapping wetlands using multi-temporal RADARSAT-1 data and a decision-based classifier. *Canadian Journal of Remote Sensing*, Vol. 28, No. 2, pp. 175–186. doi: 10.5589/m02-014.
- Ramsey, E. 1995. Monitoring flooding in coastal wetlands by using radar imagery and ground-based measurements. *International Journal of Remote Sensing*, Vol. 16, No. 13, 2495 p. doi: 10.1080/01431169508954571.
- Ramsey, E., Rangoonwala, A., Middleton, B., and Lu, Z. 2009. Satellite optical and radar data used to track wetland forest impact and short-term recovery from Hurricane Katrina. *Wetlands*, Vol. 29, No. 1, pp. 66–79. doi: 10.1672/08-103.1.
- Ruefenacht, B., Liknes, G., Lister, A., Fisk, H., and Wendt, D. 2008. Evaluation of Open Source Data Mining Software Packages. In *Forest Inventory and Analysis (FIA) Symposium 2008*, October 21–23, 2008, Park City, UT, USA. Edited by McWilliams, W. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO. Addendum.
- Slotton K.C., Crawford M.M., and Chang L.D., 2008. Modeling temporal variations in multipolarized radar scattering from intertidal coastal wetlands. *ISPRS Journal of Photogrammetry and Remote Sensing*. Vol. 63, pp. 559–577. doi: 10.1016/j.isprsjprs.2008.07.003.
- Touzi, R. 2006. Wetland characterization using polarimetric RADARSAT-2 capability. 2006. In *IEEE international Conference on Geoscience and Remote Sensing Symposium*, July 31-August 4, 2006, Denver, CO, U.S.A. 1639 p.
- Touzi, R., Deschamps, A., and Rother, G. 2009. Phase of target scattering for wetland characterization using polarimetric C-band SAR. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 47, No. 9, pp. 3241–3261. doi: 10.1109/TGRS.2009.2018626.
- Townsend, P. 2001. Mapping seasonal flooding in forested wetlands using multi-temporal radarsat SAR. *Photogrammetric Engineering & Remote Sensing*, Vol. 67, No. 7, pp. 857–864.
- Töyrä, J., Pietroniro, A., Martz, L., and Prowse, T. 2002. A multi-sensor approach to wetland flood monitoring. *Hydrological Processes*, Vol. 16, No. 8, pp. 1569–1581. doi: 10.1002/hyp.1021.
- United States Department of Agriculture (USDA) Geospatial Data Gateway. Available from <http://datagateway.nrcs.usda.gov/> [Accessed 31 March 2011].
- United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS). Cropland Data Layer. Available from <http://www.nass.usda.gov/research/Cropland/SARS1a.htm> [Accessed 31 March 2011].
- United States Fish and Wildlife Service. National wetlands inventory (NWI). Available from <http://www.fws.gov/wetlands/> [Accessed 12 October 2009].
- United States Geological Survey. Accuracy of the National Elevation Dataset (NED). Available from <http://ned.usgs.gov/Ned/accuracy.asp> [Accessed 31 March 2011].
- van der Kamp, G., and Hayashi, M. 1998. *The groundwater recharge function of small wetlands in the semi-arid northern prairies*. Center for Great Plains Studies, Lincoln, NE. Report No. 8, pp. 39–56.
- van Zyl, J. 1998. Unsupervised classification of scattering behavior using radar polarimetry data. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 27, No. 1, pp. 36–45. doi: 10.1109/36.20273.
- Vymazal, J. 2005. Constructed wetlands for wastewater treatment. *Ecological Engineering*, Vol. 25, No. 5, pp. 475–477. doi: 10.1016/j.ecoleng.2005.07.002.

Article

Influence of Multi-Source and Multi-Temporal Remotely Sensed and Ancillary Data on the Accuracy of Random Forest Classification of Wetlands in Northern Minnesota

Jennifer M. Corcoran ^{1,*}, Joseph F. Knight ¹ and Alisa L. Gallant ²

¹ Department of Forest Resources, University of Minnesota, 1530 Cleveland Ave. N, St. Paul, MN 55108, USA; E-Mail: jknight@umn.edu

² Earth Resources Observation and Science Center, US Geological Survey, Sioux Falls, SD 57198, USA; E-Mail: gallant@usgs.gov

* Authors to whom correspondence should be addressed; E-Mail: murph636@umn.edu; Tel.: +1-612-624-3459; Fax: +1-612-625-5212.

Received: 5 April 2013; in revised form: 20 June 2013 / Accepted: 20 June 2013 /

Published: 4 July 2013

Abstract: Wetland mapping at the landscape scale using remotely sensed data requires both affordable data and an efficient accurate classification method. Random forest classification offers several advantages over traditional land cover classification techniques, including a bootstrapping technique to generate robust estimations of outliers in the training data, as well as the capability of measuring classification confidence. Though the random forest classifier can generate complex decision trees with a multitude of input data and still not run a high risk of over fitting, there is a great need to reduce computational and operational costs by including only key input data sets without sacrificing a significant level of accuracy. Our main questions for this study site in Northern Minnesota were: (1) how does classification accuracy and confidence of mapping wetlands compare using different remote sensing platforms and sets of input data; (2) what are the key input variables for accurate differentiation of upland, water, and wetlands, including wetland type; and (3) which datasets and seasonal imagery yield the best accuracy for wetland classification. Our results show the key input variables include terrain (elevation and curvature) and soils descriptors (hydric), along with an assortment of remotely sensed data collected in the spring (satellite visible, near infrared, and thermal bands; satellite normalized vegetation index and Tasseled Cap greenness and wetness; and horizontal-horizontal (HH) and horizontal-vertical (HV) polarization using L-band satellite radar). We undertook this exploratory analysis to inform decisions by natural resource

managers charged with monitoring wetland ecosystems and to aid in designing a system for consistent operational mapping of wetlands across landscapes similar to those found in Northern Minnesota.

Keywords: wetland; classification; data integration; decision tree; random forest

1. Introduction

Wetlands provide many ecosystem services such as filtering polluted water [1], mitigating flood damage [2–4], recharging groundwater storage [5,6], and providing habitat for diverse flora and fauna [7–9]. Wetland quality and quantity are particularly important in light of the increasing impacts of climate change, a growing human population, and changing land cover and land use practices [10,11]. It is therefore essential that wetlands are managed appropriately and monitored frequently.

The US Army Corps of Engineers defines wetlands as: “areas that are inundated or saturated by surface or ground water at a frequency and duration sufficient to support, and that under normal circumstances do support, a prevalence of vegetation typically adapted for life in saturated soil conditions” [12]. The Corps identifies potential wetland areas using three broad categories: soils, vegetation, and hydrology, where the classification is specifically based on geological substrate (soil type, drainage), the presence and type of hydrophytic vegetation, and topographic features that influence the hydrological movement and storage of water.

Characteristics of wetland structure and position are not the only influential factors on the permanence and duration of a wetland’s capacity to store water. Regional and local climate conditions are the main driving forces behind a wetland’s hydroperiod. Hydroperiod can be defined as the seasonal pattern of water level, duration and frequency in a wetland, akin to a “hydrologic signature”. The hydroperiod of a wetland has been described by Wissinger [13] as the single most important aspect of the biodiversity within a wetland habitat, because the duration between dry and wet periods directly influences complex biological interactions and communities. The phenology of a wetland has a major influence on its classification and changes in the hydroperiod over time can thus alter a wetland’s classification.

Accurate landscape-scale wetland maps are important for stakeholders that represent many different interests in wetland ecosystems. Accurate wetland maps are needed to: better respond to and prepare for natural disasters and invasive species mediation [14,15], conserve and restore wetland areas following policy and regulation changes [16,17], address water quality and quantity concerns [18,19], and better understand the linkages and seasonality of these ecosystems to biodiversity and other natural resources [20,21]. However, many existing wetland maps are out of date and efforts for updating them tend to happen over small geographic extents or at intervals too infrequent for appropriate environmental mitigation [22]. Furthermore, traditional wetland mapping methods often rely on optical imagery and manual photo interpretation or classification using single date imagery. These maps typically under-represent ephemeral and forested wetlands, due to their possible absence during time of data acquisition and because of obscuration by vegetative canopy [18]. Even if the temporal

coverage is appropriate, optical imagery alone may not reveal wetlands obscured by clouds or haze, or a dense vegetated canopy.

The integration of multi-source (multi-platform and multi-frequency) and multi-temporal remotely sensed data can provide information for mapping wetlands in addition to the use of single date optical imagery traditionally used for wetland classification. Surface features, such as extent of inundation, vegetation structure, and likelihood of wetlands can be better resolved with the addition of longer wavelength radiometric responses, topographic derivatives [23], and ancillary data about the geological substrate [24,25]. Long-wave radar signals, such as C-band (5.6 cm) or L-band (23 cm), have been found to improve land cover classification accuracy because these wavelengths have deeper canopy penetration and are sensitive to soil moisture and inundation [26–28]. These active sensors are not as sensitive to atmospheric effects, penetrate clouds, and are operational at night, thereby increasing the temporal coverage of wetland mapping. Research has shown that data from multiple sources and over multiple seasons capture greater variation in hydroperiod and vegetative condition and thus have the potential to increase both classification accuracy and confidence [29–31].

Given the wealth of remotely sensed and ancillary data, a robust wetland classification method applied to large geographic areas needs to be computationally fast, require no assumptions about data distribution, handle nonlinearity in relations between input variables, and be capable of using numeric and categorical data. In addition, the assessment of results will be improved if the classification method identifies outliers in the training data, provides rankings of the importance of the input variables, and produces internal estimates of error and confidence of the output classification. Many decision tree classifiers fulfill all these requirements and have been used in land cover mapping for years [32–35], including several that use the meta-classifier random forest [36–38].

Our goal was to identify an optimal selection of input data from multiple sources and time periods of remotely sensed and ancillary data for accurate wetland mapping using random forest decision tree classification in a forested region of Northern Minnesota. We assessed ways of increasing classification accuracy, confidence, and practicality by assessing results from several combinations of input data. Our main questions for this study site in Northern Minnesota were: (1) how does classification accuracy and confidence of wetland mapping compare using different remote sensing platforms and ancillary data from different periods of the growing season; (2) what are the key input variables for accurate differentiation of upland, water, and wetlands, including wetland type; and (3) which datasets and seasonal imagery yield the best accuracy for wetland classification.

2. Methods

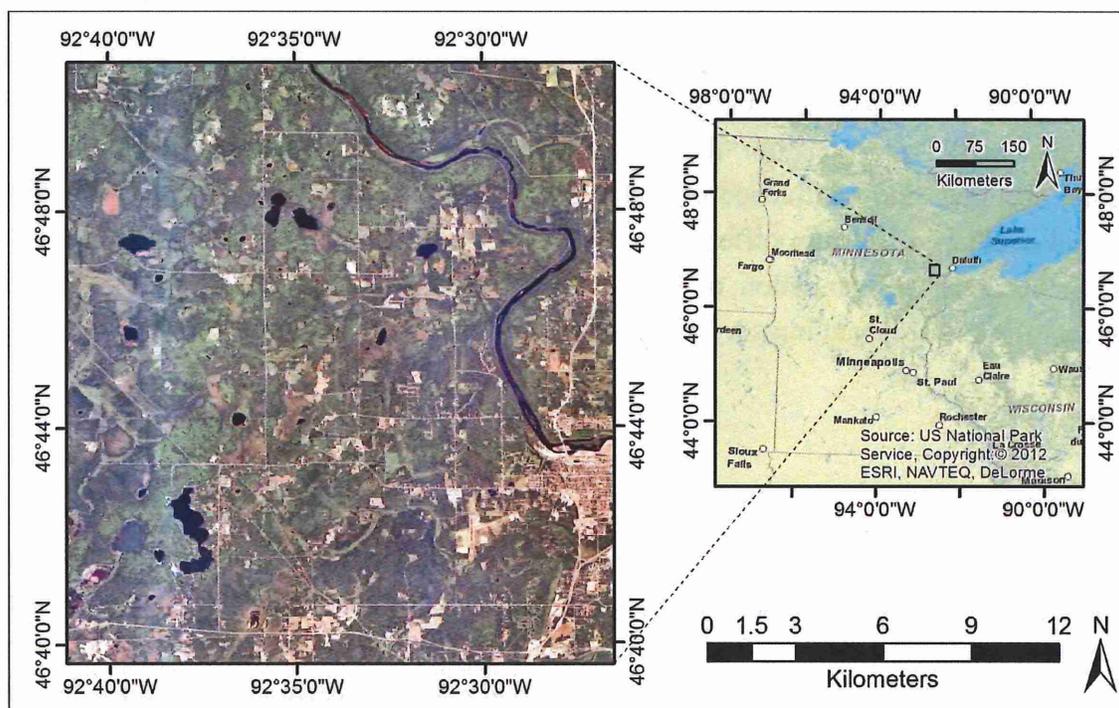
2.1. Study Area

Much of northern Minnesota (MN) is forested. The hydrographic patterns of the landscape have been influenced heavily by glacial advances and retreats over the millennia [39]. Our study centered on Cloquet, MN (Figure 1), which lies in the sparsely populated “Arrowhead” region of northeastern Minnesota. This study area is dominated by managed and natural hardwood and conifer forests, woody and herbaceous wetlands [40], and low density residential housing with a small city center (population

12,000) [41]. The elevation across the study area is 330–450 m above sea level (mean of 392 m), with the slope of the landscape averaging less than 1.7 degrees.

Given the variable nature of hydroperiod in space and over time, the weather during remotely sensed and field data acquisition is especially relevant when mapping wetlands. We collected field data in the summers of 2009 and 2010 and acquired remotely sensed data for several dates from 2008 to 2010. The 30-year normal total annual precipitation for the nearest major NOAA weather station in Duluth, MN (about 35 km away from the study site) measures between 5 and 10 cm in the spring, about 10 cm in the summer, and between 5 and 10 cm in the fall, for a total of about 79 cm annually [42]. The 30-year normal minimum precipitation in the spring is between 0.6 and 1.25 cm, with a maximum between 18 and 20 cm. In the summer the minimum precipitation is between 1.75 and 2 cm, with a maximum between 20 and 25 cm. The minimum precipitation in the fall is around 0.25 cm, with a maximum between 18 and 23 cm. Hydrologists in the northern hemisphere use the term water year to describe the period of time between 1 October and 30 September of the next calendar year. The lowest level of precipitation is in general during the fall and the landscape is typically replenished during the winter and spring of that water year. Precipitation over the study site during the 2008 water year (October 2007–September 2008) was slightly above normal, whereas the rest of that summer and well into the 2009 water year the trend was slightly below normal. Precipitation during the first part of the 2010 water year was slightly above normal around the study site and trended more towards normal throughout the north east region, whereas in the latter part of that year the trend was slightly below average [43].

Figure 1. Study area near Cloquet, Minnesota (MN). The aerial photo on the right is from the 2008 National Agricultural Imagery Program (NAIP).



2.2. Land Cover Classification Schemes

Two levels of classification were performed. The land cover classification schemes we used differentiated between upland, water, and wetland areas (Level 1) and sub-classified wetlands into wetland type (Level 2). Upland areas included all non-wetland classes, for example: urban, forest, grassland, agriculture, and barren land cover classes. Areas classified as wetland were sub-classified into a modified version of the Cowardin classification scheme [44], including the three most common wetland classes in the study area according to the National Wetlands Inventory (NWI) [45]: emergent, forested, and scrub/shrub wetlands. We merged the palustrine unconsolidated bottom class with the emergent wetland class and the riverine unconsolidated bottom class with the water class, based on visual assessment of the landscape variability in the study area (Table 1).

Table 1. Level 2 classification and our corresponding class modifications.

Level 2 Class	Modification of the Classes Used
Upland	Upland
Water	Water + Palustrine Unconsolidated Bottom
Emergent Wetland	Emergent + Riverine Unconsolidated Bottom
Forested Wetland	Forested Wetland
Scrub/Shrub Wetland	Scrub/Shrub Wetland

Any errors present in the initial Level 1 classification result prior to sub-classifying the wetland class can be propagated to the Level 2 classification [46–49]. We tested whether classification accuracy could be improved by developing a Level 2 classification directly from the full set of input data without first producing a Level 1 classification, but the results were too poor for further consideration. Thus, all subsequent Level 2 classification results and discussion represent a hierarchical sub-classification of the wetland class from the results of the corresponding Level 1 land cover classification.

2.3. Decision Tree Classification

We used random forest as the decision tree classifier for our study [50]. Generating decision trees was an efficient means of using our point reference training data to establish relations between our independent (remotely sensed and ancillary data) and dependent (field determined land cover class) variables to produce a land cover classification [51,52]. Random forest is a meta-classifier that consists of a collection (forest) of decision trees using training data. The decision trees were constructed with a random sample of input variables selected to split at each node [53]. The default number of variables selected equals the square root of the total number of input variables, which we held as a constant during forest growing. The decision trees were fully grown without pruning using a sample (with replacement) of about one-third of the training data. The cross-validation accuracy was calculated using the remaining training data (out-of-bag) and was used to evaluate the relative accuracy of each model prior to a formal accuracy assessment. Each tree produced a 'vote' for the final classification, where the final result was the class which had the highest number of votes [53]. The classification confidence, or probability, equals the ratio of the number of votes for a given class out of the total

number of trees generated, with a resulting value range of 0–1. For each model tested we ran 500 decision trees.

We built several random forest models per classification level by integrating different combinations of remotely sensed and ancillary input data to determine: (1) the most important data sources (corresponding to platform and wavelength of optical or radar data, and ancillary topographic and soils data derivatives), (2) the most significant input variables for mapping wetlands and classifying wetland type, and (3) the most effective temporal period (all data or only spring, summer, or fall season). Pre-defined combinations of input data are shown in Figure 2. We reviewed the top three models with highest overall accuracy for each classification level.

To determine if reducing the data load significantly changed the accuracy of the classification, we re-ran the top random forest models having the highest overall accuracy using only a selection of important variables—referred to as Reduced Data Load (RDL) models from this point forward. We used a combination of assessment measures from random forest (*i.e.*, mean decrease in accuracy and Gini index for the overall model and per class, explained in the Accuracy Assessment section below) and expert knowledge to assess variable importance. In the selection of important variables for the RDL, we thought it was valuable to have fair representation from all data sources and seasons, to incorporate both remote sensing and wetland science knowledge, and to utilize the measures of variable importance produced by the random forest classifier. For example, if a radar data variable was within the top 20 variables for either the Gini index or the mean decrease in accuracy, that variable was included in the RDL model based on our knowledge of the sensitivity of the radar signal to saturated conditions. Selection for the Level 2 RDL was complex. We considered variable importance measures for the overall model and for each of the three wetland classes, and we incorporated expert knowledge of specific input data layers for our final selection of the RDL. We selected 10 important variables for the Level 1 classification. We increased our selection to 15 variables for the Level 2 classification to accommodate anticipated overlap in the input data distributions between different classes.

2.4. Training and Test Reference Point Data

Reference training and test point data (Table 2) were compiled from randomly generated field sites visited in the summers of 2009 and 2010, from study sites of an existing wetland monitoring program (centroids from polygons of the 2006–2008 MN Department of Natural Resources Wetland Status and Trends Monitoring Program [19]), and from our expert knowledge in photo interpretation. The protocol for reference data collection in 2009 and 2010 involved several steps in the field: two different field crews were sent to locate random ground reference points with a GPS unit; crew members identified the dominant Cowardin wetland type [44] within a reasonable visual distance; crew members recorded basic observations about the site's characteristics; 2–5 photographs were taken per site; and crew members recorded the point ID, photo ID, Cowardin classification, and GPS coordinates in a back-up field book. Each field point represents a spatial area equal to the ground resolution of the input raster data used in the model (30 m). If the landscape surrounding the field point was not homogeneous within a reasonable visual distance, the field crew would use their discretion and move the GPS point to a new location which was more homogeneous. Empirical comparison of accuracies of results using different subdivisions of training and testing data [45] led us to use a

stratified random sample of 75% of the reference point data for training the random forest classifier and 25% of the reference point data for testing the accuracy of the results. Reference points were added to the training dataset via photo interpretation to maintain appropriate representation of land cover classes and to preserve a suitable spatial distribution of training points. Assessment of outliers in the training dataset integrated the proximity measure from random forest (described in more detail below), aerial and field photo interpretation, and expert knowledge to determine whether training sites were appropriate reference for their respective classes. We filtered only training sites; all testing sites were maintained in the reference set (Table 2). Spatial autocorrelation in either reference dataset was not formally addressed in this study.

Table 2. Summary of reference point data before and after the filtering of training sites.

Land Cover Classification	Training Sites Prior to Filtering	Final Sites for Model Training	Sites Used for Accuracy Testing	Final Total
Upland	464	305	136	441
Water	69	46	19	65
Wetland	421	402	149	551
Total	954	753	304	1057
Emergent Wetland	97	109	43	152
Forested Wetland	156	140	49	189
Scrub/Shrub Wetland	168	153	57	210
Total	421	402	149	551

The set of reference training data were evaluated for outliers using the proximity measure from the random forest classifier. Proximity was calculated by running the training dataset down each tree in the forest a second time, increasing the proximity value by one each time the training site occupied the same terminal node of the decision tree in the first and second run. The proximity measure was normalized by dividing by the total number of trees generated by random forest. Training sites with a low proximity measure may be outliers in the training data. For this study, the proximity measure was used to guide the selection and evaluation of training sites that were considered outliers. Each of the identified sites was evaluated and, subsequently, some of the sites were removed.

2.5. Input Datasets and Process Flow

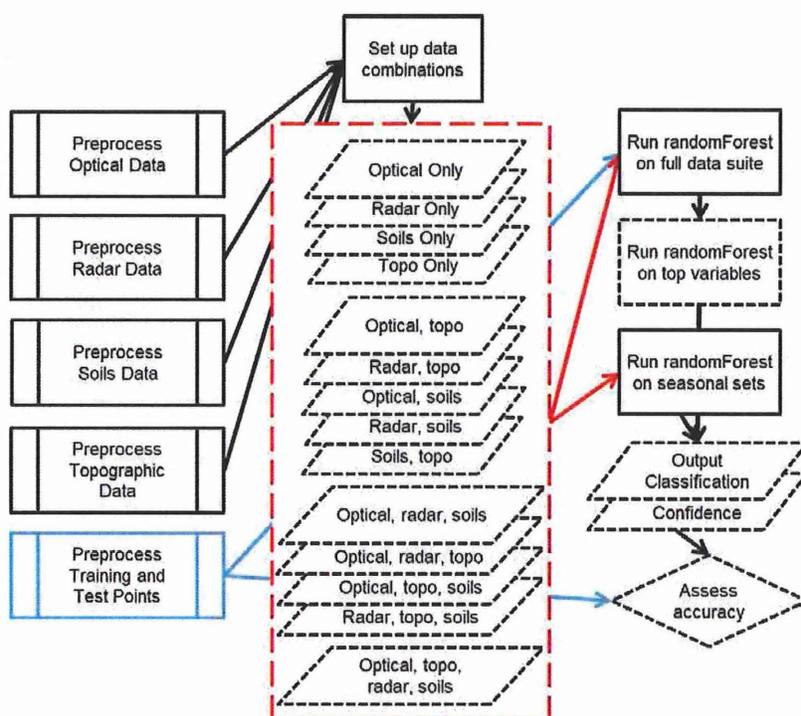
The implementation used to run random forest required that all raster data have the same spatial resolution and geographic extent. We chose to resample all raster data to match the layer with the coarsest resolution: Landsat 5 Thematic Mapper (TM) at 30 m spatial resolution. Resampling an image can introduce errors prior to classification [48], so we used the nearest neighbor sampling approach to minimize alteration of the original data values for our optical imagery. All input data were used in raster format and coregistered using ERDAS Imagine (v. 2010) with a root mean square error (RMSE) of less than 15 m.

In all of the tables and figures to follow, if a data source/platform is mentioned (e.g., “Landsat TM” or “radar”), all data layers from that source/platform are included in the tested combination. For example, the “All Season, All Data” model which uses Landsat TM, PALSAR, and Soils data includes

all Landsat TM bands and derivatives from all dates (Table 2), all PALSAR polarizations from all dates, and all Soils data layers.

Following preparation of input datasets and training point data, we ran random forest to generate classification and confidence layers based on predefined combinations of datasets, including combinations of different platforms and seasons described earlier. We used our test point data to assess accuracy of each of the output classifications (Figure 2).

Figure 2. Data process flow. Preprocessing of input datasets and reference point data (shown in blue) are in the left-hand column. Combinations of datasets used to perform random forest (shown in red), along with generation of the output classification, confidence maps, and accuracy assessment are referenced by boxes in the right-hand column.



2.5.1. Topographic Input Data

We used the US Geological Survey (USGS) National Elevation Dataset (NED) [54] (10 m resolution resampled to 30 m) to determine elevation and derive slope gradient, aspect, curvature, and flow accumulation across the study area. The accuracy of this dataset varied spatially, but the overall vertical root mean square error was 2.44 m. We applied the flow accumulation function provided by the Environmental Systems Research Institute (ESRI) ArcGIS (v. 10.0) to calculate the direction(s) of water flow across the landscape and accumulate flow for all downslope cells. Cells with high flow accumulation imply areas of concentrated flow, such as stream channels, and cells with low flow accumulation likely are ridges or plateaus [55]. The curvature metric is a second derivative of slope and influences the convergence and divergence of water flow [23]. The topography of this study area does not vary significantly (330–450 m elevation, 392 m mean elevation, 20 m standard deviation; 0–37 degree slope with an average of 1.7 degrees). Compared to the height distribution of the study area, the vertical accuracy of the dataset has a negligible RMSE.

2.5.2. Soils Input Data

Soil attributes are defining variables in all working definitions of wetland areas [44]. Though soils data are not available everywhere and the quality of the maps that are available may be questionable, we tested the effectiveness of including or not including soils data in this study. We extracted soils tabular and vector data from the US Department of Agriculture (USDA) Soil Survey Geographic Data Base (SSURGO) [56]. The following data layers were used based on their likelihood to be associated with wetland areas: soil type (e.g., mucky peat, loam), dominant and wettest drainage class (e.g., moderately well drained, poorly drained, and somewhat poorly drained), and hydric class (e.g., hydric, or partially hydric) [25,57]. We joined the tabular and vector data for these four soils data layers and then converted the layers to raster format with 30 m spatial resolution.

2.5.3. Optical Input Data

Northern Minnesota is frequently cloudy, particularly in the summer, making it a challenge to find cloud-free conditions over our study area. The only Landsat TM imagery available with adequate cloud-free conditions was from early spring and fall (Table 3). We used blue (TM Band 1, B), green (TM Band 2, G), red (TM Band 3, R), near-infrared (TM Band 4, NIR), two mid-infrared (TM Band 5, MIR1; and TM Band 7, MIR2), and thermal infrared (TM Band 6, TIR) bands from all image dates. We included NIR, MIR1, MIR2, and TIR because of their suitability for land cover mapping and detecting water content in plants and soil [58,59]. Though multi-temporal and multi-platform data were used, the acquired satellite data were not atmospherically corrected and the data remained in digital number format. All of the input data were integrated into a single dataset, from which the training data were derived to classify land cover as a single snapshot [60].

Table 3. Input optical data for decision tree classification.

Season	Date	Band Combinations	Platform-Source
Spring	17 April 2010	B, G, R, NIR, MIR1, MIR2, TIR	Satellite-Landsat 5 TM
	19 May 2010	B, G, R, NIR, MIR1, MIR2, TIR	Satellite-Landsat 5 TM
	June 2009	B, G, R, NIR	Aerial Orthophoto-NAIP
Summer	August 2008	B, G, R, NIR	Aerial Orthophoto-NAIP
	August 2010	B, G, R	Aerial Orthophoto-NAIP
Fall	21 September 2009	B, G, R, NIR, MIR1, MIR2, TIR	Satellite-Landsat 5 TM
	4 October 2008	B, G, R, NIR, MIR1, MIR2, TIR	Satellite-Landsat 5 TM

We calculated both the normalized difference vegetation index (NDVI) and Tasseled Cap transformations for each TM image date. NDVI has been useful for separating vegetated versus non-vegetated areas and wet versus dry areas [61]. The brightness, greenness, and wetness axes of the Tasseled Cap transformation [62,63] have a long record of use in improving classification results, assessing land cover change, and aiding in estimates of forest structure and disturbance [64–66].

Due to the aforementioned challenge to find cloud-free imagery during the summer season over our study area, we also acquired aerial orthophotos from the US Department of Agriculture (USDA) Farm Service Administration (FSA) National Agricultural Imagery Program (NAIP) for August 2008 and

2010 and an additional orthophoto from June 2009 (early leaf onset) to increase our temporal coverage of optical data during the summer season. The 2008 and 2009 images were acquired with visible and near infrared bands (blue, green, red, NIR), whereas the 2010 image was collected only in visible bands (blue, green, red). We used the red and near infrared bands to calculate NDVI for both 2008 and 2009. All aerial orthophotos were resampled to 30 m spatial resolution.

2.5.4. Radar Input Data

We used synthetic aperture radar (SAR) from RADARSAT-2 (C-band, 5.6 cm wavelength) and Advanced Land Observing Satellite (ALOS) Phased Array type L-band Synthetic Aperture Radar (PALSAR) (L-band, 23.6 cm wavelength) satellite systems (Table 4). We obtained two fully polarized RADARSAT-2 images (15 June 2009 and 19 September 2009) through the Canadian Space Agency's Science and Operational Applications Research (SOAR) Program. Two additional dates (9 July 2009 and 26 August 2009) were made available by the Canada Center for Remote Sensing (CCRS). Though proprietary data and licensing restrictions prohibited us from incorporating the backscatter data from the dates provided by CCRS, we were able to generate polarimetric decompositions for use in our analysis (all preprocessing steps were performed in the same manner, described below). All of the RADARSAT-2 imagery was provided by the vendor with the constant beta application look up table (LUT) applied to avoid over saturation of the data [67]. Table 4 outlines which dates included the backscatter plus polarimetric decompositions ("Full dataset") and which dates did not include backscatter ("Decomp only").

We used the software package PCI Geomatica (v. 9.1) to preprocess the RADARSAT-2 imagery and generate polarimetric decompositions. Prior to resampling the imagery, we applied a boxcar filter (7×7 moving window) to reduce speckle and increase the number of effective looks for polarimetric decomposition [68]. The data were then resampled to 30 m using a mean window after terrain correcting the imagery. We then radiometrically corrected the data, performed antennae pattern correction, converted the amplitude values to sigma naught (σ^0 ; output scaling LUT), and scaled the backscatter values in decibels for quantitative analysis [69]. After preprocessing the imagery as described above, we generated polarimetric decompositions.

We used three types of polarimetric decompositions on the RADARSAT-2 imagery to assess the benefits of radar polarimetry for mapping wetlands: van Zyl, Freeman-Durden, and Cloude-Pottier [70]. The premise behind a polarimetric decomposition is that the received signals contain important information regarding the structure of the landscape target, the scattering mechanism of the return signal, and the apparent shift in the phase of the signal from the target [71–74]. The van Zyl decomposition is a classification [70] based on the backscatter and number of phase shifts that occur in the returned signal, where each pixel is discretely classified as having a single, odd, or diffuse dominant backscatter. The Freeman-Durden decomposition [75] models the target scattering mechanisms as a continuous variable where each pixel represents relative proportions of surface scattering, double bounce, and volume scattering. The Cloude-Pottier decomposition [76] uses parameters of entropy, alpha angle, and anisotropy calculated from the eigenvalues and eigenvectors of a coherency matrix. Entropy is the randomness of scattering mechanisms, alpha angle represents the dominant scattering mechanism, and anisotropy characterizes directional dependence and importance

of the secondary scattering mechanism. Among these three polarimetric decompositions, many authors have found the Freeman-Durden decomposition in particular to be useful for wetland mapping [77–79]. These polarimetric decompositions represent the advanced analysis possible with radar polarimetry and thus were included in the random forest models which evaluated the effectiveness of RADARSAT-2 imagery for mapping wetlands.

We also acquired three dual-polarized (horizontal-horizontal (HH) and horizontal-vertical (HV)) ALOS PALSAR images (29 July and 11 September 2009 and 14 June 2010) for the study area from the Alaska Satellite Facility (ASF) archive (Table 4). We used the software package MapReady (v. 2.3), available through the ASF, for preprocessing the PALSAR data. The imagery was geocoded and resampled to 30 m spatial resolution using the default method, bilinear interpolation, which considers four neighboring pixel values. MapReady was used to perform antenna pattern correction using the beta coefficient, scale the data to decibel backscatter, and perform radiometric and geometric terrain correction using the 10 m NED elevation dataset. The RADARSAT-2 and PALSAR imagery were preprocessed using different LUTs and it is assumed that any resulting differences are negligible.

Table 4. Input radar data for decision tree classification.

Season	Date	Source	Acquisition Mode *	Incidence Angle	Product
Spring	15 June 2009	RADARSAT-2	FBQ	26.9 near, 28.7 far	Full dataset
	14 June 2010	PALSAR	FBD	34.3 center	Full dataset
Summer	09 July 2009	RADARSAT-2	FBQ	26.9 near, 28.7 far	Decomp only
	29 July 2009	PALSAR	FBD	34.3 center	Full dataset
	26 August 2009	RADARSAT-2	FBQ	26.9 near, 28.7 far	Decomp only
Fall	11 September 2009	PALSAR	FBD	34.3 center	Full dataset
	19 September 2009	RADARSAT-2	FBQ	26.9 near, 28.7 far	Full dataset

* FBQ: Fine Beam Quad-polarization; FBD: Fine Beam Dual-polarization.

2.6. Accuracy Assessment

We reserved a stratified random subset of 25% of the reference point data and implemented traditional methods to assess accuracy and evaluate results. We constructed error matrices with overall accuracy, 95% confidence intervals (CI), User's and Producer's accuracies, kappa statistic (\hat{k}), and ran significance tests of error matrix \hat{k} values [80] for all random forest classification models. We performed two error matrix significance tests for each of the land cover classification levels: (1) between the most accurate random forest model with the full data suite to the same model with only a selection of the most important variables (RDL), and (2) between the most accurate random forest model with the full data suite to the most accurate random forest model using only data from a seasonal snapshot. Asterisks were used next to table values that were significant at an alpha of 0.05. We also conducted an accuracy assessment of the original NWI for comparison to our accuracy results.

Outputs from random forest provide unique complements to traditional accuracy assessment, including: (1) cross-validation, using the out-of-bag sample of training data to evaluate relative accuracy of each model prior to a formal accuracy assessment; (2) classification confidence, or probability, calculated by the number of times a given class was designated as the final class out of the total number of trees, with a resulting value range of 0–1; (3) mean decrease in accuracy, calculated

per input data layer, giving insight to how influential a layer was on the overall accuracy; and (4) Gini index, which aids in evaluating the influence of input layers on the structure of the decision trees.

To calculate mean decrease in accuracy, the sample of reference data that was retained during the growth of each decision tree (out-of-bag) was used to determine the relative change in accuracy by including or excluding a particular variable. The normalized change in cross-validation accuracy was totaled after all decision trees were run and represents the relative importance of that variable [53]. The Gini index is calculated by, starting with an index value of 1, reducing the index value per variable every time that variable was used to make a dichotomous split in each decision tree. This index value was totaled per variable and represents the relative influence of that variable on the structure of each decision tree [53]. The most important variables in the random forest model can be inferred by evaluating both the mean decrease in accuracy and Gini index.

3. Results

3.1. Upland, Water, and Wetland Land Cover Classification (Level 1)

The most accurate full season random forest model for the Level 1 classification (85% accurate) integrated all available Landsat 5 TM, topographic, PALSAR, and soils data. The error matrix (Table 5) shows this model confused upland areas with wetland areas about 29% of the time (commission error calculated from the User's accuracy), but wetland areas were confused with upland areas only 4% of the time. In terms of Producer's accuracy (omission error), reference upland areas were more often correctly classified as uplands (94%) compared to the wetland class (78%). The water class was highly accurate in terms of both Producer's and User's accuracies (100% and 95%, respectively).

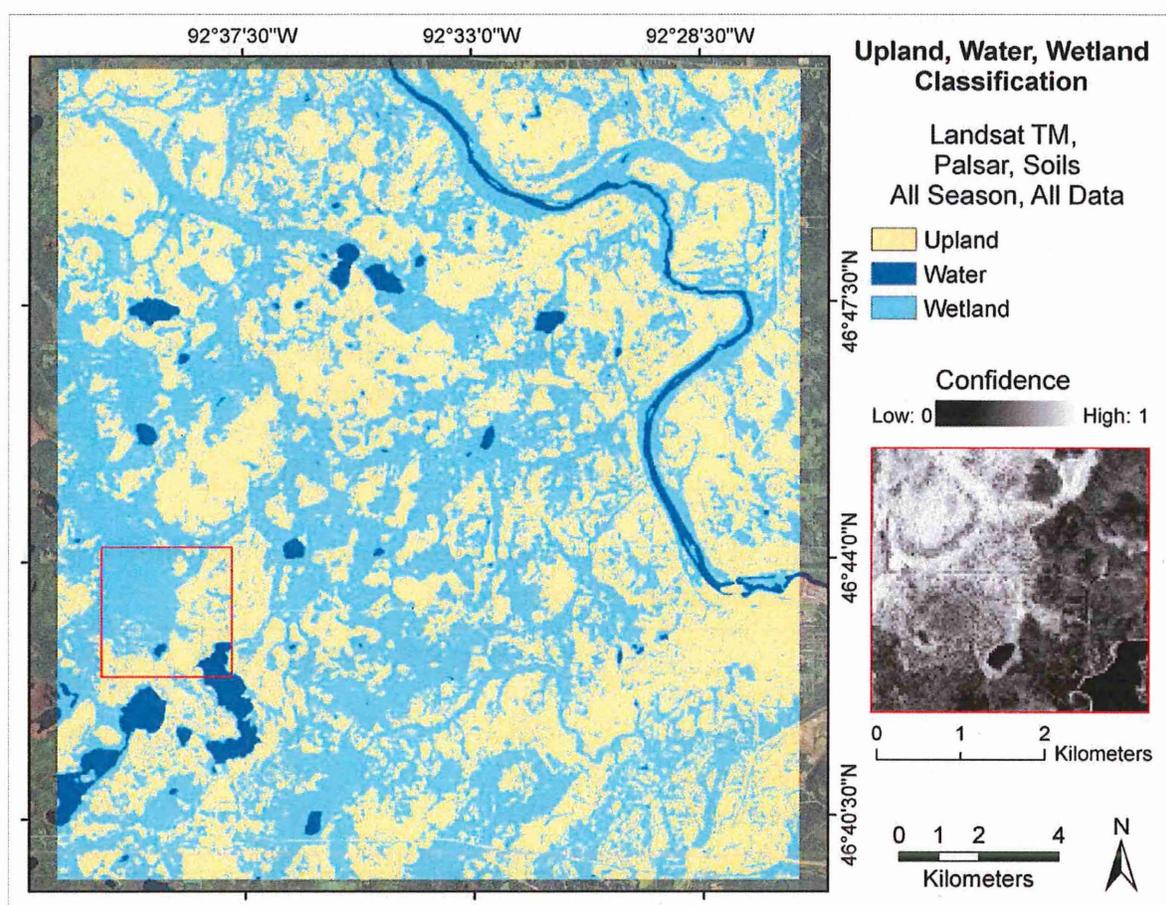
Table 5. Classification error matrix for the most accurate full season random forest model for the Level 1 classification which incorporated all available Landsat 5 TM, topographic, PALSAR, and soils data.

		Reference Data				
		Upland	Water	Wetland	Row Total	User Accuracy (%)
Classified Data	Upland	97	0	39	136	71
	Water	0	18	1	19	95
	Wetland	6	0	144	150	96
	Column Total	103	18	184	305	
	Producer Accuracy (%)	94	100	78		Overall = 85% k-hat = 0.73, 95% CI ± 4%

The second and third most accurate full season random forest models for the Level 1 classification had overall accuracies of 84% and 83%, respectively (Table 7). The second most accurate model incorporated all available Landsat 5 TM, aerial orthophoto, topographic, PALSAR, and soils data. This result shows that adding aerial orthophotos changes the accuracy by a very small amount (<1%). The third most accurate model incorporated all available Landsat 5 TM, topographic, RADARSAT-2, PALSAR, and soils data. This result shows that adding RADARSAT-2 data changes the accuracy by about 2%.

The classification map for the best Level 1 model illustrates how wetlands dominate the study landscape (Figure 3). The confidence for the resulting land cover classification (see representative area subset in Figure 3) was relatively high for most of the area classified as wetland, particularly around the shoreline of water bodies and in larger wetland complexes. Areas of lower confidence may be prone to misclassification from high variability or data redundancy in the input variables. We also tested a full season reduced data load (RDL) model to evaluate if using only the top 10 important variables significantly changed the accuracy of the classification.

Figure 3. Output classification of the most accurate full season random forest model for the Level 1 land cover classification using all available Landsat 5 TM, topographic, PALSAR, and soils data.



We identified the top 10 variables using expert knowledge and the mean decrease in accuracy and Gini index values for each variable in the Landsat 5 TM, topographic, PALSAR, and soils model (Table 6). The overall accuracy of classification results from the RDL model (Table 7) was 81% ($\pm 4\%$) with generally lower values of Producer's and User's accuracies. However, a significance test of the difference between the full data suite and RDL models was not significant at an alpha level of 0.05. There was a small difference in the resulting wetland area between the two models: the full season model had a slightly lower total wetland area (18,969 ha) than the RDL model (19,010 ha). Though the difference in wetland area was negligible, a difference map of the results from the two models revealed widespread spatial differences, without pattern, due to more isolated pixels throughout the RDL model.

When ancillary datasets were used without the addition of remotely sensed data, the accuracy was significantly reduced. Classifying upland, water, and wetlands using topographic and soils data produced a higher accuracy (74%) than a model with soils data alone (73%) or topographic data alone (62%). Conversely, the best classification result without ancillary data, using only Landsat TM and PALSAR imagery, was still less accurate (80%) than models which used both remotely sensed and ancillary data (85%). All comparisons made here were statistically significant at an alpha level of 0.05. These findings show that integrating ancillary datasets with remotely sensed data can statistically improve accuracy of mapping wetlands.

Table 6. Top 10 important variables, in order of importance, selected from the most accurate full season random forest model used in a Reduced Data Load (RDL) model for the Level 1 classification.

Data Type	Date	Source
NIR Band	19 May 2010	Landsat 5 TM
Hydric Soils	NA	USDA SSURGO
MIRI Band	21 September 2009	Landsat 5 TM
Elevation	NA	USGS NED
Curvature	NA	USGS NED
Green Band	4 October 2008	Landsat 5 TM
Red Band	4 October 2008	Landsat 5 TM
Blue Band	17 April 2010	Landsat 5 TM
NDVI	17 April 2010	Landsat 5 TM
HV Polarization	14 June 2010	PALSAR

Table 7. Error matrix summary of the three best full season random forest models for the Level 1 land cover classification, as compared to the NWI.

Model	Overall Accuracy (%)	Kappa Statistic	Z Statistic
Best: TM, topo, PALSAR, soils (Table 5)	85	0.73	19.4*
RDL: top variables in best model (Table 6)	81	0.67	16.3*
2 nd Best: TM, aerial, topo, PALSAR soils	84	0.71	18.3*
3 rd Best TM, topo, RSAT-2, PALSAR soils	83	0.68	17.2*
National Wetlands Inventory	70	0.46	9.6*

* Values were significant at an alpha of 0.05.

We also evaluated results of different models from a temporal perspective to determine the influence of season for data acquisition on classification results (Table 8). Input data from different platforms were available for different periods of the growing season (Tables 3 and 4), a situation typical of multi-platform analyses and worth investigating. The seasonal model with the best accuracy (85%) was constructed from spring season data and had an overall accuracy comparable with the full season model. When we compared the full season and spring season models, the full season model had a lower total wetland area (18,969 ha) than the spring season model (19,679 ha). A difference map of the results from the two models did not reveal significant widespread spatial differences, but there was an observed pattern of differences occurring along roads and land cover transition zones; meaning, the two models have slight differences in feature boundaries. The most accurate model using fall data had

an overall accuracy of 82% and the best model constructed from summer data had the least accurate results at 79%.

Table 8. Summary of results for the best seasonal random forest models for the Level 1 land cover classification.

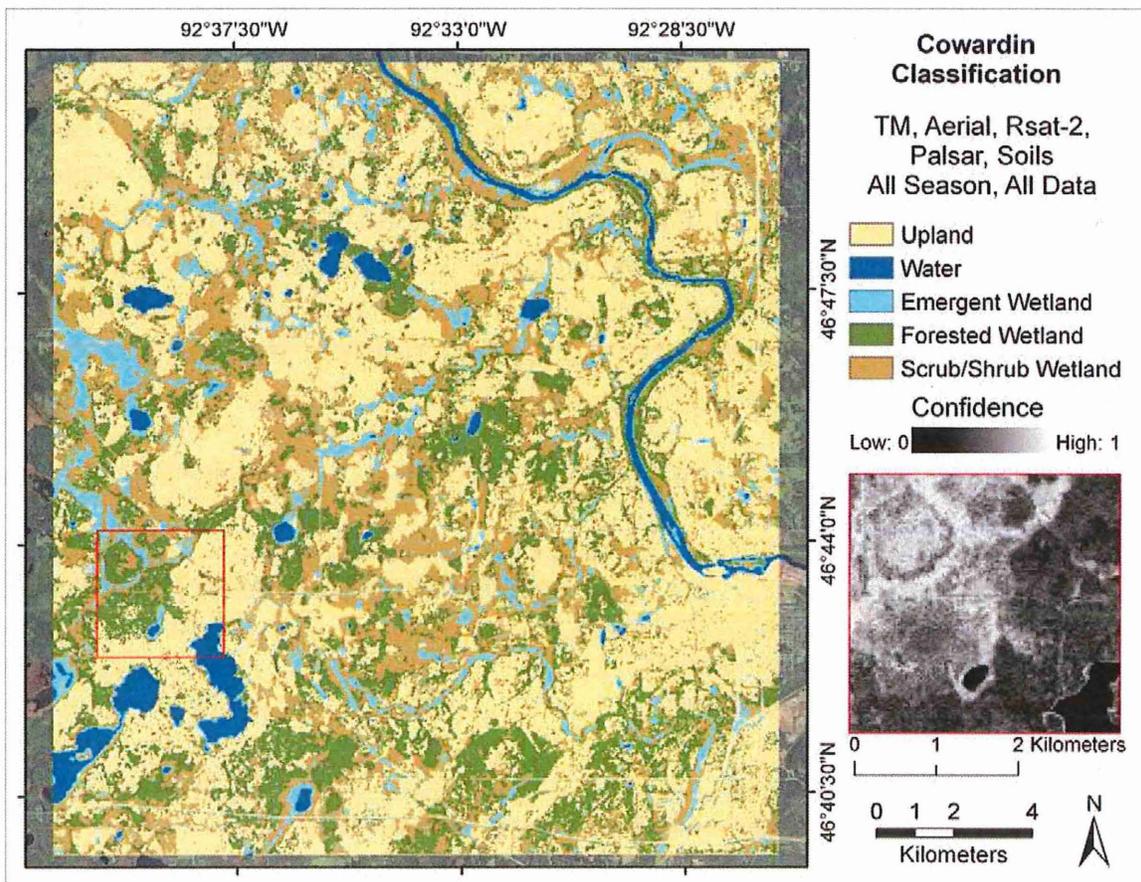
Season	Model	Overall Accuracy (%)	Kappa Statistic	Z Statistic
Spring	TM, topo, PALSAR, soils	85	0.72	19.1*
Summer	Aerial, topo, PALSAR, soils	79	0.63	14.5*
Fall	TM, topo, RSAT-2, PALSAR, soils	82	0.67	16.3*
Full Season	TM, topo, PALSAR, soils	85	0.73	19.4*

* Values were significant at an alpha of 0.05.

3.2. Cowardin Wetland Classification (Level 2)

The most accurate full season random forest model for the Level 2 classification integrated all available Landsat 5 TM, aerial orthophoto, topographic, RADARSAT-2, PALSAR, and soils data to yield an overall accuracy of 69% ($\pm 5\%$) (Figure 4). The overall accuracy for this model prior to sub-classifying the wetland class was 84% ($\pm 5\%$), with the Producer’s and User’s accuracies for the wetland class at 79% and 93%, respectively ($\pm 6\%$ and 4%, respectively).

Figure 4. Output classification of the most accurate full season random forest model for the Level 2 land cover classification using all available Landsat 5 TM, aerial orthophoto, topographic, RADARSAT-2, PALSAR, and soils data.



The error matrix for results from the best Level 2 classification model (Table 9) shows that upland areas were confused with wetland areas about 28% of the time (User's accuracy was $72\% \pm 8\%$). The forested wetland class had the highest User's accuracy ($71\% \pm 13\%$) and the emergent wetland class had the highest Producer's accuracy ($65\% \pm 5\%$). Reference upland sites were classified correctly as uplands 92% of the time ($\pm 5\%$). Reference emergent wetlands were classified correctly 65% of the time ($\pm 5\%$), but forested and scrub/shrub wetlands were classified correctly only about half of the time (49% and 48%, respectively, $\pm 12\%$ for each). Both forested and emergent wetlands tended to be confused with scrub/shrub wetlands. The water class was highly accurate for both Producer's and User's accuracies (95% for each, $\pm 11\%$ for each).

Table 9. Classification error matrix for the most accurate full season random forest model for the Level 2 classification which incorporated all available Landsat 5 TM, aerial orthophoto, topographic, RADARSAT-2, PALSAR, and soils data.

		Reference Data						Row Total	User Accuracy
		Upland	Water	Emergent Wetland	Forested Wetland	Scrub/Shrub Wetland			
Classified Data	Upland	98	0	4	21	14	137	72	
	Water	0	18	1	0	0	19	95	
	Emergent Wetland	5	1	24	1	12	43	56	
	Forested Wetland	3	0	0	35	11	49	71	
	Scrub/Shrub Wetland	1	0	8	14	34	57	60	
	Column Total	107	19	37	71	71	305		
	Producer Accuracy	92	94	65	49	48		Overall = 69% k-hat = 0.58 95% CI $\pm 5\%$	

The second and third most accurate full season random forest models for the Level 2 classification had overall accuracies of 66% and 65%, respectively (Table 10). The second most accurate model incorporated all available Landsat 5 TM, topographic, RADARSAT-2, PALSAR, and soils data. This result shows that when we do not include aerial orthophotos, the overall accuracy in sub-classifying wetlands changes by about 3%. The third most accurate model incorporated all available Landsat 5 TM, aerial orthophoto, topographic, RADARSAT-2, and soils data. This result shows that when we do not include PALSAR data, the overall accuracy in sub-classifying wetlands decreases by about 4%.

Table 10. Error matrix summary of the three best full season random forest models for the Level 2 classification.

Model	Overall Accuracy (%)	Kappa Statistic	Z Statistic
Best: TM, aerial, topo, RSAT-2, PALSAR, soils (Table 9)	69	0.58	16.4*
RDL: top variables in best model (Table 11)	63	0.50	13.7*
2 nd Best: TM, topo, RSAT-2, PALSAR soils	66	0.55	15.3*
3 rd Best: TM, aerial, topo, RSAT-2, soils	65	0.53	14.6*
National Wetlands Inventory	55	0.38	11.0*

* Values were significant at an alpha of 0.05.

We assessed a full season RDL model to evaluate whether using the most important variables significantly changed the accuracy of the results for the Level 2 classification. We used expert knowledge and the mean decrease in accuracy and Gini index values for each variable in the full model to identify the top 15 variables for a RDL model (Table 11). The accuracy of the RDL model (Table 10) was 63% ($\pm 5\%$), which did not differ significantly from the accuracy of the full model at an alpha level of 0.05. There was a difference in the resulting wetland area between the two models: the full season model had a lower total wetland area (18,351 ha) than the RDL model (20,376 ha). Most of the difference in area between the two models was from forested and scrub/shrub wetland classes erroneously classified as upland areas in the full season model. A difference map of the results from the two models revealed widespread spatial differences with an observed pattern of classification differences occurring along roads and land cover transition zones.

Table 11. Top 15 important variables, in order of importance, selected from the most accurate full season random forest model used in a RDL model for Level 2 classification.

Data Type	Date	Source
TC Greenness	19 May 2010	Landsat 5 TM
NDVI	19 May 2010	Landsat 5 TM
TIR Band	17 April 2010	Landsat 5 TM
MIR1 Band	21 September 2009	Landsat 5 TM
TC Wetness	21 September 2009	Landsat 5 TM
MIR1 Band	4 October 2008	Landsat 5 TM
HH Polarization	21 September 2009	PALSAR
HV Polarization	21 September 2009	PALSAR
NDVI	17 April 2010	Landsat 5 TM
NDVI	Summer 2008	NAIP
TC Wetness	19 May 2010	Landsat 5 TM
TC Wetness	4 October 2008	Landsat 5 TM
HH Polarization	14 June 2010	PALSAR
HV Polarization	14 June 2010	PALSAR
HV Polarization	29 July 2009	PALSAR

We evaluated several models to determine the extent to which season and corresponding data platforms could influence results for the most accurate Level 2 classification (Table 12). Satellite data from the spring yielded the most accurate results (71%), exceeding the level of accuracy produced by the full season model (69%). When we compared the full season and spring season models, the full season model had a higher total wetland area (18,351 ha) than the spring season model (17,162 ha). Most of the difference in area between these two models was from forested and scrub/shrub wetland areas erroneously classified as the upland class in the spring season model. A difference map of the results from the two models revealed less significant widespread spatial differences and no apparent pattern. The most accurate model using summer data had an overall accuracy of 65%. The best model constructed from fall data had the least accurate results at 62%.

Table 12. Error matrix summaries of the best seasonal random forest models for the Level 2 classification.

Season	Model	Overall Accuracy (%)	Kappa Statistic	Z Statistic
Spring	TM, topo, PALSAR, soils	71	0.60	17.3*
Summer	Aerial, topo, soils	65	0.51	13.8*
Fall	TM, topo, RSAT-2, PALSAR, soils	62	0.48	13.0*
Full Season	TM, aerial, topo, RSAT-2, PALSAR soils	69	0.58	16.4*

* Values were significant at an alpha of 0.05.

4. Discussion

A key challenge in mapping and monitoring the landscape with remotely sensed data is that temporal coverage can be limited because of cloud contamination of imagery and because overpass schedules and return frequencies vary from platform to platform. Conditions for our research were no exception. This motivated us to examine the importance of type and seasonal timing of source data for classifying wetland-dominated landscapes in a forested region of the Upper Midwest.

4.1. Upland, Water, and Wetland Land Cover Classification (Level 1)

Our best Level 1 classification (85%) relied on ancillary soils, topographic, and remotely sensed data from satellite optical (Landsat 5 TM) and radar (PALSAR) platforms. This most accurate model used remotely sensed variables from fewer data sources than did the second (84%) and third best (83%) models, and did not require full temporal coverage (Table 7). A possible reason the best model did not place importance on summer data (aerial orthophotos), according to the Gini index and mean decrease in accuracy values, was that the fully developed tree canopy obscures underlying landscape features (*i.e.*, inundation, wetland plant species, *etc*) that could otherwise reduce confusion in classifying vegetated upland and wetland areas [81]. The third most accurate model included RADARSAT-2 imagery and polarimetric decompositions, along with ancillary soils and topographic data. The fact that these particular radar datasets were not incorporated in the most accurate model implies that the C-band imagery was not as appropriate as L-band imagery for mapping wetlands in a forested region, primarily due to better propagation of the longer wavelength radar signal through the tree canopy. These results echo findings elsewhere that, though filtering techniques may vary, the high variability from radar backscatter in C-band imagery can confuse the model and cause a reduction in accuracy [34]. Though none of the three best models or the RDL model were significantly different from each other (at an alpha level of 0.05), all four models were significantly more accurate than the original NWI.

Reducing the number of variables in the Level 1 model to only the 10 most important variables produced results that were 4% less accurate than obtained with the full data suite model. However, this accuracy still was relatively high (81%) and enabled us to remove nearly 50 variables from the full model, thereby increasing classification efficiency and reducing cost without sacrificing a significant level of accuracy. Furthermore, our assessment of seasonal data sources suggests that imagery from spring alone can provide comparable results with imagery distributed throughout the entire growing season (Table 8). Most of the spring input data used in the model corresponded with above-normal

precipitation conditions, confirming findings from other research that precipitation conditions are highly relevant to differentiating upland, water, and wetland classes [61,82]. Our results show the effectiveness of targeting input variables acquired during the spring season in this geographic region to improve land cover classification accuracy and confidence.

Results of the RDL model for this classification level showed that in addition to elevation, curvature, and hydric soils data, the most important spring season data included: satellite blue and NIR bands, satellite NDVI, and HV polarization using L-band radar. The satellite blue band, which had a high importance based on the mean decrease in accuracy for the upland class, was acquired on an especially clear day (17 April 2010) and thus had very little atmospheric interference, which typically makes this band noisy and not as useful. Others have found the blue band to be useful in classifying upland classes such as bare soil and in masking out shadowed areas [83,84]. Other studies have confirmed these remotely sensed variables, particularly near infrared and NDVI, are important for land cover classification and land cover change mapping. Such variables are particularly important when discriminating between forest structural condition (*i.e.*, open or closed canopy), monitoring stand age and regrowth, and estimating species composition and richness [85–87]. Studies have also established that the multiple scattering and subsequent depolarization of the radar signal explains the importance of HV polarization for classifying land cover and estimating biomass, particularly in forested regions [72,78,87]. It is important to note that even though our best results included ancillary soils and topographic input data, without the inclusion of ancillary data, the selected remotely sensed layers in the RDL model retain their level of importance.

4.2. Cowardin Wetland Classification (Level 2)

The second and third most accurate models (66% and 65%, respectively) developed for the Level 2 classification relied on fewer data sources than used by the best model and performed better than the RDL model. None of the three best models or the RDL model were significantly different from each other (at an alpha level of 0.05), but all four were a statistical improvement over the NWI (Table 10). Sub-classifying wetlands accurately required ancillary soils and topographic data, as well as increasing the temporal and spectral coverage of remotely sensed data with optical and L-band radar, the latter undoubtedly because of deeper canopy penetration and increased interaction of the signal which has been known to be useful for distinguishing differences in vegetative land cover [26–28,88].

Our attempts to produce a RDL model using the top 15 variables from the full data suite indicated too great a reduction in classification accuracy for distinguishing between wetland types, even with the inclusion of ancillary soils and topographic data. The top 15 variables used in the model, though important, do not sufficiently represent the variation in characteristics needed to sub-classify wetlands. However, results from our seasonal analysis suggest output from a RDL model might be improved if we selected for spring data, as the spring model produced the highest accuracy for the Level 2 classification (Table 12).

We observed fewer differences from a visual comparison of results between the full season and spring models than between the full season and RDL models, but wetland class confidence was somewhat higher with full data suite (118 input data layers) than with the spring season model (33 input data layers). Though in some cases classification accuracy can be improved by increasing the

number of input data layers [89], research has also shown that increasing the number of discrete classes requires comparable increases in training data to improve the sensitivity of classifiers to more refined class differences [90,91]. Results from our efforts to model Cowardin wetland classes indicate that our model might benefit from additional reference training sites, particularly for the forested and scrub/shrub wetland classes which had very low accuracy compared to the emergent wetland class.

The most important variables selected for a RDL model of the Level 2 classification incorporated a rather different set of data sources and seasons (Table 11) than were selected for the Level 1 classification (Table 6). The most important variables for sub-classifying wetlands included remotely sensed data from a broader temporal range than for simply differentiation between upland, water, and wetland areas. Many studies have found multi-temporal data to aid in land cover classification, particularly for wetland mapping [69,73,88,92]. The Level 2 model made use of thermal data and Tasseled Cap transformation derivatives, as well as a much greater use of radar data. Other studies have confirmed that thermal data is important for land cover classification, particularly in separating vegetated and impervious areas and different moisture levels throughout the landscape [58,93]. The Tasseled Cap transformation also has been used by others to improve wetland mapping [81,94]. We found that using radar backscatter was more useful than using the polarimetric decompositions; in particular, our findings further confirm those of others documenting the importance of co- and cross-polarization radar backscatter (HH and HV, respectively) in classifying land cover [95–97].

5. Conclusions

One of our main goals was to identify an optimal selection of input data from various sources of remotely sensed and ancillary data to accurately map wetland areas in Northern Minnesota. We accomplished this goal by rigorously testing the results from several combinations of data at two classification levels. We found that the key input variables for accurately differentiating between upland, water, and wetland areas include satellite red, near infrared (NIR), and middle infrared (MIR1) bands and normalized vegetation index (NDVI), elevation and curvature, hydric soils ancillary data, and L-band horizontal-vertical (HV) polarization. We conclude that, in addition to the variables used for the Level 1 classification, the key input variables for a Level 2 classification of wetlands include Tasseled Cap Greenness and Wetness, satellite thermal band, and L-band horizontal-horizontal (HH) polarization. Our sound methods have generated an important set of results for the remote sensing community, describing in detail the differences in accuracy of wetland mapping in a forested region using specific data sources and combinations.

Weather conditions over the study site during the water years October 2007-September 2010 were relevant to conclusions made regarding seasonal data importance. This is because precipitation, and any subsequent deviation from the 30 year normal, influences the site's hydrologic characteristics prior to data acquisition. The important spring datasets identified in Tables 5 and 9 all correspond to above normal precipitation conditions. With the exception of the summer of 2008, the rest of the important summer and fall datasets were acquired during below normal precipitation conditions. Though it is possible to plan spring data acquisition knowing the water year trends from the fall and winter before, it is difficult to fully anticipate precipitation events that will obscure optical data acquisition.

To accurately identify wetland areas in a forested region, such as Northern Minnesota, we found accuracy is improved when incorporating only spring season data for both Level 1 and Level 2 classifications. We conclude that, provided multi-temporal satellite optical, L-band radar (PALSAR), topographic, and soils data are included, identifying wetland areas in this region is more accurate when quad-polarization C-band radar (RADARSAT-2) and higher resolution aerial orthophotos are left out of the random forest model. However, we found that once wetland areas are identified, classifying wetland type is more accurate when C-band radar and broader temporal coverage of optical data are included. These findings are unique because through rigorous testing of different sources of remotely sensed data, a task that has not been done before in this region, we found that different wavelengths of radar data are beneficial for different levels of land cover classification.

The results of this study suggest that wetland mapping in a forested region such as Northern Minnesota can be improved by targeting the selection of important input variables from essential data platforms (such as L-band PALSAR) and by allocating more complete spectral coverage during the spring season. The way forward for further improvements to wetland classification in a forested region may include: analysis and utilization of classification confidence to target areas for future field reference data collection, using additional topographic information derived from light detection and ranging (lidar) such as canopy height and other parameters that relate to vegetation structure (e.g., standard deviation of height and number of returns within a grid cell, intensity), and incorporating spatial context and geometry of features through use of image segmentation and object based image analysis.

Acknowledgments

The authors gratefully acknowledge valuable assistance provided by Marvin Bauer of the University of Minnesota, Bruce Wiley of the US Geological Survey, and Rudi Gens of the Alaska Satellite Facility for their generous offer to review this manuscript. Funding for this research was provided in part by several sources and agencies: the Great Lakes Restoration Initiative and the US Fish & Wildlife Service; the Legislative Citizen Commission on Minnesota Resources, and the Environment & Natural Resources Trust Fund, and the Minnesota Department of Natural Resources; the Canadian Space Agency Science and Operational Applications Research (SOAR) Program and the Canadian Center for Remote Sensing; and the Alaska Satellite Facility and the Japan Aerospace Exploration Agency's (JAXA) Japanese Ministry of Economy, Trade, and Industry (METI).

Conflict of Interest

The authors declare no conflict of interest.

References

1. Vymazal, J. Constructed wetlands for wastewater treatment. *Ecol. Eng.* **2005**, *25*, 475–477.
2. Mitsch, W.J.; Gosselink, J.G. The value of wetlands: Importance of scale and landscape setting. *Ecol. Econ.* **2000**, *35*, 25–33.

3. Lane, C.R.; D'Amico, E. Calculating the ecosystem service of water storage in isolated wetlands using LiDAR in North Central Florida, USA. *Wetlands* **2010**, *30*, 967–977.
4. Deschamps, A.; Greenlee, D.; Pultz, T.J.; Saper, R. Geospatial Data Integration for Applications in Flood Prediction and Management in the Red River Basin. In Proceedings of IEEE International Geosciences and Remote Sensing Symposium and the 24th Canadian Symposium on Remote Sensing, Toronto, Canada, 24–28 June, 2002; pp. 3338–3340.
5. Van der Kamp, G.; Hayashi, M. The groundwater recharge function of small wetlands in the semi-arid northern prairies. *Wetlands* **1998**, *8*, 39–56.
6. Acharya, G.; Barbier, E.B. Valuing groundwater recharge through agricultural production in the Hadejia-Nguru Wetlands in Northern Nigeria. *Agr. Econ.* **2000**, *22*, 247–259.
7. Hamilton, S.K.; Kellndorfer, J.; Lehner, B.; Tobler, M. Remote sensing of floodplain geomorphology as a surrogate for biodiversity in a tropical river system (Madre De Dios, Peru). *Geomorphology* **2007**, *89*, 23–38.
8. Richardson, D.M.; Holmes, P.M.; Esler, K.J.; Galatowitsch, S.M.; Stromberg, J.C.; Kirkman, S.P.; Pyšek, P.; Hobbs, R.J. Riparian vegetation: Degradation, alien plant invasions, and restoration prospects. *Divers. Distrib.* **2007**, *13*, 126–139.
9. Mayer, P.M.; Galatowitsch, S.M. Diatom communities as ecological indicators of recovery in Restored Prairie Wetlands. *Wetlands* **1999**, *19*, 765–774.
10. Solomon, S.; Qin, D.; Manning, M.; Chen, Z.; Marquis, M.; Averyt, K.; Tignor, M.; Miller, H. *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK, 2007.
11. Hearne, R. Evolving water management institutions in the Red River Basin. *Environ. Manage.* **2007**, *40*, 842.
12. Environmental Laboratory. *Corps of Engineers Wetlands Delineation Manual*; Wetland Research Program; US Army Corps of Engineers: Vicksburg, MS, USA, 1987; Y-87-1.
13. Wissinger, S.A. Ecology of Wetland Invertebrates. In *Invertebrates in Freshwater Wetlands of North America: Ecology and Management*; Batzer, D.P., Rader, R.B., Wissinger, S.A., Eds.; Wiley: New York, NY, USA, 1999; pp. 1043–1086.
14. Van der Sande, C.J.; de Jong, S.M.; de Roo, A.P.J. A segmentation and classification approach of IKONOS-2 imagery for land cover mapping to assist flood risk and flood damage assessment. *Int. J. Appl. Earth Obs. Geoinf.* **2003**, *4*, 217–229.
15. Ramsey, E.; Ragoonwala, A.; Middleton, B.; Lu, Z. Satellite optical and radar data used to track wetland forest impact and short-term recovery from Hurricane Katrina. *Wetlands* **2009**, *29*, 66–79.
16. Nel, J.; Roux, D.; Abell, R.; Ashton, P.; Cowling, R.; Higgins, J.; Thieme, M.; Viers, J. Progress and challenges in freshwater conservation planning. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **2009**, *19*, 474–485.
17. Mayer, A.L.; Lopez, R.D. Use of remote sensing to support forest and wetlands policies in the USA. *Remote Sens.* **2011**, *3*, 1211–1233.
18. Dahl, T.E. *Wetlands Losses in the United States 1780's to 1980's*; US Department of the Interior, Fish and Wildlife Service, Office of Biological Services: Washington, DC, USA, 1990; p. 13.
19. Minnesota Department of Natural Resources. Wetlands Status and Trends. Available online: http://www.dnr.state.mn.us/eco/wetlands/wstm_prog.html (accessed on 20 August 2010).

20. Haas, E.M.; Bartholomé, E.; Lambin, E.F.; Vanacker, V. Remotely sensed surface water extent as an indicator of short-term changes in ecohydrological processes in Sub-Saharan Western Africa. *Remote Sens. Environ.* **2011**, *115*, 3436–3445.
21. Goetz, S. Remote sensing of riparian buffers: Past progress and future prospects. *J. Am. Water Resour. Assoc.* **2006**, *42*, 133–143.
22. Dahl, T.E.; Watmough, M.D. Current approaches to wetland status and trends monitoring in Prairie Canada and the Continental United States of America. *Can. J. Remote Sens.* **2007**, *33*, S17–S27.
23. Moore, I.; Grayson, R.; Landson, A. Digital Terrain Modeling: A review of hydrological, geomorphological, and biological applications. *Hydrol. Process.* **1991**, *5*, 330.
24. Sader, S.A.; Ahl, D.; Liou, W. Accuracy of Landsat-TM and GIS rule-based methods for forest wetland classification in Maine. *Remote Sens. Environ.* **1995**, *53*, 133–144.
25. Li, R.; Zhu, A.; Song, X.; Li, B.; Pei, T.; Qin, C. Effects of spatial aggregation of soil spatial information on watershed hydrological modeling. *Hydrol. Process.* **2012**, *26*, 1390–1404.
26. Evans, T.L.; Costa, M.; Telmer, K.; Silva, T.S.F. Using ALOS/PALSAR and RADARSAT-2 to map land cover and seasonal inundation in the Brazilian Pantanal. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2010**, *3*, 560–575.
27. Li, G.; Lu, D.; Moran, E.; Dutra, L.; Batistella, M. A Comparative analysis of ALOS PALSAR L-band and RADARSAT-2 C-band data for land-cover classification in a tropical moist region. *ISPRS J. Photogramm.* **2012**, *70*, 26–38.
28. Pereira, F.R.S.; Kampel, M.; Cunha-Lignon, M. Mapping of mangrove forests on the Southern Coast of São Paulo, Brazil, using Synthetic Aperture Radar data from ALOS/PALSAR. *Remote Sens. Lett.* **2012**, *3*, 567–576.
29. Bwangoy, J.B.; Hansen, M.C.; Roy, D.P.; Grandi, G.D.; Justice, C.O. Wetland mapping in the Congo Basin using optical and radar remotely sensed data and derived topographical indices. *Remote Sens. Environ.* **2010**, *114*, 73–86.
30. Castañeda, C.; Ducrot, D. Land cover mapping of wetland areas in an agricultural landscape using SAR and Landsat imagery. *J. Environ. Manage.* **2009**, *90*, 2270–2277.
31. Li, J.; Chen, W. Clustering Synthetic Aperture Radar (SAR) imagery using an automatic approach. *Can. J. Remote Sens.* **2007**, *33*, 303–311.
32. Bolstad, P.V.; Lillesand, T.M. Rule-based classification models: Flexible integration of satellite imagery and thematic spatial data. *Photogramm. Eng. Remote Sensing* **1992**, *58*, 965–971.
33. Friedl, M.A.; Brodley, C.E. Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* **1997**, *61*, 399–409.
34. Li, J.; Chen, W. A rule-based method for mapping Canada's wetlands using optical, radar and DEM data. *Int. J. Remote Sens.* **2005**, *26*, 5051–5069.
35. Pal, M.; Mather, P. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sens. Environ.* **2003**, *86*, 554–565.
36. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a Random Forest classifier for land-cover classification. *ISPRS J. Photogramm.* **2012**, *67*, 93–104.

37. Naidoo, L.; Cho, M.A.; Mathieu, R.; Asner, G. Classification of savanna tree species, in the Greater Kruger National Park Region, by integrating hyperspectral and lidar data in a Random Forest data mining environment. *ISPRS J. Photogramm.* **2012**, *69*, 167–179.
38. Guo, L.; Chehata, N.; Mallet, C.; Boukir, S. Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests. *ISPRS J. Photogramm.* **2011**, *66*, 56–66.
39. Minnesota Department of Natural Resources. Natural History–Minnesota’s Geology. Available online: <http://www.dnr.state.mn.us/snass/naturalhistory.html> (accessed on 31 March 2011).
40. United States Department of Agriculture. National Agricultural Statistics Service (NASS) Cropland Data Layer. Available online: <http://www.nass.usda.gov/research/Cropland/SARS1a.htm> (accessed on 31 March 2011).
41. Minnesota Department of Administration (AdminMN). Office of Geographic and Demographic Analysis State Demographic Center, 2010 Census: Minnesota City Profiles. Available online: <http://www.demography.state.mn.us/CityProfiles2010/index.html> (accessed on 12 September 2012).
42. National Oceanic and Atmospheric Administration. National Climatic Data Center (NCDC). Normals, Means, and Extremes for Duluth, MN. Available online: http://www.ncdc.noaa.gov/cdo-web/datasets/NORMAL_MLY/stations/GHCND:USW00014913/detail (accessed on 12 August 2012).
43. Minnesota Department of Natural Resources. State Climatology Office, MN Climatology Working Group Historical Climate Data. Available online: <http://climate.umn.edu/doc/historical.htm> (accessed on 9 April 2012).
44. Cowardin, L.; Carter, V.; Golet, F.; LaRoe, E. *Classification of Wetlands and Deepwater Habitats of the United States*; US Department of the Interior, Fish and Wildlife Service: Washington, DC, USA, 1979; pp. 1–79.
45. United States Fish and Wildlife Service. National Wetlands Inventory (NWI). Available online: <http://www.fws.gov/wetlands/> (accessed on 11 October 2011).
46. Foody, G.M. Status of land cover classification accuracy assessment. *Remote Sens. Environ.* **2002**, *80*, 185–201.
47. Yuan, F.; Sawaya, K.E.; Loeffelholz, B.C.; Bauer, M.E. Land cover classification and change analysis of the Twin Cities (Minnesota) metropolitan area by multitemporal Landsat remote sensing. *Remote Sens. Environ.* **2005**, *98*, 317–328.
48. Lunetta, R.; Congalton, R.; Fenstermaker, L.; Jensen, J.; McGwire, K.; Tinny, L. Remote Sensing and Geographic Information System data integration: Error sources and research issues. *Photogramm. Eng. Remote Sensing* **1991**, *57*, 677–687.
49. Janssen, L.L.F.; van der Wel, F.J.M. Accuracy assessment of satellite derived land-cover data: A review. *Photogramm. Eng. Remote Sensing* **1994**, *60*, 419–426.
50. Liaw, A.; Wiener, M. Classification and Regression by randomForest. *R News* **2002**, 18–22.
51. Harken, J.; Sugumaran, R. Classification of Iowa wetlands using an airborne hyperspectral image: A comparison of the Spectral Angle Mapper classifier and an object-oriented approach. *Can. J. Remote Sens.* **2005**, *31*, 167–174.
52. Hogg, A.R.; Todd, K.W. Automated discrimination of upland and wetland using terrain derivatives. *Can. J. Remote Sens.* **2007**, *33*, S68–S83.

53. Breiman, L. Random Forests. *Machine Learn.* **2001**, *45*, 5-32.
54. US Geological Survey. Accuracy of the National Elevation Dataset (NED). Available online: <http://ned.usgs.gov/Ned/accuracy.asp> (accessed on 10 March 2011).
55. Tarboton, D.G.; Bras, R.L.; Rodriguez-Iturbe, I. On the extraction of channel networks from digital elevation data. *Hydrol. Process.* **1991**, *5*, 81–100.
56. Natural Resources Conservation Service, US Department of Agriculture. *Soil Survey Geographic (SSURGO) Database for St. Louis and Carlton County, MN*. Available online: <http://soildatamart.nrcs.usda.gov> (accessed on 1 September 2009).
57. Zhu, Z.; Woodcock, C.E.; Rogan, J.; Kellndorfer, J. Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peri-urban land cover classification using Landsat and SAR data. *Remote Sens. Environ.* **2012**, *117*, 72–82.
58. Baker, C.; Lawrence, R.; Montagne, C.; Patten, D. Mapping wetlands and riparian areas using Landsat ETM+ imagery and decision tree based models. *Wetlands* **2006**, *26*, 465.
59. Shih, S.F.; Jordan, J.D. Landsat mid-infrared data and GIS in regional surface soil-moisture assessment. *J. Am. Water Resour. Assoc.* **1992**, *28*, 713–719.
60. Song, C.; Woodcock, C.E.; Seto, K.C.; Lenney, M.P.; Macomber, S.A. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sens. Environ.* **2001**, *75*, 230–244.
61. Ozesmi, S.L.; Bauer, M.E. Satellite remote sensing of wetlands. *Wetlands Ecol. Manage.* **2002**, *10*, 381–402.
62. Crist, E.P.; Cicone, R.C. A physically-based transformation of Thematic Mapper data—The TM tasseled cap. *IEEE Trans. Geosci. Remote Sens.* **1984**, *22*, 256.
63. Kauth, R.; Thomas, G. The Tasseled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as seen by Landsat. In Proceedings of Symposium on Machine Processing of Remotely Sensed Data, West Lafayette, IN, USA, 29 June–1 July 1976; pp. 4B 41–51.
64. Cohen, W.B.; Spies, T.A. Estimating structural attributes of Douglas-fir/western Hemlock Forest stands from Landsat and SPOT imagery. *Remote Sens. Environ.* **1992**, *41*, 1–17.
65. Dymond, C.C.; Mladenoff, D.J.; Radeloff, V.C. Phenological differences in tasseled cap indices improve deciduous forest classification. *Remote Sens. Environ.* **2002**, *80*, 460–472.
66. Jin, S.; Sader, S.A. Comparison of time series tasseled cap wetness and the Normalized Difference Moisture Index in detecting forest disturbances. *Remote Sens. Environ.* **2005**, *94*, 364–372.
67. Kaya, S. Personal Communication, Environment Canada, Canada Center for Remote Sensing, Ottawa, ON, Canada, 30 June 2010.
68. Bouchemakh, L.; Smara, Y.; Boutarfa, S.; Hamadache, Z. A Comparative Study of Speckle Filtering in Polarimetric Radar SAR Images. In Proceedings of 3rd International Conference on Information and Communication Technologies: From Theory to Applications, Damascus, Syria, 7–11 April 2008; pp. 1–6.
69. Parmuchi, M.G.; Karszenbaum, H.; Kandus, P. Mapping wetlands using multi-temporal RADARSAT-1 data and a decision-based classifier. *Can. J. Remote Sens.* **2002**, *28*, 175–186.
70. van Zyl, J. Unsupervised classification of scattering behavior using radar polarimetry data. *IEEE Trans. Geosci. Remote Sens.* **1989**, *27*, 36–45.

71. Henderson, F.M.; Lewis, A.J. Radar detection of wetland ecosystems: A review. *Int. J. Remote Sens.* **2008**, *29*, 5809–5835.
72. Baghdadi, N.; Bernier, M.; Gauthier, R.; Neeson, I. Evaluation of C-band SAR data for wetlands mapping. *Int. J. Remote Sens.* **2001**, *22*, 71–88.
73. Slatton, K.C.; Crawford, M.M.; Chang, L.D. Modeling temporal variations in multipolarized radar scattering from intertidal coastal wetlands. *ISPRS J. Photogramm.* **2008**, *63*, 559–577.
74. Wang, Y.; Davis, F.W. Decomposition of polarimetric Synthetic Aperture Radar backscatter from upland and flooded forests. *Int. J. Remote Sens.* **1997**, *18*, 1319–1332.
75. Freeman, A.; Durden, S. A three-component scattering model for polarimetric SAR data. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 963–973.
76. Cloude, S.; Pottier, E. An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 68–78.
77. Corcoran, J.; Knight, J.; Brisco, B.; Kaya, S.; Cull, A.; Murnaghan, K. The integration of optical, topographic, and radar data for wetland mapping in Northern Minnesota. *Can. J. Remote Sens.* **2011**, *37*, 564.
78. Sartori, L.R.; Imai, N.N.; Mura, J.C.; Novo, E.M.L.M.; Silva, T.S.F. Mapping Macrophyte Species in the Amazon Floodplain wetlands using fully polarimetric ALOS/PALSAR data. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4717–4728.
79. Brisco, B.; Kapfer, M.; Hirose, T.; Tedford, B.; Liu, J. Evaluation of C-band polarization diversity and polarimetry for wetland mapping. *Can. J. Remote Sens.* **2011**, *37*, 82–92.
80. Congalton, R., Green, K., Eds. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 2nd ed.; CRC Press: Boca Raton, FL, USA; 2008; p. 200.
81. Wright, C.; Gallant, A. Improved wetland remote sensing in Yellowstone National Park using classification trees to combine TM imagery and ancillary environmental data. *Remote Sens. Environ.* **2007**, *107*, 582–605.
82. Jensen, J.R.; Christensen, E.J.; Sharitz, R. Nontidal wetland mapping in South Carolina using airborne multispectral scanner data. *Remote Sens. Environ.* **1984**, *16*, 1–12.
83. Nobrega, R.A.; O'Hara, C.G.; Quintanilha, J.A. An Object-Based Approach to Detect Road Features for Informal Settlements near Sao Paulo, Brazil. In *Object Based Image Analysis*; Blaschke, T., Lang, S. and Hay, G., Eds.; Springer: Heidelberg/Berlin, Germany, 2008; pp. 589–607.
84. Doxani, G.; Siachalou, S.; Tsakiri-Strati, M. An object-oriented approach to urban land cover change detection. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *XXXVII*, 1655–1660.
85. Sesnie, S.E.; Gessler, P.E.; Finegan, B.; Thessler, S. Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments. *Remote Sens. Environ.* **2008**, *112*, 2145–2159.
86. Steininger, M.K. Tropical secondary forest regrowth in the Amazon: Age, area and change estimation with Thematic Mapper data. *Int. J. Remote Sens.* **1996**, *17*, 9–27.
87. Tuomisto, H.; Poulsen, A.D.; Ruokolainen, K.; Moran, R.C.; Quintana, C.; Celi, J.; Cañas, G. Linking floristic patterns with soil heterogeneity and satellite imagery in Ecuadorian Amazonia. *Ecol. Appl.* **2003**, *13*, 352–371.

88. Whitcomb, J.; Moghaddam, M.; McDonald, K.; Podest, E.; Chapman, B. Decadal Change in Northern Wetlands Based on Differential Analysis of JERS and PALSAR Data. In Proceedings of 2009 IEEE International Geoscience and Remote Sensing Symposium, Cape Town, South Africa, 12–17 July 2009; pp. III-951–III-954.
89. Carrão, H.; Gonçalves, P.; Caetano, M. Contribution of multispectral and multitemporal information from MODIS images to land cover classification. *Remote Sens. Environ.* **2008**, *112*, 986–997.
90. Jackson, Q.; Landgrebe, D.A. An adaptive classifier design for high-dimensional data analysis with a limited training data set. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 2664–2679.
91. Ho, T.K.; Basu, M. Complexity measures of supervised classification problems. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 289–300.
92. Townsend, P.A. Mapping seasonal flooding in forested wetlands using multi-temporal RADARSAT SAR. *Photogramm. Eng. Remote Sensing* **2001**, *67*, 857–864.
93. Yang, L.; Huang, C.; Homer, C.; Wylie, B.; Coan, M. An approach for mapping large-area impervious surfaces: Synergistic use of Landsat-7 ETM+ and high spatial resolution imagery. *Can. J. Remote Sens.* **2003**, *29*, 230–240.
94. Baker, C.; Lawrence, R.; Montagne, C.; Patten, D. Change detection of wetland ecosystems using Landsat imagery and change vector analysis. *Wetlands* **2007**, *27*, 610–619.
95. Augusteijn, M.F.; Warrender, C.E. Wetland classification using optical and radar data and neural network classification. *Int. J. Remote Sens.* **1998**, *19*, 1545–1560.
96. Hess, L.L.; Melack, J.M.; Filoso, S.; Wang, Y. Delineation of inundated area and vegetation along the Amazon Floodplain with the SIR-C Synthetic Aperture Radar. *IEEE Trans Geosci. Remote Sens.* **1995**, *33*, 896–904.
97. Hess, L.L.; Melack, J.M.; Novo, E.M.L.M.; Barbosa, C.C.F.; Gastil, M. Dual-season mapping of wetland inundation and vegetation for the Central Amazon Basin. *Remote Sens. Environ.* **2003**, *87*, 404–428.

The Effects of Data Selection and Thematic Detail on the Accuracy of High Spatial Resolution Wetland Classifications

Joseph F. Knight, Bryan P. Tolcser, Jennifer M. Corcoran, and Lian P. Rambi

Abstract

Accurate wetland maps are of critical importance for preserving the ecosystem functions provided by these valuable landscape elements. Though extensive research into wetland mapping methods using remotely sensed data exists, questions remain as to the effects of data type and classification scheme on classification accuracy when high spatial resolution data are used. The goal of this research was to examine the effects on wetland mapping accuracy of varying input datasets and thematic detail in two physiographically different study areas using a decision tree classifier. The results indicate that: topographic data and derivatives significantly increase mapping accuracy over optical imagery alone, the source of the elevation data and the type of topographic derivatives used were not major factors, the inclusion of radar and leaf-off imagery did not improve mapping accuracy, and increasing thematic detail resulted in significantly lower mapping accuracies i.e., particularly in more diverse wetland areas.

Introduction

Wetlands are a valuable natural resource and play a crucial role in the ecological systems of a landscape. Wetlands provide important ecosystem functions such as maintaining water quality by filtering nutrients and pollutants, storing floodwater and mitigating its effects, and providing habitat for a variety of wildlife adapted to saturated environments. Wetlands also play a role in the global carbon cycle, acting as both carbon sources and sinks (Keddy, 2000; Mitsch and Gosselink, 2000).

Wetland loss has occurred at a rapid rate in the United States. In the years between European settlement and the 1980s, the 48 conterminous states lost an estimated 53 percent of wetland acreage due to human activities such as agriculture, urbanization, and pollution (Dahl, 1990). In the state of Minnesota, United States, over 50 percent of the estimated pre-settlement 3.6 million ha of wetlands have been lost statewide. However, the degree of wetland loss is greatest, over 80 percent, in southern and western Minnesota where wetlands were drained primarily for agriculture. Urbanization has caused comparatively smaller wetland area losses, but has

significantly altered wetlands' physical, biological, and chemical properties (Johnston, 1989). The loss of wetlands continues, but some studies suggest that wetland loss is slowing due to regulatory controls (Dahl and Johnson, 1991). Despite the critical importance of accurate mapping of the spatial distributions of wetlands for making policy decisions related to preservation of existing wetlands (Baker *et al.*, 2006), the National Wetlands Inventory (NWI) in Minnesota is as much as 38 years out of date in some areas (MNGeo, 2012).

Accurate mapping of wetlands can be achieved through a variety of approaches ranging from field investigation to remote wetland assessment. Due to the high costs of performing field wetland mapping, remote sensing-based approaches have been used for several decades (Cowardin and Myers, 1974). Numerous studies have examined remote sensing based data sources and approaches for wetland mapping.

Frequently examined methods include aerial photograph interpretation and satellite image analysis of both single and multi-date optical satellite imagery, in which optical properties (e.g., reflectance) of wetland vegetation and land forms are assessed (Baker *et al.*, 2006; Harvey *et al.*, 2001; Hodgson *et al.*, 1987; Lunetta and Balogh, 1999; Ozesmi and Bauer, 2002; Pope, 1994; Sader *et al.*, 1995; Tiner, 1990; Townsend and Walsh, 2001; Wang *et al.*, 1998a; Wright and Gallant, 2007). A notable example of a project incorporating these techniques is the National Oceanic and Atmospheric Administration's (NOAA) Coastal Change Analysis Program (C-CAP). C-CAP provides periodic land-use/land-cover classifications of areas near coastlines and the Great Lakes, with the goal of studying change, including in wetlands, in those areas. A somewhat less studied optical method involves the use of hyperspectral imagery to map wetlands based on fine details in vegetation spectral response. Though hyperspectral imagery can be used to derive accurate wetland maps, the data acquisition, storage, and processing requirements are greater than those of multispectral imagery (Becker *et al.*, 2005 and 2007; Hirano *et al.*, 2003; Jollineau and Howarth, 2008; Neuenschwander *et al.*, 1998; Wang *et al.*, 1998b). Though useful, hyperspectral data were not available for inclusion in this project. In recent years, high spatial resolution satellite and aerial imagery have been assessed for wetland mapping potential. Maxa and Bolstad (2009) used Ikonos imagery and lidar data to map northern wetlands, which outperformed an

Joseph F. Knight, Jennifer M. Corcoran, and Lian P. Rambi are with the Department of Forest Resources, University of Minnesota, 1530 Cleveland Ave N., Saint Paul, MN 55108 (jknight@umn.edu).

Bryan P. Tolcser is with Short, Eliot, Hendrickson, Inc., 3535 Vadnais Center Dr., St. Paul, MN 55110, and formerly with the Department of Forest Resources, University of Minnesota, Saint Paul, MN 55108.

Photogrammetric Engineering & Remote Sensing
Vol. 79, No. 7, July 2013, pp. 613–623.

0099-1112/13/7907-613/\$3.00/0
© 2013 American Society for Photogrammetry
and Remote Sensing

existing wetland inventory for the State of Wisconsin. Laba *et al.* (2008) used QuickBird imagery to map invasive wetland species. Bowen *et al.* (2010) used high-resolution aerial images and ancillary data to map playa wetlands in Kansas. Halabisky *et al.* (2011) used a combination of high-resolution imagery and object-based classification to map semi-arid wetlands. Many other studies have examined issues such as wetland vegetation analysis and coastal wetland mapping with high spatial resolution imagery (Dechka *et al.*, 2002; Ramsey and Laine, 1997; Wei and Chow-Fraser, 2011).

Radar imagery has been shown to have utility for wetland remote sensing. Unlike optical sensors, radar sensors operate in the microwave portion of the electromagnetic spectrum and are insensitive to most atmospheric and low light conditions. Radar backscatter is sensitive to soil and vegetation moisture properties and can, to some degree, penetrate the forest canopy and provide sub-canopy vegetation and soil saturation information (Whitcomb *et al.*, 2007). Because radar is sensitive to moisture, techniques using interferometric analysis of radar data have been shown to identify changes in water levels to within a centimeter (Wdowinski, 2008). Numerous researchers report that careful selection of the timing of image acquisition with respect to soil moisture levels, radar band(s) to be used, and the combination of radar and optical imagery results in higher wetlands mapping accuracies (Costa *et al.*, 2006; Dobson *et al.*, 1995; Henderson and Lewis, 2008; Hess *et al.*, 1990; Hess *et al.*, 1995; Hess *et al.*, 2003; Kasischke, 1997; Lozano-Garcia and Hoffer, 1993; Ramsey, 1998; Rosenqvist *et al.*, 2004; Wang *et al.*, 1995). Others caution that radar imagery may be only situationally useful due to the effects of speckle and forest canopy interference on classification results (Corcoran *et al.*, 2012; Li and Chen, 2005).

Non-image geospatial data sets may provide valuable information for wetland mapping. Digital Elevation Models (DEM) are commonly used, both for elevation information and a number of topographic derivatives including slope, flow accumulation, and probability of soil wetness. Studies on the use of DEMs for land-cover mapping include the effects of DEM resolution on wetland mapping accuracy (Creed *et al.*, 2003), determination of soil characteristics (NRCS, 2010; Thompson *et al.*, 2001), use of a depth-to-water index for modeling of wet areas (Murphy *et al.*, 2007), and the suitability of several DEM derivatives for identification of wetlands (Hogg and Todd, 2007).

A large number of studies have focused on the effects of classification algorithm choice on wetland mapping accuracy, with rule-based and decision tree algorithms emerging as strong alternatives to traditional approaches such as maximum likelihood estimation (Bolstad and Lillesand, 1992; Rodriguez-Galiano, 2012). Hogg and Todd (2007) compared several statistical methods and found the Classification and Regression Tree (CART) algorithm to result in the highest accuracy. Baker *et al.* (2006) compared the accuracy of Classification Tree Analysis (CTA) and Stochastic Gradient Boosting (SGB) classifiers and reported that the SGB method performed best. Liu *et al.* (2008) used a decision tree approach to successfully map mangrove forests. Rover *et al.* (2011) determined hydrologic function of wetlands using a decision tree classifier. Li and Chen (2005), Parmuchi *et al.* (2002), and Phillips *et al.* (2005) developed rule-based wetland mapping methods for combining inputs from a variety of geospatial sources. In terms of more general (i.e., not wetland specific) land-cover/land-use mapping, a notable example is the 2001 National Land Cover Database (NLCD), which was created using a decision tree classifier with inputs composed of several dates of imagery, topography and topographic derivatives, and other ancillary data sets such as impervious surface maps (Homer *et al.*, 2004).

Despite the aforementioned extensive research in wetland mapping, many questions remain i.e., particularly with respect to studies using high spatial resolution imagery. An important question is which data types, among the many that are available to geospatial researchers, should one prioritize for inclusion in high spatial resolution mapping projects. In the research described here, we examined the effects of data type selection on wetland mapping accuracy using multiple classification schemes in two physiographically different study areas. The specific goals of this study were: (a) to examine the advantages and disadvantages of using several input geospatial datasets for mapping wetlands, (b) to describe the suitability of geospatial data for classifying wetlands according to three schemes (wetland/non-wetland, Cowardin class, and MNDNR), and (c) To compare classification accuracies of wetlands in two very different physiographic regions.

Study Areas

Two study areas in Minnesota were selected for this research, one located in the Minneapolis-St. Paul metropolitan area and one located in the northeast forested region (Figure 1). These areas were selected because they represent a wide range of wetland types and because geospatial datasets and field reference data (described below) were available. The metro study area encompassed the limits of the City of Chanhassen, Minnesota, a southwestern suburb of Minneapolis with an area of approximately 60 km². Land-use within the city is primarily medium density residential with some areas of industrial and dedicated open space. Wetlands, lakes, ponds, and rivers account for approximately 27 percent of the city's surface area (City of Chanhassen, 2006).

The Fond du Lac Reservation (FDL), located northwest of the City of Cloquet, Minnesota, is part of the boreal forest biome. FDL has an area of approximately 390 km². The land-cover is dominated by both deciduous and evergreen forests and low density residential. Wetlands and water bodies account for approximately 38 percent of FDL's surface area. The FDL area experienced dryer than normal weather conditions during 2009 when some of the FDL imagery used in this study were acquired. Drought conditions persisted throughout the spring and summer of 2009, which may have affected the study results with respect to the measured utility of those images.

Methods

Classification Schemes

Three classification schemes were used in this study: (a) A simple wetland versus upland discrimination, (b) Wetlands classified to the Cowardin class level (Cowardin *et al.*, 1974; Table 1), and (c) A simplified plant community classification (hereafter termed "MNDNR"). The MNDNR scheme was developed by the Minnesota Department of Natural Resources (DNR) and is based on Eggers and Reed (1997), with modifications to make the scheme more appropriate for remote sensing-based mapping of wetlands. The full classification scheme including class definitions can be found in Kloiber and MacLeod, 2011. The scheme is in official use in Minnesota within the wetland mapping group of the DNR. In addition, it is dissimilar to the Cowardin scheme; thus it provides a useful and applicable base for comparison of the various input data types used.

Tables 2 and 3 show the wetland composition in Chanhassen and FDL by Cowardin and MNDNR classes, respectively. Wetland data for the City of Chanhassen were collected during the 2006 Surface Water Management Plan (SWMP) update (described below); data for FDL were derived from a

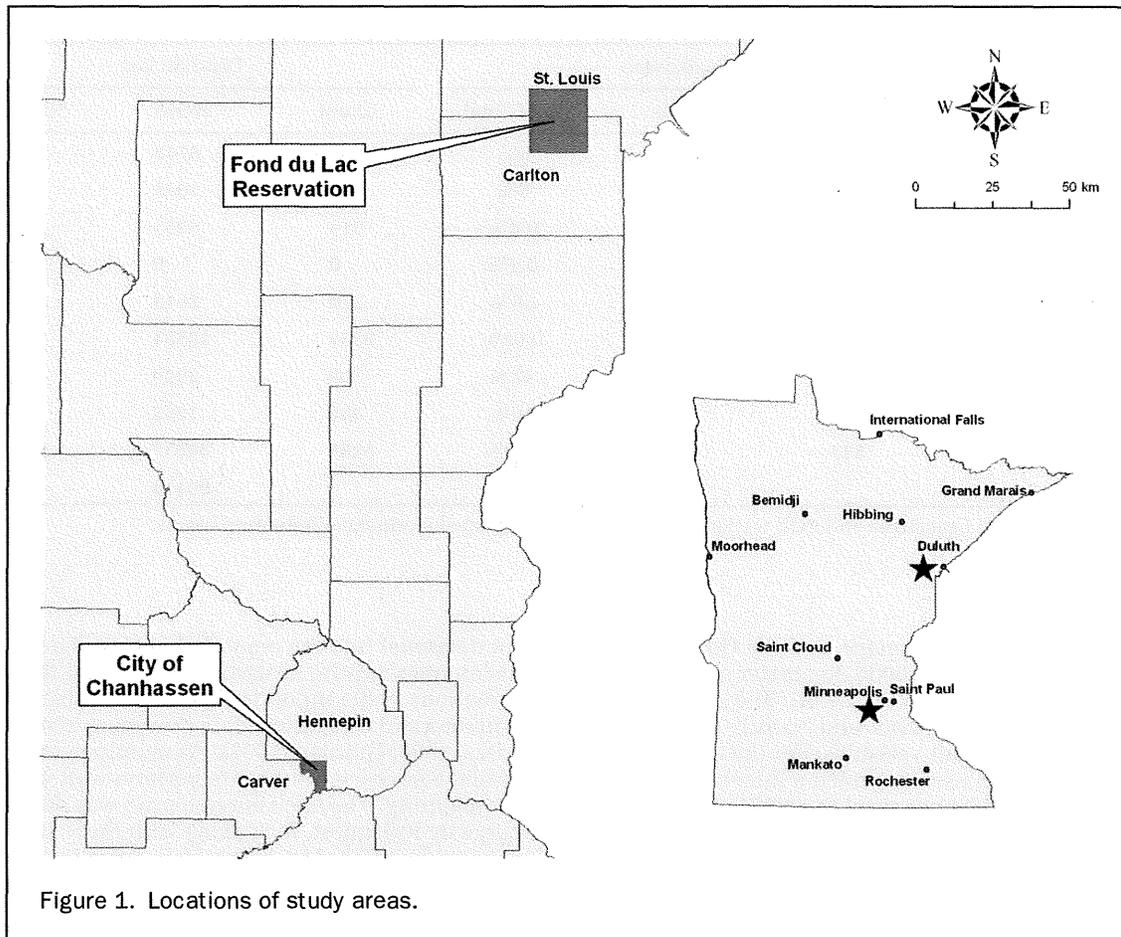


Figure 1. Locations of study areas.

TABLE 1. COWARDIN WETLAND CLASSES

Cowardin Code ¹	Description
PEM	Palustrine Emergent
PSS	Palustrine Scrub Shrub
PFO	Palustrine Forested
L	Lacustrine
PUB	Palustrine Unconsolidated Bottom

¹ Cowardin codes are taken from Cowardin *et al.* (1974).

2008 one meter aerial image-based wetland inventory provided by FDL. Note that this latter FDL inventory was used for qualitative purposes only, as it was found to be of insufficient quality for either classifier training or results validation.

Data Used

The data used in this research varied depending on the study area. Data common to both study areas included: National Agriculture Imagery Program (NAIP) images (acquired in summer of 2008, one meter spatial resolution, color infra-red, 5 m horizontal accuracy), US Department of Agriculture (USDA)

TABLE 2. SUMMARY OF WETLAND TYPES BY COWARDIN CLASS

Class	Chanhassen			Fond du Lac		
	Count	Acres	% of Total	Count	Acres	% of Total
PEM	305	2304	58.4%	826	4311	11.8%
PFO	40	19	0.5%	1797	15776	43.1%
PSS	3	1	0.02%	2334	13584	37.1%
W ¹	189	1621	41.1%	309	2949	8.1%
Total Features	537	3944	100.0%	5266	36619	100.0%
Study Area		14515	27.2%		96119	38.1%

¹ Water class included Lacustrine and PUB wetlands as well as non-vegetated stormwater detention basins.

TABLE 3. SUMMARY OF WETLAND TYPES BY MNDNR CLASSIFICATION SCHEME

Class	Chanhassen			Fond du Lac		
	Count	Acres	% of Total	Count	Acres	% of Total
Coniferous Wetland	0	0	0%	883	9743	27%
Deep Marsh	52	228	6%	148	1045	3%
Hardwood Wetland	47	25	0.6%	914	6033	17%
Seasonally Flooded	10	5	0.1%	0	0	0%
Shallow Marsh	132	1410	36%	270	2013	6%
Shrub Wetland	3	1	0.02%	2334	13584	37%
Water ¹	191	1635	42%	309	2949	8%
Wet Meadow	102	641	16%	408	1253	3%
Total Features	537	3944	100%	5266	36619	100.0%
Study Area		14515	27%		96119	38%

¹ Water class included Lacustrine and PUB wetlands as well as non-vegetated stormwater detention basins.

Soil Survey Geographic (SSURGO) maps, and the US National Elevation Data (NED; acquired in 2008, ten meter spatial resolution, vertical accuracy estimated at 2.4 m). The extracted SSURGO drainage class (e.g., “hydric” and “poorly drained”) served as the soil-related input variables. For the Chanhassen study area, additional data used included a lidar-derived digital elevation model (lidar acquired in spring of 2006, three meter spatial resolution, 15 cm vertical resolution). For the FDL area, additional data used included: Radarsat-2 C-band radar imagery (acquired 15 June 2009, “Fine” 4.7 m spatial resolution, quad-polarization, backscattering coefficients scaled in decibels), and leaf-off digital aerial images (acquired mid-May to early-June 2009, 0.5 m spatial resolution, color infra-red, 3.5 m horizontal accuracy). Though C-band radar imagery is not optimal for wetland mapping under forest canopy, we included such imagery for completeness and because it was available at no cost. We recognize that additional polarimetric processing of the radar imagery may have yielded improved results, but software to perform that processing was not available. All data were projected to the Universal Transverse Mercator coordinate system, Zone 15, NAD83 datum. Both the NED and lidar-derived DEMs were hydrologically corrected before use. Topographic derivatives were computed for the elevation datasets: slope, Compound Topographic Index (CTI), and Curvature. The CTI is a well known measure of the likely wetness of an area. It is computed with the formula $CTI = \ln(A_s / \tan(B))$, where A_s is the upstream contributing area to the pixel and B is the slope in radians (Gessler *et al.*, 1995). The d-infinity flow model was used to create the CTI. Curvature indicates local convexity or concavity for each pixel, with positive values indicating concavity, zero values indicating linearity, and negative values indicating convexity (Parsons, 1979).

Wetland Classifications

Since the focus of this research was to test the effects on accuracy of data type selection and classification scheme rather than classification method, a common classification approach was used throughout the various trials. We chose to use a decision tree classifier in this research. A decision tree is a supervised algorithm that produces a classification by developing a set of decision points, or nodes, that are created by identifying diagnostic features in the training data. The resulting “tree” of nodes is then used to partition the

input dataset(s) into the requested output classes (“leaves”). Decision trees require no assumptions about the underlying distributions of the input datasets and are able to use both continuous and categorical data (Breiman, 2001; Friedl and Brodley, 1996; Quinlan, 1993). As mentioned above, these algorithms have been shown to perform well in land use/cover mapping. The decision tree used was the See5 software package by Rulequest, Inc., along with the NLCD Mapping Tool developed by MDA, Ltd. Three steps were involved in the decision tree classification: data sampling, data mining/tree creation, and classification.

The first step, data sampling, involved assembling training data points and collecting values from the input data layers. Training data for the Chanhassen classification were derived from the city’s 2006 SWMP data (described in *Accuracy Assessment* below). Before the SWMP data were used for training, the polygons were edited to correct for changes resulting from a 2008 highway construction project. Five thousand simple random points were generated throughout the study area in each of wetland and upland areas, for a total of 10,000 training points.

Training data for the FDL location were created from manual interpretation of the 2008 NAIP and 2009 leaf-off imagery. A Minnesota Certified Wetland Delineator manually delineated 140 training polygons. The polygons ranged in size from approximately 50 image pixels to 300 pixels. Sample pixels representing the range of wetland and upland types present in the study area were selected from the training polygons using a simple random sampling method. This procedure resulted in a total of 5,412 wetland and upland training samples.

The NLCD Sampling Tool ver. 2.0, a utility included in the NLCD Mapping Tool, was used to create an input data file for use in See5. The NLCD Sampling Tool extracted values from each input dataset at each sampling point. The utility generated a tabular file which contained a row for each sampling point with comma separated values for each input data layer.

In the second step, data mining/tree creation, the See5 software package was used to create decision trees derived from the data tables created with the NLCD Sampling Tool. The boost, fuzzy thresholds, and global pruning options were enabled for classifier construction. The boost option caused See5 to create decision trees using a recursive algorithm, in which results from previous trees were weighted more heavily in subsequent trees. The fuzzy threshold option established upper and lower bounds for each independent variable rather

than using hard values. When constructing the decision tree, a value between the upper and lower bound was assigned a class by See5. The global pruning option allowed the See5 algorithm to remove (prune) parts of the trees exhibiting relatively high error rates. The results of the data mining processes were decision trees that were used to produce the various classification trials (Table 4). Decision trees were constructed for use in wetland versus upland classification, wetland classification to the Cowardin class level, and wetland classification using the MNDNR scheme. Additional decision trees were constructed to evaluate the mapping accuracy differences resulting from varying the type and number of input data sets (described below).

The final step was to produce the classifications. The output classes were those drawn from the training data. The area classified was the geometric intersection of all input datasets. The classifications were performed using the See5 Classifier Tool, a part of the NLCD Mapping Tool. The See5 output included an internal validation done using the input sampling points as a measure of error inherent in the resultant decision tree (i.e., contingency accuracy). Cross-validation was enabled to provide validation estimates using "out of bag" sampling.

Several classifications were performed to determine the effects of various input datasets on classification accuracy (Table 4). The first classification was a wetland versus upland discrimination using all of the available data types (e.g., imagery, topography, etc.). Then, areas identified as wetlands were classified at higher thematic detail according to the Cowardin and MNDNR schemes. These wetland type classifications were created using a variety of inputs, including: the available data of all types, the available data without topography, and the NAIP/optical imagery alone. Additional classifications in Chanhassen were performed to compare differences between high spatial resolution (3m) and lower resolution (10m) topography data as well as differences between the CTI and Curvature topographic derivatives. Additional classifications in Fond du Lac were performed to determine the effects of including C-band radar data and leaf-off imagery on the classification accuracy.

Because some areas identified as wetland in the initial wetland/upland discrimination may not have been wetlands, upland was included as an output class in the wetland type classifications. A small percentage of pixels initially classified as wetland in the wetland/upland classification were

TABLE 4. DATA USED SCENARIOS FOR WETLAND CLASSIFICATIONS

Data Layer	Classification Scenario													
	Chanhassen							Fond du Lac						
	All Data	Hi-Res - CTI	Hi-Res - Curve	NED Topo	NED - CTI	NED - Curve	No Topo	NAIP Only	All Data	No Radar	No Leaf Off	No Topo	Optical Only	NAIP Only
Imagery														
2008 NAIP Leaf On Imagery (R,G,B,IR)	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2009 Spring Leaf Off Imagery (R,G,B,IR)									X	X		X	X	
RaDAR Imagery (Quad Pol)									X		X	X		
Imagery - Derived														
2008 NAIP NDVI	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2009 Leaf Off NDVI									X	X		X	X	
NDVI Difference									X	X		X	X	
Topography														
10m NED DEM				X	X	X			X	X	X			
2-ft Hi-Res LiDAR Based DEM	X	X	X											
Topography Derivations														
CTI (3m LiDAR derived)	X	X												
CTI (10m NED derived)				X	X				X	X	X			
CTI (24m LiDAR degrade derived)	X	X												
Slope (3m LiDAR derived)	X	X	X											
Slope (10m NED derived)				X	X	X			X	X	X			
Curvature (3m LiDAR derived)	X		X											
Curvature (10m NED derived)				X		X			X	X	X			
Other Data														
SSURGO (Drainage Class)	X	X	X	X	X		X		X	X	X	X		

subsequently classified as upland in the wetland type classifications and were maintained as such in the accuracy assessment. Areas incorrectly identified as upland were included in the error matrices, but the accuracy of the upland class within the wetland type classifications was not assessed. This approach was necessary so that the overall accuracy estimates of the various classification trials would better reflect the performance of the wetland type classifications rather than wetland/upland discrimination.

Accuracy Assessment

The accuracy of each classification was assessed by comparison with ground and image-based reference data. Error matrices were calculated using the methods described in Congalton and Green (1999). For the Chanhassen pilot area, the city's SWMP was used as the reference data source. In the SWMP, uplands, wetlands, and water features throughout the city were identified and observed in the field. Mapping for all areas within city was completed using a combination of field GPS delineation and image interpretation. A Minnesota Certified Wetland Delineator validated all polygons. Further methodology is described in City of Chanhassen (2006). To create the reference data for this study, a random sample of 10,000 points was generated throughout the city. *This sample was independent of the 10,000 samples used in training the classifiers.* Wetland classes were extracted from the SWMP for each point. Wetland polygons in the SWMP with two or more wetland types noted were considered to be the dominant wetland type. A simple random sampling scheme resulted in 7,343 upland points and 2,657 wetland points. Wet features in Chanhassen consisted of water, forested wetlands, and emergent wetlands, as listed in Tables 1 and 2. Wetland type classification by Cowardin class included water (L, PUB, PAB), emergent (PEM), scrub/shrub (PSS), and forested (PFO) wetlands. Scrub/shrub comprised a very small area of the wetland cover in the study area and contained only five field validation points; therefore that class was not included in the accuracy assessment. The MNDNR classification included water, wet meadow, shallow marsh, deep marsh, shrub wetland, seasonally flooded, and hardwood wetland classes. Seasonally flooded and shrub wetlands each had fewer than ten field validation points and were removed from the accuracy assessment to maintain statistical validity.

For FDL, field reference data were collected 13-17 July 2009 by a team from the University of Minnesota, which was led by a Minnesota Certified Wetland Delineator. A stratified random sampling scheme based on the existing NWI classes was used within wetland types to generate a sample of 250 wetland sites. An additional 150 sites were randomly generated within uplands. Data collected at each site included land-cover/land-use type, vegetative species present, crown closure percent, neighboring land-cover/land-use, panoramic and canopy photographs, and general notes about the site. A total of 195 points was collected during one week of field work. These points were used as reference data for the accuracy assessment for the FDL study area. *These reference data were independent of the training polygons used in decision tree development.* Wet features in Fond du Lac consisted primarily of forested and scrub/shrub type wetlands. Wetland type classification by Cowardin class included water (L, PUB, PAB), emergent (PEM), scrub/shrub (PSS), and forested (PFO) wetlands. Most of the field validation points were scrub/shrub and forested wetlands, so emergent wetlands were not included in the accuracy assessment due to insufficient validation points to create statistically significant error estimates for that class. Wetland type classification by MNDNR included water, wet meadow, shallow marsh, deep marsh, shrub wetland, hardwood wetland, and coniferous wetland.

The wet meadow, shallow marsh, and deep marsh classes each had fewer than ten field validation points and so were not included in the accuracy assessment.

The See5 software's internal cross-validation process was employed to provide a measure of the agreement of the classifications' outputs with the training data. With cross-validation, See5 performed a user-determined number of iterations of decision tree construction (i.e., folds) with a subset of the total training points and used the remainder of the points for validation. In this study, a 10-fold cross-validation was used, in which 10 percent of the training points were randomly set aside and the decision tree was constructed using the other 90 percent of points. Repeated iterations were performed with different subsets of points set aside such that after ten iterations each point had been used once in cross-validation.

Results

The results of this research are summarized in Tables 5 through 14. Full error matrices are presented for the "All Data" scenario for each of the different classification schemes and study areas. Due to space constraints, only the overall percent accuracy is given for the many other trials described in Table 4. Unless otherwise noted, all of the accuracy assessment results in each study area were compared using the kappa-based z-statistic tests described in Congalton and Green (1999) and were found to be significantly different at an alpha level of 0.05. Although there has been controversy surrounding the kappa coefficient (e.g., Foody, 1992; Pontius and Millones, 2011; Stehman and Czaplewski, 1998), we believe that kappa retains value in thematic accuracy assessment i.e., especially for comparison of error matrices.

Table 5 shows the results of the wetland/upland discrimination in the Chanhassen study area using the All Data scenario. The overall accuracy was 93 percent, with low errors of omission and commission. When the decision tree classifier was trained to identify Cowardin classes rather than the simpler wetland/upland determination (Table 6), overall performance remained strong at 86 percent; however the user's and producer's accuracies of the PFO class were both relatively low. The MNDNR error matrix is shown in Table 7. The overall accuracy of this trial was lower than the preceding Cowardin class mapping, at 77 percent. Two of the classes exhibited very high errors of omission and/or commission: hardwood wetland and deep marsh. Tables 8 and 9 show the See5 cross-validation (X-Val) and accuracy assessment (Assess) results for the other trials conducted in the Chanhassen area. The combination of high-resolution optical imagery, SSURGO, topography, and a topographical derivative performed significantly better (as measured by the z-test) than the trials without topographical information; and, much better than the optical imagery and SSURGO data alone. However, the differences between the lidar versus NED trials and the CTI versus Curvature trials were not statistically significant. Thus, the source of the topographic data and the choice of topographic derivative were not important influences on the overall accuracies of the trials.

TABLE 5. CHANHASSEN ALL DATA SCENARIO – WETLAND/UPLAND ERROR MATRIX

		Reference Data			
		Upland	Wetland	Map Total	User's Acc.
Map Data	Upland	6945	296	7241	96
	Wetland	398	2361	2759	86
	Ref. Total	7343	2657	10000	
	Prod. Acc.	95	89		93

TABLE 6. CHANHASSEN ALL DATA SCENARIO – COWARDIN CLASS ERROR MATRIX

		Reference Data					Map Tot	User.Acc.
		UPL	PEM	W	PFO	PSS		
Map Data	UPL	0	230	34	11	0	276	0
	PEM	0	1262	53	8	0	1323	95
	W	0	41	1013	0	0	1054	96
	PFO	0	1	0	2	0	3	67
	PSS	0	0	0	0	0	0	0
	Ref. Total	0	1534	1101	21	0	2656*	
Prod.Acc.		0	82	92	10	0		86

*One wetland sample was removed because the Cowardin class reference label was incorrect.

TABLE 7. CHANHASSEN ALL DATA SCENARIO – MNDNR ERROR MATRIX

		Reference Data							Map Total	User Acc	
		Upl	Shall. Mrsh.	Water	Wet Mead	Deep Marsh	Hdwd Wet	Seas. Flood			Shrub Wet.
Map Data	Upland	0	126	42	104	37	12	0	0	321	-
	Shal Marsh	0	743	23	64	30	2	0	0	862	86
	Water	0	22	1005	11	39	0	0	0	1077	93
	Wet Mead	0	37	14	251	14	3	0	0	319	79
	Deep Mrsh	0	9	17	7	31	2	0	0	68	46
	Hdwd Wet	0	2	0	2	0	2	0	0	6	33
	Seas Flood	0	0	0	1	0	0	0	0	1	-
	Shrub Wet	0	0	0	0	1	0	0	0	1	-
	Ref. Total	0	939	1101	442	152	21	0	0	2655*	
	Prod.Acc.		-	79	91	57	20	10	-	-	

*Two wetland samples were removed because the MNDNR class reference labels were incorrect.

TABLE 8. CHANHASSEN TRIALS - SEE5 CROSS VALIDATION VERSUS ACCURACY ASSESSMENT, PART 1; ACCURACY ASSESSMENT VALUES ("ASSESS") MARKED WITH † WERE NOT SIGNIFICANTLY DIFFERENT FROM EACH OTHER WITHIN CLASSIFICATION SCHEMES

Classification Scheme	All Data		NED Topo		No Topo		NAIP Only	
	X-Val	Assess	X-Val	Assess	X-Val	Assess	X-Val	Assess
Wetland/Upland	90	93 [†]	86	92 [†]	82	89	69	78
Cowardin Class	85	86 [†]	82	84 [†]	77	80	64	55
MNDNR	81	77 [†]	77	76 [†]	67	61	60	43

TABLE 9. CHANHASSEN TRIALS - SEE5 CROSS VALIDATION VERSUS ACCURACY ASSESSMENT (CONTINUED)

Classification Scheme	Hi-Res Topo Curvature Only		NED Topo Curvature Only		Hi-Res Topo CTI Only		NED Topo CTI Only	
	X-Val	Assess	X-Val	Assess	X-Val	Assess	X-Val	Assess
Wetland/Upland	86	93 [†]	85	91 [†]	90	92 [†]	86	91 [†]
Cowardin Class	81	84 [†]	81	84 [†]	85	84 [†]	82	85 [†]
MNDNR	77	76 [†]	76	77 [†]	79	75 [†]	77	77 [†]

Tables 10 through 14 present the results of the FDL study area trials. Table 10 shows the wetland/upland discrimination using the All Data scenario. Both the overall and most of the user's/producer's accuracies were lower than in the Chanhasen area. We attribute these differences to complexity introduced by the greater variety of wetland types and the large extent of forest canopy in FDL. Tables 11 and 12 show the Cowardin and MNDNR classification results. The class-specific accuracy estimates were also generally lower than in Chanhasen. In both schemes, upland class commission errors had greater impacts on the overall accuracy than in Chanhasen. Tables 13 and 14 show cross-validation and

overall percent accuracy estimates for the FDL trials. The FDL study area trials included input data types that were not available in Chanhasen, such as spring leaf-off aerial images and Radarsat-2 imagery. As was the case in Chanhasen, the exclusion of topography data and its derivatives had the largest negative impact on mapping accuracy. An unexpected result was that the inclusion of leaf-off and radar imagery had no statistically significant impact on accuracy in any of the classification schemes.

Discussion and Conclusions

The overarching goal of this research was to examine the wetland mapping accuracy effects of varying input data types and classification schemes using high spatial resolution datasets. To that end, several classification trials were performed in two physiographically different study areas. In aggregate, the results from both areas broadly suggest that more and varied input data can improve mapping accuracy, but there were unexpected findings along with those more typically seen in the relevant literature.

First, as expected, including topography information significantly improved classification accuracy across the different classification schemes and the input data trials in both study areas. Since the ability to at least temporarily hold water is a defining characteristic of wetlands, topographic position is important to discriminate both wetland versus

TABLE 10. FDL ALL DATA SCENARIO - WETLAND/UPLAND ERROR MATRIX

		Reference Data			User's Acc.
		Upland	Wetland	Map Total	
Map Data	Upland	27	37	64	42
	Wetland	4	127	131	97
	Ref. Total	31	164	195	
	Prod. Acc.	87	77		79

TABLE 11. FDL ALL DATA SCENARIO - COWARDIN CLASS ERROR MATRIX

		Reference Data					Map Tot	User's Acc.
		UPL	PSS	PFO	PEM	W		
Map Data	UPL	0	14	20	0	1	35	-
	PSS	0	31	14	0	0	45	69
	PFO	0	8	47	0	0	55	85
	PEM	0	8	1	0	0	9	-
	W	0	0	0	0	14	14	100
	Ref. Total	0	61	82	0	15	158*	
	Prod. Acc.	-	51	57	-	93		58

*Six wetland samples were removed because the Cowardin class reference labels were incorrect.

TABLE 12. FDL ALL DATA SCENARIO - MNDNR ERROR MATRIX

		Reference Data							Map Tot	User's Acc.	
		Upl	Shrub Wet.	Conif. Wet.	Shal. Marsh	Water	Hdwd Wet	Deep Marsh			Wet Mead
Map Data	Upland	0	13	7	0	1	12	0	0	33	-
	Shrb. Wet	0	34	11	0	0	6	0	0	51	67
	Conif. Wet	0	3	25	0	0	2	0	0	30	83
	Shl. Mrsh	0	8	1	0	0	1	0	0	10	-
	Water	0	0	0	0	14	0	0	0	14	100
	Hdwd W.	0	2	2	0	0	15	0	0	19	79
	Dp. Mrsh	0	0	0	0	0	0	0	0	0	-
	W. Mead.	0	1	0	0	0	0	0	0	1	-
	Ref. Total	0	61	46	0	15	36	0	0	158*	
	Prod. Acc.	-	56	54	-	93	42	-	-		55

*Six wetland samples were removed because the MNDNR reference class labels were incorrect.

TABLE 13. FDL TRIALS - SEE5 CROSS VALIDATION VERSUS ACCURACY ASSESSMENT, PART 1; ACCURACY ASSESSMENT VALUES ("Assess") MARKED WITH † WERE NOT SIGNIFICANTLY DIFFERENT FROM EACH OTHER WITHIN CLASSIFICATION SCHEMES

Classification Scheme	All Data		No Leaf Off		No Radar		No Topo	
	X-Val	Assess	X-Val	Assess	X-Val	Assess	X-Val	Assess
Wetland/Upland	96	79 [†]	96	77 [†]	96	78 [†]	92	71
Cowardin Class	93	58 [†]	93	60 [†]	93	54	87	44
MNDNR	93	56 [†]	92	58 [†]	93	53	86	43

TABLE 14. FDL TRIALS - SEE5 CROSS VALIDATION VERSUS ACCURACY ASSESSMENT (CONTINUED)

Classification Scheme	Optical Only		NAIP Only	
	X-Val	Assess	X-Val	Assess
Wetland/Upland	84	50	76	42
Cowardin Class	80	29	73	26
MNDNR	78	32	71	23

upland and different wetland types. Two unexpected results were that the source of the elevation information, and the choice of topographic derivative did not have statistically significant influences on accuracy. The NED and lidar-derived ("Hi-Res") topography data produced similar results, which suggest that in areas like Chanhassen, with relatively low wetland and topographic diversities, the coarser resolution NED (10m) may be sufficient. The comparisons of the CTI with Curvature also indicated no significant differences in accuracy, no matter whether the topographic derivatives were computed from the NED or lidar-derived DEMs, or which classification scheme or study area was examined. The CTI is a well known method of determining the wetness potential of an area, but it requires significant computational resources to create for large areas at high spatial resolution. It also requires a hydrologically corrected DEM, while Curvature does not. These results indicate that Curvature may represent a suitable alternative in some situations.

Second, the choice of classification scheme had a significant effect on classification accuracy. The relatively simple wetland versus upland discrimination unsurprisingly resulted in the highest accuracy estimates in both study areas. In contrast, the accuracies of the wetland type classifications were lower, and the results were not consistent between the study areas. In Chanhassen, the Cowardin and MNDNR accuracy estimates (Tables 8 and 9: "Assess") were significantly lower than the wetland/upland discrimination, but were not as different as in FDL. In Chanhassen, the Cowardin class scheme performed much better than did the MNDNR scheme, while in FDL both type classifications' accuracy estimates were low. We attribute these differences to the lower wetland diversity in Chanhassen and the higher difficulty of mapping the forested wetlands in FDL. Of particular note is that the MNDNR classification, which was developed to be suitable for mapping with remotely sensed data, performed worse than the Cowardin scheme in nearly every classification trial. A likely reason for this discrepancy is that the MNDNR scheme is more thematically detailed i.e., especially in specifying multiple types of emergent wetlands.

Third, the differences between the internal See5 cross-validation and the accuracy assessment results raise interesting questions. In Chanhassen, the cross-validation values were much more similar to the accuracy estimates than they

were in FDL. This indicates that the out-of-bag sampling performed during See5 classifications fairly closely represented the actual accuracies of the Chanhassen results as measured by comparison with the reference data. However, the cross-validation values in FDL were substantially higher in every combination of trial and classification scheme. These large discrepancies may have been caused by the added mapping complexity in FDL; however another important factor may be that field-based reference data were used to assess the accuracy of classifications created with image-based training data. By comparison, in Chanhassen the same dataset was used for training and assessment (though with independent training and reference samples). In FDL, the training data were collected by interpreting aerial images, while the reference data were collected by a field crew over one week. Thus, it is likely that the field team made determinations based on information that was not visible on the aerial images, such as counts of obligate wetland plants. In addition, the image interpreter had access to imagery collected on multiple dates, which may have further increased the effective differences between the training and reference databases. Finally, while the training database contained representatives of all Cowardin and MNDNR classes present in the study area, the field database lacked sufficient representatives of some classes, which resulted in a small downward bias in the overall accuracy estimates.

Fourth, these results show that the inclusion of leaf-off and C-band radar imagery did not increase classification accuracies in FDL. Based on existing literature, challenges to the use of C-band radar in forested areas were expected. Incorporating derivatives of the radar imagery such as polarimetric analyses may have improved the results, but software to perform such tests was not available. Lack of accuracy improvement with the inclusion of leaf-off imagery was surprising. The FDL area contains a mix of coniferous and deciduous vegetation. Spring leaf-off imagery was expected to allow for better viewing of ground features and wetness in deciduous areas; however both the cross-validation and error matrices results showed no significant change in accuracy with its inclusion. It is possible that both the radar and leaf-off results were affected by the especially dry conditions present in FDL during the time of image acquisition (spring/summer 2009), since normal spring wetness that would have been visible on the imagery may not have been as evident. A further potential complicating factor with both the leaf-off and NAIP imagery is differences in illumination levels of the forest canopy. Such differences can be somewhat ameliorated with smoothing of the imagery; however we chose not to degrade the spatial resolution out of concern that doing so would decrease the mapping accuracies of non-forest classes.

In summary, this research suggests the following conclusions related to mapping wetlands with high spatial resolution geospatial data: (a) Mapping accuracy is greatly improved by including topography data with optical imagery; (b) The source of the topography data is less important than its presence or absence; (c) Simple topographic derivatives like

Curvature can provide mapping accuracy similar to the more complex and labor intensive CTI; (d) C-band radar and leaf-off imagery did not improve mapping accuracy in an area with significant forest canopy; (e) Simpler wetland classifications schemes are more likely to perform well than more complex schemes i.e., even those designed with remote sensing in mind; and (f) Mapping wetlands in forested areas is challenging even with the inclusion of several different geospatial data types.

Acknowledgments

This research was funded by the Minnesota Environment and Natural Resources Trust (ENRTF) Fund, the Minnesota Department of Natural Resources (MNDNR), and the United States Fish and Wildlife Service (USFWS: Award 30181AJ194). The authors gratefully acknowledge the valuable help provided by Rick Gitar in the Resource Management Division of the Fond du Lac Band of Lake Superior Chippewa, Steve Kloiber at the MN DNR, the excellent comments of our three peer reviewers, and the service to the remote sensing community by the editor and staff of PE&RS.

References

- Baker, C., R. Lawrence, C. Montagne, and D. Patten, 2006. Mapping wetlands and riparian areas using Landsat ETM+ imagery and decision-tree-based models, *Wetlands*, 26(2):465–474.
- Becker, B.L., D.P. Lusch, and J. Qi, 2005. Identifying optimal spectral bands from in situ measurements of Great Lakes coastal wetlands using second-derivative analysis, *Remote Sensing of Environment*, 97(2):238–248.
- Becker, B.L., D.P. Lusch, and J. Qi, 2007. A classification-based assessment of the optimal spectral and spatial resolutions for Great Lakes coastal wetland imagery, *Remote Sensing of Environment*, 108(1):111–120.
- Bolstad, P.V., and T.M. Lillesand, 1992. Improved classification of forest vegetation in northern Wisconsin through a rule-based combination of soils, terrain, and Landsat Thematic Mapper data, *Forest Science*, 38(1):5–20.
- Bowen, M.W., W.C. Johnson, S.L. Egbert, and S.T. Klopfenstein, 2010. A GIS-based approach to identify and map playa wetlands on the high plains, Kansas, USA, *Wetlands*, 30:675–684.
- Breiman, L., 2001. *Random Forests*, Springer, Netherlands. doi:10.1023/A:1010933404324.
- City of Chanhassen., 2006. Second generation surface water management plan, URL: <http://www.ci.chanhassen.mn.us/?nid=588> (last date accessed: 04 March 2013).
- Congalton, R.G., and K. Green, 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Lewis Publishers, Boca Raton, Florida.
- Corcoran, J.M., J.F. Knight, B. Brisco, S. Kaya, A. Cull, and K. Murhaghan, 2012. The integration of optical, topographic, and radar data for wetland mapping in northern Minnesota, *Canadian Journal of Remote Sensing*, in press.
- Costa, M.P.F., and K.H. Telmer, 2006. Utilizing SAR imagery and aquatic vegetation to map fresh and brackish lakes in the Brazilian Pantanal wetland, *Remote Sensing of Environment*, 105(3):204–213.
- Cowardin, L.M., V. Carter, F.C. Golet, and E.T. LaRoe, 1974. Classification of wetlands and deepwater habitats of the United States, U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C.
- Cowardin, L.M., and V.I. Myers, 1974. Remote sensing for identification and classification of wetland vegetation, *The Journal of Wildlife Management*, 38(2):308–314.
- Creed, I.F., S.E. Sanford, F.D. Beall, L.A. Molot, and P.J. Dillon, 2003. Cryptic wetlands: Integrating hidden wetlands in regression models of the export of dissolved organic carbon from forested landscapes, *Hydrological Processes*, 17(18):3629–3648.
- Dahl, T.E., 1990. Wetlands losses in the United States 1780's to 1980's, U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C.
- Dahl, T.E., C.E. Johnson, 1991. Status and trends of wetlands in the conterminous United States, mid-1970's to mid-1980's, U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C.
- Dechka, J.A., S.E. Franklin, M.D. Watmough, R.P. Bennett, and D.W. Ingstrup, 2002. Classification of wetland habitat and vegetation communities using multi-temporal Ikonos imagery in southern Saskatchewan, *Canadian Journal of Remote Sensing*, 28(5):679:685.
- Dobson, C.M., F.T. Ulaby, and L.E. Pierce, 1995. Land-cover classification and estimation of terrain attributes using synthetic aperture radar, *Remote Sensing of Environment*, 51(1):199–214, doi: DOI: 10.1016/0034-4257(94)00075-X.
- Eggers, S.D., and D.M. Reed, 1997. Wetland plants and communities of Minnesota and Wisconsin, U.S. Army Corps of Engineers, St. Paul District, 263 p.
- Footy, G.M., 1992. On the compensation for chance agreement in image classification accuracy assessment, *Photogrammetric Engineering & Remote Sensing*, 58(10):1459–1460.
- Friedl, M.A., and C.E. Brodley, 1997. Decision tree classification of land cover from remotely sensed data, *Remote Sensing of Environment*, 61(3):399–409.
- Halabisky, M., L.M. Moskal, and S.A. Hall, 2011. Object-based classification of semi-arid wetlands, *Journal of Applied Remote Sensing*, 5:053511-1–053511-13.
- Harvey, K.R., and G.J.E Hill, 2001. Vegetation mapping of a tropical freshwater swamp in the northern territory, Australia: A comparison of aerial photography, Landsat TM and SPOT satellite imagery, *International Journal of Remote Sensing*, 22(15):2911.
- Henderson, F.M., and A.J. Lewis, 2008. Radar detection of wetland ecosystems: A review, *International Journal of Remote Sensing*, 29(20):5809–5835.
- Hess, L.L., J.M. Melack, S. Filoso, and W. Yong, 1995. Delineation of inundated area and vegetation along the Amazon floodplain with the SIR-C synthetic aperture radar, *IEEE Transactions on Geoscience and Remote Sensing*, 33(4), 896–904.
- Hess, L.L., J.M. Melack, E.M.L.M. Novo, and C.C.F. Barbosa, and M. Gastil, 2003. Dual-season mapping of wetland inundation and vegetation for the central Amazon basin, *Remote Sensing of Environment*, 87(4):404–428.
- Hess, L.L., J.M. Melack, and D.S. Simonett, 1990. Radar detection of flooding beneath the forest canopy: A review, *International Journal of Remote Sensing*, 11(7):1313.
- Hirano, A., M. Madden, and R. Welch, 2003. Hyperspectral image data for mapping wetland vegetation, *Wetlands*, 23(2):436–448.
- Hodgson, M.E., J.R. Jensen, H.F. Mackey, and M.C. Coulter, 1987. Remote sensing of wetland habitat: A wood stork example, *Photogrammetric Engineering & Remote Sensing*, 53(10):1075–1080.
- Hogg, A.R., and K.W. Todd, 2007. Automated discrimination of upland and wetland using terrain derivatives, *Canadian Journal of Remote Sensing*, 33(1):68–83.
- Homer, C., J. Dewitz, J. Fry, M. Coan, N. Hossain, and C. Larson, 2007. Completion of the 2001 National Land Cover Database for the conterminous United States, *Photogrammetric Engineering & Remote Sensing*, 73(4):337–341.
- Johnston, C.A., 1989. *Human Impacts to Minnesota Wetlands*, No. PB-91-183160/XAB; EPA-600/J- 89/519), Duluth, Minnesota: U.S. Environmental Protection Agency.
- Jollineau, M.Y., and P.J. Howarth, 2008. Mapping an inland wetland complex using hyperspectral imagery, *International Journal of Remote Sensing*, 29(12):3609–3631.
- Kasischke, E.S., J.M. Melack, and M.C. Dobson, 1997. The use of imaging radars for ecological applications - A review, *Remote Sensing of Environment*, 59(2):141–156.
- Keddy, P.A. 2000. *Wetland Ecology: Principles and Conservation*, Cambridge University Press, New York.

- Kloiber, S., and R. MacLeod, 2011. Supplemental guidance for the classification of wetlands for the update of the National Wetlands Inventory for Minnesota, Minnesota Department of Natural Resources, St. Paul, Minnesota, URL: ftp://ftp.dnr.state.mn.us/pub/eco/nwi/SupplementalMappingGuidance/Supplemental_Guidance_MN_NWI_061411.pdf (last date accessed: 05 April 2012).
- Laba, M., R. Downs, S. Smith, S. Welsh, C. Neider, S. White, M. Richmond, W. Philpot, P. Baveye, 2008. Mapping invasive wetland plants in the Hudson River National Estuarine Research Reserve using QuickBird satellite imagery, *Remote Sensing of Environment*, 112:286–300.
- Li, J., and W. Chen, 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.
- Liu, K., X. Li, X. Shi, and S. Wang, 2008. Monitoring mangrove forest changes using remote sensing and GIS data with decision-tree learning, *Wetlands*, 28(2):336–346.
- Lozano-Garcia, D.F., and R.M. Hoffer, 1993. Synergistic effects of combined Landsat-TM and SIR-B data for forest resources assessment. *International Journal of Remote Sensing*, 14(14):2677.
- Lunetta, R., and M. Balogh, 1999. Application of multi-temporal Landsat-5 TM imagery for wetland identification, *Photogrammetric Engineering & Remote Sensing*, 65(12):1303–1310.
- Maxa, M., and P. Bolstad, 2009. Mapping northern wetlands with high resolution satellite imagers and lidar, *Wetlands*, 29(1):248–260.
- Mitsch, W.J., and J.G. Gosselink, 2000. *Wetlands*, Third edition, John Wiley and Sons, Inc., New York.
- MNGeo, 2012. Minnesota Geospatial Information Office, National Wetlands Inventory Metadata. Retrieved from <http://www.mngeo.state.mn.us/chouse/metadata/nwi.html> (last date accessed; 05 March 2013).
- Murphy, P.N.C., J. Ogilvie, K. Connor, and P.A. Arp, 2007. Mapping wetlands: A comparison of two different approaches for New Brunswick, Canada, *Wetlands*, 27(4):846–854.
- Natural Resources Conservation Service, United States Department of Agriculture, 2010. Soil survey geographic (SSURGO) database for Minnesota, URL: <http://soildatamart.nrcs.usda.gov> (last date accessed: 05 March 2013).
- Neuenschwander, A.L., M.M. Crawford, and M.J. Provanca, 1998. Mapping of coastal wetlands via hyperspectral AVIRIS data, *Proceedings of the Geoscience and Remote Sensing Symposium, IGARSS '98*.
- Ozesmi, S.L., and M.E. Bauer, 2002. Satellite remote sensing of wetlands, *Wetlands Ecology and Management*, 10(5):381–402.
- Parmuchi, M.G., H. Karszenbaum, and P. Kandus, 2002. Mapping wetlands using multi-temporal RADARSAT-1 data and a decision-based classifier, *Canadian Journal of Remote Sensing*, 28(2):175–186.
- Parsons, A.J., 1979. Plan form and profile form of hillslopes, *Earth Surface Processes*, 4:395–402.
- Phillips, R.L., O. Beerli, and E.S. DeKeyser, 2005. Remote wetland assessment for Missouri Coteau prairie glacial basins, *Wetlands*, 25(2):335–349.
- Pontius, R.G., and M. Millones, 2011. Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment, *International Journal of Remote Sensing*, 32(15):4407–4429.
- Pope, K.O., E. Rejmankova, H.M. Savage, J.I. Arredondo-Jimenez, M.H. Rodriguez, and D.R. Roberts, 1994. Remote sensing of tropical wetlands for malaria control in Chiapas, Mexico, *Ecological Applications*, 4(1):81–90.
- Quinlan, J.R., 1993. *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, California.
- Ramsey III, E.W., and S.C. Laine, 1997. Comparison of Landsat Thematic Mapper and high resolution photography to identify change in complex coastal wetlands, *Journal of Coastal Research*, 13(2):281–292.
- Ramsey III, E.W., 1998. Radar remote sensing of wetlands, *Remote Sensing Change Detection: Environmental Monitoring Methods and Applications* (R. Lunetta, and C. Elvidge, editors.) Ann Arbor, Michigan, Ann Arbor Press, Inc., pp. 211–243.
- Rodriguez-Galiano, V.F., B. Ghimire, J. Rogan, M. Chica-Olmo, and J.P. Rigol-Sanchez, 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.
- Rosenqvist, A., M. Shimada, B. Chapman, K. McDonald, G. De Grandi, and H. Jonsson, 2004. An overview of the JERS-1 SAR global boreal forest mapping (GBFM) project, *Proceedings of the Geoscience and Remote Sensing Symposium*.
- Rover, J., C.K. Write, N.H. Euliss, D.M. Mushet, and B.K. Wylie, 2011. Classifying the hydrologic function of prairie potholes with remote sensing and GIS, *Wetlands* 31:319–327.
- Sader, S.A., D. Ahl, and W. Liou, 1995. Accuracy of Landsat-TM and GIS rule-based methods for forest wetland classification in Maine, *Remote Sensing of Environment*, 53(3):133–144.
- Stehman, S.V., and R.L. Czaplewski, 1998. Design and analysis for thematic map accuracy assessment: Fundamental principles, *Remote Sensing of Environment*, 64(3):331–344.
- Thompson, J.A., J.C. Bell, and C.A. Butler, 2001. Digital elevation model resolution: Effects on terrain attribute calculation and quantitative soil-landscape modeling, *Geoderma*, 100(1-2), 67–89.
- Tiner Jr., R.W., 1990. Use of high-altitude aerial photography for inventorying forested wetlands in the United States, *Forest Ecology and Management*, 33-34:593–604.
- Townsend, P., and S. Walsh, 2001. Remote sensing of forested wetlands: Application of multitemporal and multispectral satellite imagery to determine plant community composition and structure in southeastern USA, *Plant Ecology*, 157(2):129–149.
- Wang, Y., L.L. Hess, S. Filoso, and J.M. Melack, 1995. Understanding the radar backscattering from flooded and nonflooded Amazonian forests: Results from canopy backscatter modeling, *Remote Sensing of Environment*, 54(3):324–332.
- Wang, J., J. Shang, B. Brisco, and R.J. Brown, 1998a. Evaluation of multirate ERS-1 and multispectral Landsat imagery for wetland detection in southern Ontario, *Canadian Journal of Remote Sensing*, 24(1):60–68.
- Wang, J., L. Zheng, and Q. Tong, 1998b. Derivative spectra matching for wetland vegetation identification and classification by hyperspectral image, *Proceedings of SPIE*, 3502:280–288.
- Wdowinski, S., S. Kim, F. Amelung, T.H. Dixon, F. Miralles-Wilhelm, and R. Sonenshein, 2008. Space based detection of wetlands' surface water level changes from L-band SAR interferometry, *Remote Sensing of Environment*, 112(3):681–696.
- Wei, A., and P. Chow-Fraser, 2011. Use of IKONOS imagery to map coastal wetlands of Georgian Bay, *Fisheries*, 32(4):167–173.
- Whitcomb, J., M. Moghaddam, K. McDonald, E. Podest, and J. Kellendorfer, 2007. Wetlands map of Alaska using L-band radar satellite imagery, *Proceedings of the Geoscience and Remote Sensing Symposium*.
- Wright, C., and A. Gallant, 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.

(Received 13 April 2012; accepted 03 July 2012; final version 18 January 2013)

Wetland Mapping in the Upper Midwest United States: An Object-Based Approach Integrating Lidar and Imagery Data

Lian P. Rambi, Joseph F. Knight, and Keith C. Pelletier

Abstract

This study investigated the effectiveness of using high resolution data to map wetlands in three ecoregions in Minnesota. High resolution data included multispectral leaf-off aerial imagery and lidar elevation data. These data were integrated using an Object-Based Image Analysis (OBIA) approach. Results for each study area were compared against field and image interpreted reference data using error matrices, accuracy estimates, and the kappa statistic. Producer's and user's accuracies were in the range of 92 to 96 percent and 91 to 96 percent, respectively, and overall accuracies ranged from 96-98 percent for wetlands larger than 0.20 ha (0.5 acres). The results of this study may allow for increased accuracy of mapping wetlands efforts over traditional remote sensing methods.

Introduction

Wetlands are naturally dynamic systems of important value to the environment and society. The US Army Corps of Engineers (USACE) in cooperation with the US Environmental Protection Agency (EPA) have defined wetlands, incorporating technical and policy considerations, as "...those areas that are inundated or saturated by surface or ground water at a frequency and duration to support and under normal circumstances do support, a prevalence of vegetation typically adapted for life in saturated soil conditions" (Federal Register, 1980 and 1982). Wetlands can reduce some of the negative effects of flooding and recharge groundwater by gradually releasing flood water and snow melt. Wetlands offer habitat that supports wildlife and fishing activities. Wetlands also provide ecosystem services, including educational, aesthetic, and economic opportunities. For example, intact freshwater marshes in Canada have a total economic value of approximately 5,800 USD per hectare compared to 2,400 USD when those lands are drained and used for agriculture (Millennium Ecosystem Assessment, 2005; Turner *et al.*, 2000).

Due to wetland loss and degradation, many of the preceding benefits have been reduced and are increasingly impacted. About 215 million acres of wetlands existed in the United States at the time of European settlement. However, by the mid-1970s, only 99 million acres of the original wetlands remained. Many of the lost wetlands were drained and are currently used for agriculture, resource extraction, urbanization, and other commercial purposes (Dahl and Johnson, 1991; Frayer *et al.*, 1983; Stedman and Dahl, 2008). Minnesota is not an exception to this large national wetland loss. Nearly half of Minnesota's original wetlands were lost due to extensive agricultural drainage and urban development. According to the Minnesota Pollution Control Agency (MPCA) (2006), many original natural

wetlands were changed into local storm-water ponds to make additional land available for urban development.

Currently in Minnesota only a few cities have updated wetland inventories. For the rest of Minnesota the only wetland inventory available is the National Wetlands Inventory (NWI). The Minnesota NWI maps were completed in the late 1980s using aerial photos (some black and white) collected between 1979 and 1988 (LMIC, 2007). Several 7.5' quadrangles in northwestern Minnesota and a much larger area in northeastern Minnesota were mapped based on 1970s 1:80 000 scale black-and-white photos (MPCA, 2006). Changes in the landscape have occurred which limit the use of the NWI maps due to the outdated data and techniques used to create them. Thus, there is a need to update wetland inventories with accurate boundaries and improved delineation of smaller wetlands. An updated wetland inventory would provide information that could be used to make accurate decisions for the conservation, protection, and restoration of wetlands. Although a Minnesota statewide update is underway, it is a heavily image interpretation-based project that is not expected to be completed until 2020. Thus, more automated techniques may be useful in the near term.

A fast and effective method to identify accurate wetland boundaries involves the use of remote sensing data and techniques (Butera, 1983; Corcoran *et al.*, 2011; Knight *et al.*, 2013). To the present time, the majority of wetland mapping efforts using remote sensing data and techniques has been focused on evaluating traditional pixel-methods with medium to coarse resolution data. In many cases, the use of remote sensing for wetland mapping has resulted in low accuracy estimates, often due to mixed pixels and insufficient spectral resolution (Grenier *et al.*, 2007; Fournier *et al.*, 2007; Lunetta and Balogh, 1999; Ozesmi and Bauer, 2002). Integration of high resolution optical and elevation data has been shown to reduce the mixed pixel problem (Frohn *et al.*, 2009; Maxa and Bolstad, 2009). Some studies have integrated optical and elevation data to map wetlands using traditional pixel-based methods. However, their accuracy results were low for wetland classification due to the use of low to medium spatial resolution data and pixel-based techniques (Baker *et al.*, 2006; Ozesmi and Bauer, 2002).

An object-based approach may be a better option to integrate high resolution data and overcome some limitations, including the mixed pixel problem and salt-and-pepper effect of traditional pixel-based techniques (Myint *et al.*, 2011; Zhou

Photogrammetric Engineering & Remote Sensing
Vol. 80, No. 5, May 2014, pp. 439–449.
0099-1112/14/8005-439

© 2014 American Society for Photogrammetry
and Remote Sensing

doi: 10.14358/PERS.80.5.439

The Department of Forest Resources, University of Minnesota, 1530 Cleveland Ave, N., Saint Paul, MN 55108 (ortiz073@umn.edu).

and Troy, 2008). Object Based Image Analysis (OBIA) segmentation and classification techniques have been considered as an alternative to pixel-based methods since the late-1990s because of their ability to include contextual information, human knowledge, and experience to interpret the objects of interest (Baatz *et al.*, 2008; Blaschke, 2003; Blaschke, 2010). The foundation of the OBIA approach is an initial image segmentation that uses pixel-based features to create statistically homogeneous image objects (Benz *et al.*, 2004; Fournier *et al.*, 2007). These homogeneous objects, also called geo-objects or segments, can be classified into land-cover classes using attributes of the objects such as spectral, textural, contextual and shape characteristics (Burnett and Blaschke, 2003; Bruzzone and Carlin, 2006; Hay and Castilla, 2008). The OBIA approach can be used to generate vector polygons from the classification and directly incorporate them into a geographic information system (GIS) (Castilla, *et al.*, 2008; O'Neil-Dunne *et al.*, 2012).

The aim of this research was to investigate the effectiveness of using high resolution leaf-off aerial imagery and lidar data to map wetlands in three different ecoregion study areas in Minnesota.

Study Area and Data

Study Area Description

Due to the complexity and variety of wetlands in Minnesota, we selected three study areas to evaluate the OBIA approach to map wetlands. The first study area was the Minnesota River Headwaters watershed located in the Northern Glaciated Plains ecoregion and within Big Stone, Traverse, and Stevens counties (Figure 1). This watershed is 717 km² in size and the main land use is agriculture. A large portion of the watershed is characterized by a rolling prairie of till plain, clay loam soils and a combination of poorly and well drained soils (Minnesota

Department of Natural Resources, 2006). The average precipitation is 640 mm/year and 360 mm during the growing season (May to September). Many shallow lakes and wetlands are common features of the landscape in this watershed. These lakes and wetlands are perfect settings to support and nurture wildlife habitat and viewing opportunities for a variety of bird and duck species (Midwest Community Planning LLC, 2012).

The second study area was the Swan Lake watershed located in the Western Corn Belt Plains ecoregion and within Nicollet County (Figure 1). It has an area of 204 km², and the main land use is agriculture. A large portion of the watershed consists of glacial till plain with level to gently rolling prairie uplands. This area is characterized by clay loam soils and fertile deep soils with a high level of organic matter (Minnesota Department of Natural Resources, 2006). The average precipitation is 740 mm/year and 460 mm during the growing season (May to September). This watershed has one of the biggest prairie pothole marshes in the United States, providing habitat for different species, storm water retention and education opportunities (Nicollet County, 2008).

The third study area is the Thompson Reservoir St. Louis River watershed located in the Northern Lakes and Forest, between St. Louis and Carlton counties (Figure 1). It is 53 km² in size and the main land use is forested land. A large portion of the watershed is characterized by drumlins covered with forest, poorly drained wetland depressions, and fine sandy loam soils. The average precipitation is 710 mm/year and 440 mm during the growing season (May to September).

Data Acquisition

We used two data sources to investigate the effectiveness of integrating multiple datasets to map wetlands in the three study areas. These sources included lidar data and orthorectified digital aerial photography (0.5 m). The half-meter orthorectified imagery used for Swan Lake and the Minnesota River

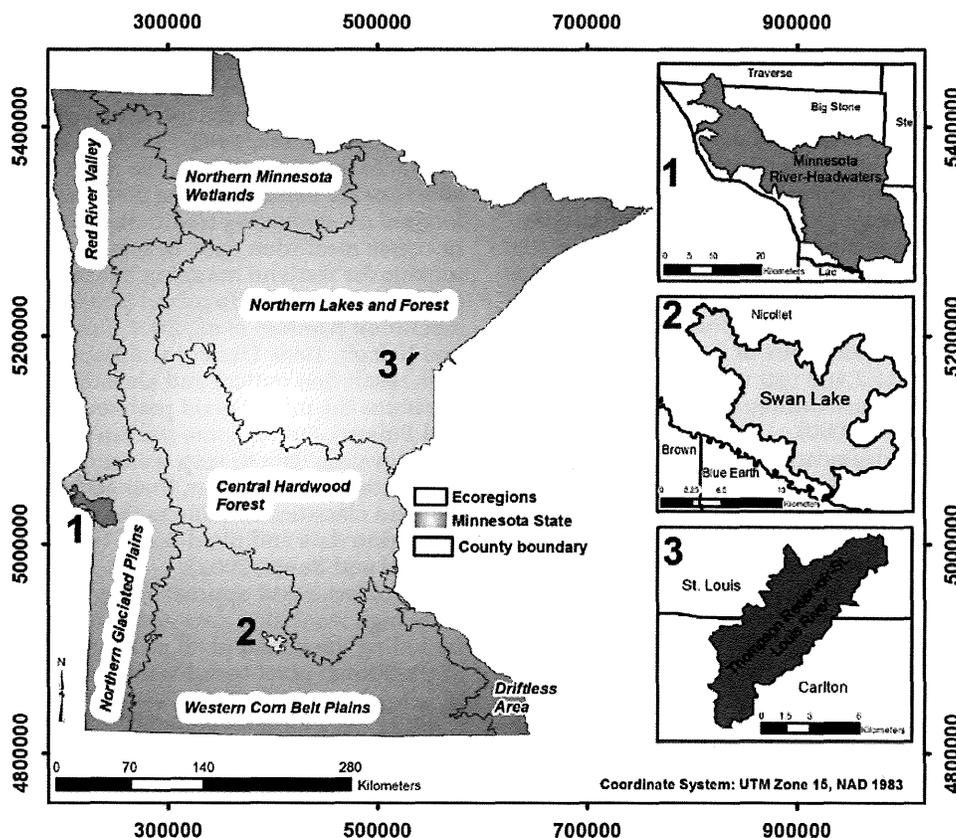


Figure 1. Location of the three watershed study areas in the state of Minnesota, USA.

Headwaters was collected by Surdex Corporation between 12 April 2011 and 16 May 2011. This imagery was provided by the vendor to the Minnesota Department of Natural Resources (MDNR) as a radiometric/orthorectified ready product. The Minnesota Department of Transportation (MnDOT) separately tested the horizontal positional accuracy of this imagery and obtained a root mean square error (RMSE) of 0.819 m with an NSSDA of 1.418 m (95 percent confidence level). This imagery was acquired with an Intergraph DMC[®] (digital mapping camera) from an altitude of about 3,000 m, capturing four multispectral bands (red, green, blue, and near infrared). The half-meter orthorectified imagery used for the Thompson Reservoir St. Louis River watershed was collected by Keystone Aerial Surveys, Inc. in May 2009. This imagery was provided by the vendor to the DNR as a radiometric/orthorectified ready product. The DNR separately tested the horizontal positional accuracy of this imagery and obtained an RMSE of 2 m with an NSSDA of 3.5 m (95 percent confidence level). This imagery was acquired with a Vexcel UltraCamX[®] camera from an altitude of about 7,300 m, capturing four multispectral bands (red, green, blue, and near infrared).

The lidar data (point cloud data, lidar DEM, and lidar hybrid intensity) used for the Minnesota River Headwaters study area was obtained through the International Water Institute (IWI) lidar download portal. The lidar data for the Minnesota River Headwaters was collected in the spring of 2010 during leaf-off conditions by Fugro Horizons, Inc. The data were collected with a Leica sensor ALS50-II MPiA[®] (Multiple Pulses in Air), at an altitude of 2,400 m above mean terrain (AMT), and with an average post spacing of 1.35 m. The horizontal accuracy for these data was of ± 1 m (95 percent confidence level), and a vertical accuracy RMSE of 15.0 cm. For this study area, we used the 1 m DEM and hybrid intensity images provided by the IWI. The DEM was produced by interpolating the bare earth LAS files delivered by the vendor using the "Raster to ASCII" command in ArcGIS[®] 10.1. The hybrid intensity layers were created from lidar intensity and raw lidar/hillshade by the vendor. Hybrid intensity images were created by interpolating the infrared reflectance value attributed for each point. The lidar data (point cloud data, lidar DEM and lidar intensity) used for Swan Lake and Thompson Reservoir St. Louis River watershed were acquired from the Minnesota Geospatial Information Office (MnGeo) FTP site.

The lidar data for the Swan Lake study area was collected between 26 April and 28 April 2010 by AeroMetric, Inc. The data were collected using a multiple fixed wing aircraft lidar system at an altitude of 1,700 m AMT, and an average post spacing of 1.3 m. The horizontal accuracy for these data was of ± 0.3 m (95 percent confidence level), and a vertical accuracy RMSE of 10.0 cm. The lidar data collected for the Thompson Reservoir St. Louis River study area was collected between 03 May and 05 May 2011 by Woolpert, Inc. The data were collected at an altitude of about 2,400 m AMT with an average post spacing of 1.5 m. The horizontal accuracy for these data was ± 1.2 m (95 percent confidence level), and vertical accuracy RMSE was 5 cm. In this study we used the 1 m DEM provided by the Minnesota DNR, which produced the DEM by extracting bare earth points from the point cloud data. The DEM was also hydro flattened using the edge of the water breaklines.

Methods

We mapped wetlands by using an OBIA approach through the creation of rule sets for each study area. We used the Cognition Network Language (CNL) within the software package Definiens eCognition[®] Developer version 8.8.0 to develop the three rule sets. The eCognition Server 64-bit package was used to execute in a batch mode all the tile stacks for each study area. The first subsection of the methods used in this

study describes the data preparation performed for each study area. The next subsection explains the design of the rule set created for each study area. Finally, the last subsection addresses the accuracy assessment procedures used to evaluate results in each study area.

Data Preparation

Before the creation of the three rule set, we performed four data preparation steps needed prior to develop the OBIA approach. First, we generated several raster layers from the lidar point cloud data and DEM. The raster layers included: a digital surface model (DSM), a lidar intensity layer, the compound topographic index (CTI). These raster layers were chosen because of their topographic information, which is useful to differentiate wetland from other cover classes. We used Quick Terrain (QT) Modeler[®] version 7.1.6 to generate the 3 m DSM raster layer using the point cloud data for each study area. The natural-neighbor interpolation algorithm method, the maximum Z value of the first return for all the classes were used to create the DSM layer. We exported the DSM into a raster GeoTIFF file with 3 m spatial resolution.

The lidar intensity images for Swan Lake and the Thompson Reservoir St. Louis River study areas were also generated in QT Modeler with the grid statistic tool, using the mean intensity values of all the lidar returns. We exported the intensity grid layer into a raster GeoTIFF file with 3 m spatial resolution. The lidar intensity image for the Minnesota River Headwaters study area was obtained directly from the IWI download website.

The CTI layers for each study area were created using the DEM layer for each study area. We used the following formula to compute the CTI given by Beven and Kirkby (1979) study: $CTI = \ln [(\alpha) / (\tan (\beta))]$. In this equation α represents the local upslope area draining through each cell, and β represents the local slope gradient. The CTI represents the potential distribution of the water movement and water accumulation across the landscape (Moore *et al.*, 1991). The CTI is used to identify parts of the landscape where sufficient wetness could allow for the formation of wetlands (Rodhe and Seibert, 1999).

Figure 2 shows a map of the Minnesota River-Headwaters study area representing CTI values, where higher CTI values represent water accumulation (potential wetland formation), and lower CTI values represent dryness or steep places where water would not likely accumulate based on topography. The choice of the flow direction algorithm used to calculate α (local upslope area) can affect the accuracy of the CTI. For example, single flow direction algorithms allow flow to pass only to one neighboring downslope cell while multiple flow direction algorithms allow water to flow into multiple neighboring cells. This multidirectional flow effect creates more realistic water flow patterns in different topographic settings, including convex and concave hillslopes (Erskine *et al.*, 2006; Gruber and Peckham, 2008; Wilson *et al.*, 2008). Thus, in this study we used the triangular multiple flow direction algorithm proposed by Seibert and McGlynn (2007) to compute the local upslope area. We used the Whitebox open-source software version 1.0.7 to calculate the contributing area (local upslope area) and local slope layers needed for the CTI. The slope layer was modified by adding a minimum value of 0.0001 to avoid division by zero for CTI calculations.

It is important to clarify that the DEM for the Swan Lake and Thompson Reservoir St. Louis River areas was obtained directly as a raster layer, already mosaicked, from the MnGeo FTP site. However, for the Minnesota River Headwaters areas, we had to mosaic each DEM and hybrid intensity tile contained within this area. Mosaicking was necessary because these data were provided by the IWI in raster tiles of 2,000 m by 2,000 m. We used ERDAS Image[®] 2011 to mosaic and exported the DEM and intensity layers as GeoTIFF files. We also

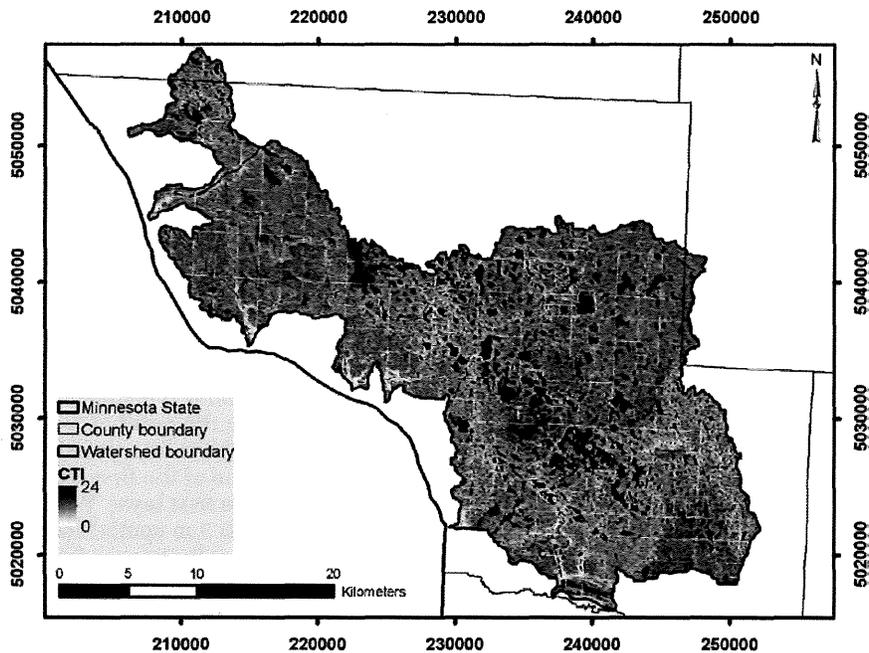


Figure 2. CTI index for Minnesota River-Headwaters study area.

exported the DEM for the Swan Lakes and the Thompson Reservoir St. Louis River study areas to GeoTIFF format.

Second, after calculating all the lidar layers, we used the MosaicPro tool from the ERDAS Image 2011 software to mosaic the orthorectified aerial imagery for each study area.

Third, once all the previous lidar and imagery layers were prepared, we used a watershed boundary shapefile layer for each study area to subset all the raster layers in ERDAS Image version 2011. The watershed boundaries were obtained from the Minnesota Department of Natural Resources (DNR). Finally, we produced a tile generation for each study area. The tile generation was carried out in ERDAS Image version 2011, using the Dice tool with the following parameters: tile size of 3,000 m × 3,000 m and an overlap of 300 m between adjacent tiles on all four sides. Each study area had a tile stack of four lidar product layers (DEM, DSM, CTI, and Intensity) and four bands of imagery layers (Figure 3). The following tile stacks were created: 224 for the Minnesota River Headwaters, 49 for Swan Lake, and 20 for the Thompson Reservoir St. Louis River.

Rule set Creation and Classification

Before the creation of rule sets for each study area, we developed a customized import routine in eCognition developer software to import all the tile stacks of layers for each study area. Each rule set was developed through a trial and error process using small subset areas (500 × 500 pixels). We used a *divide and conquer* approach (Quinlan, 1990; O'Neil-Dunne *et al.*, 2012), which is a multiscale iterative method where objects vary in size, shape, and spectral attributes. While the two major steps performed in the rule set development were the creation of objects and the classification of those objects, further steps were required for the classification of each object to be assigned to the class of interest (wetland class versus non-wetland class). Each rule set consisted of four main components: (a) image processing, (b) segmentation and classification, (c) export operation, and (d) cleanup operation.

In the image processing phase, we carried out the following tasks: calculation of the normalized Digital Surface Model (nDSM) = DSM - DEM, and application of a median filter to the nDSM and intensity layers, and computation of the Green Ratio Vegetation Index (GRVI) using the eCognition developer software tools for object features. The GRVI was computed using

the NIR and green bands of the aerial imagery as the ratio of the NIR divided by the green band (Sripada *et al.*, 2006).

This index was chosen for two reasons: first, it is known that vegetation indices such as the Normalized Difference Vegetation Index (NDVI) can be useful for discriminating wetlands from other upland classes (Hodgson *et al.*, 1987, Wright and Gallant, 2007). Second, after testing several vegetation indices including the NDVI, the Green Normalized Difference Vegetation Index (GNDVI), the Difference Vegetation Index (DVI), and the GRVI to determine which index would be more helpful in differentiating wetland features from non-wetland features. Our unpublished results indicated that the GRVI was more accurate than the other indices to differentiate and exclude areas that were topographically suitable for wetlands but contain impervious cover (non-vegetated).

In the segmentation and classification phase, we performed the following tasks: we created preliminary objects using the multi-resolution segmentation algorithm (Baatz and Schape, 2000) with the following parameter values: scale 30, shape 0.3, and compactness 0.5. A weight value of 1 was given to the three visible optical bands and a weight value of 2 to the NIR band. The scale value of 30 was chosen because we wanted medium size preliminary objects. The shape value of 0.3 was chosen because more weight was given to the influence of color on the segmentation process. The NIR band was given a higher weight value because of its ability to spectrally separate potential non-water objects from water objects.

After creating the preliminary objects, the second step was to refine those objects by applying a spectral difference segmentation algorithm, based on a maximum spectral difference value. The spectral difference algorithm merges neighboring objects based on a maximum spectral difference value parameter (Definiens Imaging, 2009). A value of 14 was chosen as the maximum spectral difference parameter for this difference segmentation. This value was chosen after visually assessing different values.

The third step was to classify the preliminary objects into temporary classes, including wet versus dry, bright versus dark, and short versus tall. We used the following attributes of each dataset to create the temporary classes: min, max, and mean threshold values of the CTI, nDSM, intensity, NIR band, imagery brightness, and GRVI.

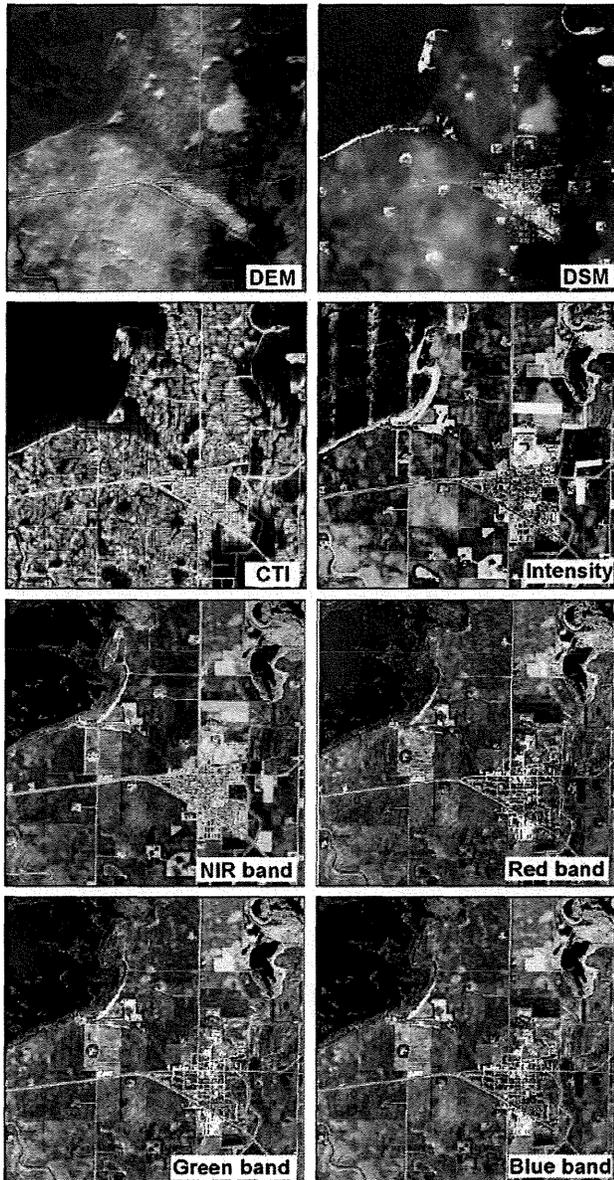


Figure 3. Tile stack of the dataset used for the OBIA approach.

The CTI, GRVI, and NIR layers were specifically used to separate wet versus dry classes with the following threshold values: NIR ≤ 45 , GRVI ≤ 0.9 , and CTI ≥ 10.78 . These threshold parameters were determined through a series of trial-and-error efforts in combination with photo-interpretation to determine whether different “wet classes” (potential wetland classes) across the three different ecoregions were sufficiently separated from dry classes (potential non-wetland classes). The threshold of 10.78 resulted after testing several threshold values at different DEM resolutions including 3 m lidar data. Our unpublished results indicated that the most accurate CTI threshold values to separate wetness (potential wetland) from dryness (upland) was the mean plus one-half standard deviation of the CTI range of values. Also, this CTI threshold value of 10.78 agrees with the value that Galzki *et al.*, (2008) found in their study based on field work.

The imagery brightness, intensity, and GRVI layer were used to classify bright versus dark objects using the spectral difference segmentation algorithm with a maximum spectral difference parameter of 12. The nDSM layer was used to separate short versus tall objects using the contrast split segmentation

algorithm with the following parameters: a minimum threshold value of 2, a max threshold value of 5, and a step size of 1. Previous parameters were determined after several trial-and-error experiments and a detailed visual assessment for separating bright versus dark classes and short versus tall classes.

Finally, we used contextual information from the different temporary classes to achieve the final desired classes. Final classes included wetlands, agriculture, forest, and urban classes. These classes were chosen to allow for easier discrimination between wetland boundaries and upland boundaries due to the spectral, contextual, and shape differences between classes. The contextual information was based on the spatial relationships of an individual object to neighboring objects. For example, small bright objects located in the middle of agriculture fields (unlikely to be impervious surfaces) were reclassified as agriculture classes based on contextual information (neighboring relationship).

In the export operation phase, we exported the final classes into raster and vector formats. In addition, we improved the wetland polygon’s appearance by applying the smoothing and generalizing tools from the advanced editing toolbar in the ArcGIS software.

Accuracy Assessment

We assessed the classification results for the three study areas using a single pixel based approach based on the analysis of the error matrix (Congalton and Green, 2009). The following accuracy assessment estimators were computed in ERDAS Imagine for each study area: overall accuracies, producer’s accuracy, user’s accuracy, and kappa coefficient.

The classification results were evaluated using independent stratified randomly generated points for each study area. Each sample point was interpreted by a trained analyst, who gave the point a value of forest, agriculture, impervious, or wetland. The analyst used aerial photos and field data. In the summer of 2009 and 2011, a team from the Remote Sensing and Geospatial Analysis Laboratory at the University of Minnesota collected field reference data of independent randomly selected points of wetland/upland from different parts of Minnesota including the three study areas used in this study. The field data collected contained the following information: Plant type and percent coverage, land-cover/land-use type, UTM coordinates, five to six photos of the area, and Cowardin wetland type (Cowardin *et al.*, 1974). Upland types included crop fields, other agriculture, forests, grasslands, urban areas, construction areas, bare areas, and others. We generated 289 reference data points for the Minnesota River Headwaters: 118 for Swan Lake and 117 for the Thompson Reservoir St. Louis River study areas.

Results

Results for the three study areas are summarized in Tables 1 through 5, Plate 1, and Figure 4. Overall accuracy results for the OBIA classification were consistently high (90 to 93 percent), throughout the three study areas, with little confusion between the four classes. Within the classification scheme of the four classes, we obtained producer and user accuracies of 92 to 96 percent respectively for the wetland class that included wetlands larger than 0.20 ha (0.5 acres) across the three ecoregions. In addition to the OBIA accuracy assessment classification, a comparison assessment was performed to compare the accuracy of the original NWI and the OBIA classification using only two classes (wetland/upland) for the same study areas. It is important to acknowledge that this comparison of the NWI results and our OBIA results is not a direct and fair comparison. The temporal and methodological differences between the two datasets are significant. Thus, our main objective was to offer an alternative method (OBIA) that will allow for improvements to the accuracy of wetland

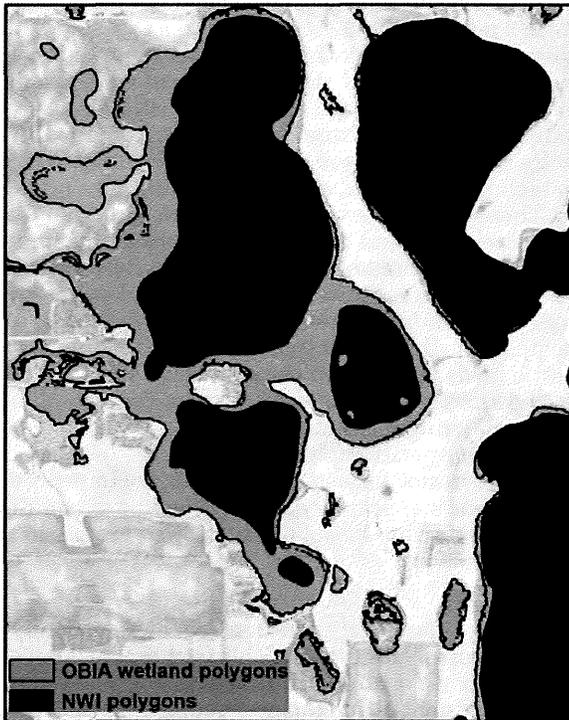


Figure 4. Comparison map of the original NWI polygons and OBIA polygons for a small portion of the Minnesota River-Headwaters with a background of an aerial image.

classification boundaries compared to current wetland boundaries. Updated accurate boundaries of wetlands are necessary, particularly for organizations that currently use older NWI maps as a tool to monitor and regulate wetland management and conservation.

The comparison assessment was done using the kappa-based Z-statistic test described in Congalton and Green (2009). Additionally, the overall accuracy, user's accuracy, and producer's accuracies for wetland and upland classes were computed for both classifications. These comparison results demonstrated that there was a statistically significant difference between the OBIA and the NWI classification at an alpha level of 0.05. For this classification scheme of two classes, the comparison results also indicated that the OBIA wetland class always had a higher user's accuracy (92 to 94 percent) and producer's accuracy (91 to 96 percent) across the three study areas compared to the NWI user's accuracy (56 to 71 percent) and producer's accuracy (57 to 79 percent) for the wetland class.

Table 1 shows a full error matrix and accuracy estimates of the four classes in the Minnesota River-Headwater study area using the OBIA method. The overall accuracy was 90 percent, with a kappa score of 0.84 and low errors of commission and omission. The wetland class was accurately identified with producer's and user's accuracies values at 92 percent. Plate 1a shows a final OBIA classification map with four classes for the Minnesota River-Headwater study area.

Table 2 shows a full error matrix and accuracy estimates of the four classes for the Swan Lake study area. The overall accuracy was 93 percent with a kappa score of 0.90, and with low errors of commission and omissions for the majority of the classes. Plate 1c displays a final OBIA classification map with four classes for this study area. The wetland class in this study area was the least confused compared to other classes (Table 2). Overall, the most confused class pair was agriculture and urban because these classes can be relatively similar spectrally and spatially close in proximity to each other (e.g., an unpaved road bordering or in the middle of an agricultural field).

TABLE 1. OBIA CLASSIFICATION ERROR MATRIX FOR MINNESOTA RIVER-HEADWATER STUDY AREA

		Reference Data				Row Total	User's Accuracy
		Wetlands	Agriculture	Forest	Urban		
Map data	Wetlands	47	4	0	0	51	92%
	Agriculture	2	148	1	5	156	95%
	Forest	1	10	31	0	42	74%
	Urban	1	5	0	34	40	85%
	Column Total	51	167	32	39	289	Overall Accuracy
Producer's Accuracy		92%	89%	97%	87%		90%

Overall Kappa Statistic: 0.84

TABLE 2. OBIA CLASSIFICATION ERROR MATRIX FOR SWAN LAKE STUDY AREA

		Reference Data				Row Total	User's Accuracy
		Wetlands	Agriculture	Forest	Urban		
Map data	Wetlands	27	1	0	0	28	96%
	Agriculture	1	46	0	0	47	98%
	Forest	0	0	23	0	23	100%
	Urban	0	5	1	14	20	70%
	Column Total	28	52	24	14	118	Overall Accuracy
Producer's Accuracy		96%	88%	96%	100%		93%

Overall Kappa Statistic: 0.90

TABLE 3. OBIA CLASSIFICATION ERROR MATRIX FOR THOMPSON RESERVOIR ST. LOUIS RIVER STUDY AREA

		Reference Data				Row Total	User's Accuracy
		Wetlands	Agriculture	Forest	Urban		
Map data	Wetlands	32	0	2	0	34	94
	Agriculture	1	20	3	0	24	83
	Forest	2	0	37	0	39	95
	Urban	0	2	1	17	20	85
	Column Total	35	22	43	17	117	Overall Accuracy
Producer's Accuracy		91%	92%	86%	100%		91%

Overall Kappa Statistic: 0.87

Table 3 shows a full error matrix and accuracy estimates of the four classes for the Thompson Reservoir St. Louis River study area. The overall accuracy was 91 percent with a kappa value of 0.87, and with low errors of commission and omissions for all the classes. Plate 1b displays a final OBIA classification map with four classes for the third study area.

Table 4 shows accuracy estimators of the NWI classification and OBIA classification with two classes (upland versus wetland) for the three study areas, indicating a higher overall accuracy for the OBIA classifications (97 to 98 percent) compared to the NWI classification (74 to 85 percent). In addition, the total amount (hectares) of wetlands for the Minnesota River-Headwaters area, revealed an underestimation of wetlands within the NWI classification. This underestimation also is confirmed by the wetland omission error (43 percent) and low wetland producer's accuracy (57 percent) obtained for the NWI

TABLE 4. OVERALL ACCURACY AND WETLAND USER'S AND PRODUCER'S ACCURACY FOR TWO MAPPING CLASSIFICATION RESULTS (CLASSIFICATION SCHEME: WETLAND/UPLAND)

Land cover classification	Overall accuracy	Wetland user's accuracy	Wetland producer's accuracy	Total area for wetlands in ha
OBIA-Minnesota River-Headwaters	97%	92%	92%	7,620.90
NWI-Minnesota River-Headwaters	88%	71%	57%	6,526.38
OBIA-Swan Lake	98%	96%	96%	4,794.52
NWI-Swan Lake	85%	65%	79%	5,812.04
OBIA- Thompson Reservoir St. Louis River	96%	94%	91%	1,927.29
NWI-Thompson Reservoir St. Louis River	74%	56%	66%	2,233.42

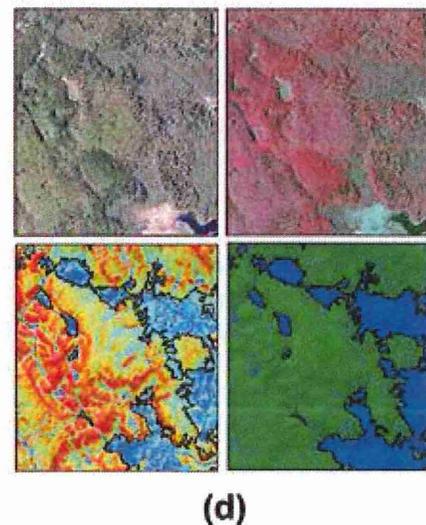
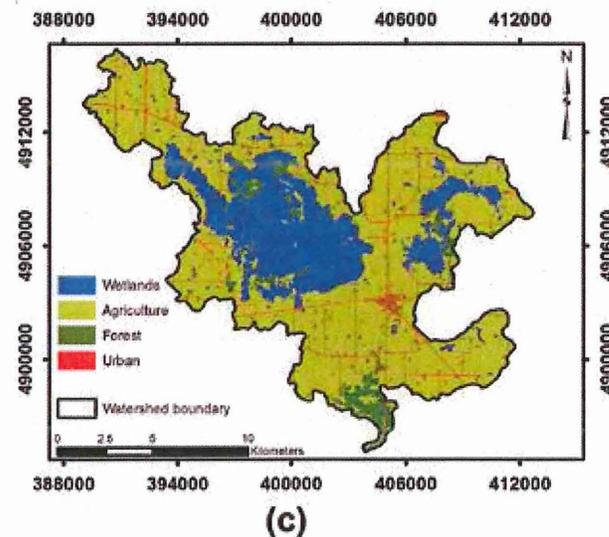
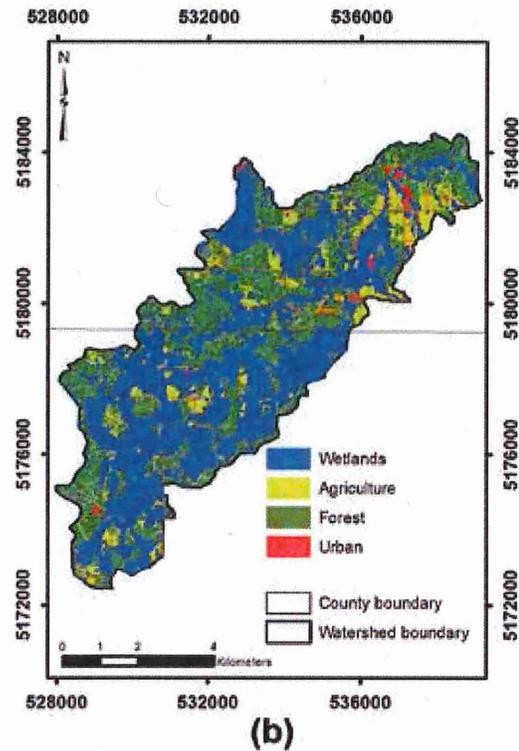
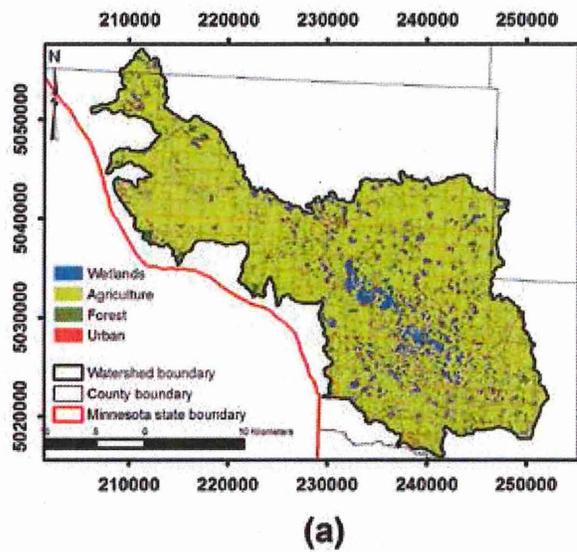


Plate 1. OBIA classification maps for (a) Minnesota River-Headwaters, (b) Thompson Reservoir St. Louis River, and (c) Swan Lake. Comparison of layers for a small portion of the Thompson Reservoir St. Louis River study area, top left visible bands, top right NIR band, bottom left CTI, and (d) bottom right OBIA classes (wetland/upland).

TABLE 5. SIGNIFICANCE TEST (Z-TEST) FOR COMPARING TWO MAPPING CLASSIFICATION SCHEME (WETLAND/UPLAND) USING THE SAME INDEPENDENT REFERENCE DATA POINTS FOR EACH STUDY AREA

Land cover classification	Kappa1 versus Kappa2	Z-Value
OBIA versus NWI for Minnesota River-Headwaters	0.91 versus 0.56	4.61*
OBIA versus NWI for Swan Lake	0.98 versus 0.61	3.88*
OBIA versus NWI for Thompson Reservoir St. Louis River	0.90 versus 0.42	4.83*

* A Z-value over 1.96 indicates that there is a significant difference at 95% confidence level.

wetland class in this area. The total amount (hectares) of wetlands for the Swan Lake and Thompson Reservoir St. Louis River areas exposed an overestimation of the current amount of wetlands compared to the total area amount of wetlands within the NWI classification. This overestimation also is confirmed by the wetland commission error (35 to 44 percent) and low wetland user's accuracy (56 to 65 percent) estimators obtained for the NWI wetland class in these areas.

Table 5 shows the significance test (Z-test) comparison of the two classification methods for each study area; the results were found to be statistically different in each study area at a 95 percent confidence level. Figure 4 shows a comparison of the NWI polygons and OBIA polygons for a small portion of the Minnesota River-Headwaters area. This figure exposes significant differences between NWI and OBIA wetland boundaries, revealing greater amount of wetland omission area for the NWI classification. Although this comparison may be unfair between the NWI and our OBIA results, this comparison confirms the assumption that NWI maps are of limited utility due to their inaccuracy in wetlands versus upland boundaries.

Discussion and Conclusions

In this study, we have evaluated an OBIA approach to map and differentiate wetlands from other classes through the design of a rule set for each study area. The OBIA approach used in this study, across three different ecoregions, provided an adequate platform to integrate different types of high resolution data for accurately detecting wetlands that were greater than 0.20 ha (0.5 acres). OBIA classification maps corresponded well with the reference data for each study area, obtaining high overall accuracy percentages between 90 to 93 percent for the four classes. The results of this study reinforced previous findings regarding the value and importance of high resolution data to improve wetland classification accuracy. Previous studies have concluded that high resolution data including lidar, aerial, and satellite imagery are very advantageous to distinguish between wetlands and non-wetlands classes. These studies have found less confusion between wetlands and upland classes due to the reduction in mixed pixels and addition of high resolution elevation data to separate wetlands from uplands (Everitt *et al.*, 2004; Huan and Zhang 2008; Laba *et al.*, 2008).

The integration of high resolution imagery and lidar data helped to improve classification of wetlands in two ways. First, the use of high resolution data including optical and lidar through an OBIA approach helped to improve the accuracy of wetland classification over traditional pixel-based techniques. For example, Corcoran *et al.* (2011) integrated high resolution imagery with coarse topographic data using a decision-tree classifier to map wetlands, in a similar area to our third study area in the Northern lakes and forest ecoregion area in Minnesota. The Corcoran *et al.* (2011) results

were lower in overall accuracy (72 percent) for wetland/upland classification compared to our OBIA results for wetland/upland classification (96 percent). Sader *et al.*, (1995) compared four satellite image classification methods, including a GIS rule-based model to delineate forest wetlands and other wetlands in Maine. Their results were lower in overall accuracy, ranging from 72 percent to 82 percent for their two study areas. Similarly, other studies have used coarse resolution imagery data including satellite data to map wetlands, but obtained low accuracy estimates for wetland classification because of mixed pixels with similar spectral reflectance (Jensen *et al.*, 1993; Lunetta and Balogh, 1999).

Our study demonstrated that an OBIA approach is more suitable than traditional pixel-based methods to take advantage of the high resolution data available to map wetlands (Dechka *et al.*, 2002; Halabisky *et al.*, 2011; Knight *et al.*, 2013; Maxa and Bolstad, 2009). The OBIA approach used in this study incorporated contextual, spectral, and shape information that came from homogenous objects instead of pixel units. It is important to note that all the high resolution data used in this study were available to the public free of charge. This free high resolution data can be advantageous to many governmental and non-governmental organizations interested in wetland conservation and protection.

Second, the integration of high resolution imagery and lidar data helped to improve classification of wetlands because of the use of high resolution lidar to calculate derivatives such as the CTI. In a qualitative visual assessment of all the data layer inputs, the CTI layer provided additional discrimination between wetland and other non-wetland classes because of its ability to separate low terrain areas from steep terrain areas based on topography (Figure 4). For example, forested vegetation in local low areas were often confused spectrally with forested vegetation in upland areas, but were easier to separate with the addition of the CTI data layer. Other studies have shown similar results when adding topographic data and optical data, resulting in a greater improvement of the wetland accuracy classification. For example, in a study by Knight *et al.* (2013), in an area similar to our third study area, different input datasets including optical and topographic data were evaluated to determine if the addition of different data types would improve the accuracy of wetland classification. The Knight *et al.* (2013), results indicated that topographic data and derivatives including the CTI helped to significantly improve the accuracy of wetland/upland classification compared to other data type scenarios including radar and optical data. That and other similar studies (e.g., Baker *et al.*, 2006; Murphy *et al.*, 2007) reinforce our results regarding the value of using topographic data, which can be categorized as one of the major factors to determine and accurately predict the potential location of wetlands across different ecoregions settings.

It is important to acknowledge that most existing research (e.g., Frohn *et al.*, 2009; Moffett and Gorelick, 2013) using

an OBIA approach to classify wetlands and other land cover has focused more on segmentation techniques, while our study focused more on the development of a customized rule set appropriate to each specific ecoregion setting. The OBIA multiscale iterative approach used in this study involved the design of a customized rule set, allowing us the incorporation of contextual and expert knowledge information through the CNL in the eCognition Developer software. Rule sets can be complex and unique to each area; however, they are adaptable with newer data and transferable to similar areas. Despite the fact that traditional pixel-based techniques are often preferred to study wetlands because of the reduction in analyst time over the classification process, OBIA offers a way to combine the experience and knowledge of the analyst with computer assistance to classify wetlands more accurately in a semi-automated way. Experience and expert knowledge are critical for mapping wetlands, because these ecosystems tend to have a high variability of physical properties. In addition, this experience and knowledge were necessary in our study to obtain and develop crucial contextual information that was not available through traditional pixel-based techniques. In addition to the high accuracy of the results, the output maps were more aesthetically pleasing than pixel-based maps.

Our OBIA results were significantly improved over the original NWI for the three study areas, with lower rates of wetland omissions. Though it is not fair to make a direct comparison between the NWI and the OBIA results, the OBIA approach used in this study suggests an alternative technique to improve the accuracy of wetlands boundaries.

Results from this study included a land-cover classification map with four classes and wetlands polygons for each study area. Lidar data in combination with high resolution imagery significantly improved the accuracy of wetland classification across the three different ecoregions in Minnesota. Our results provide evidence that diverse ecosystems such as wetlands of different sizes can be identified and classified accurately using an OBIA approach with high resolution data across the three different ecoregions studied in this paper. These results are encouraging and useful as an initial classification of wetland habitats but further research is encouraged to classify wetland types, using recently acquired remote sensing data and OBIA rule sets techniques. The OBIA approach presented here to map wetlands offers an alternative, semi-automated and improved method over traditional pixel based techniques and the original NWI. Furthermore, this OBIA approach may be suitable for extension to a larger range of wetlands located in areas such as the ones used in this study, with similar land-use, topography and ecoregion.

Acknowledgments

This research was funded by the Minnesota Environment and Natural Resources Trust (ENRTF), the Minnesota Department of Natural Resources (MNDNR), and the United States Fish and Wildlife Services (USFWS: Award 30181AJ194).

References

- Baatz, M., and A. Schäpe, 2000. Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation, *Angewandte Geographische Informationsverarbeitung XII* (J. Strobl and T. Blaschke, editors), Wichmann, Heidelberg, pp. 12–23.
- Baatz, M., M. Hoffmann, and G. Willhauck, 2008. Progressing from object-based to object-oriented image analysis, *Object Based Image Analysis* (T. Blaschke, S. Lang, and G.J. Hay, editors), Springer, Heidelberg, Berlin, New York, pp. 12–23.
- Baker, C., R. Lawrence, C. Montagne, and D. Patten, 2006. Mapping wetlands and riparian areas using Landsat ETM+ imagery and decision-tree-based models, *Wetlands*, 26(2):465–474.
- Beven, K.J., and M.J. Kirkby, 1979. A physically based, variable contributing area model of basin hydrology, *Hydrological Sciences Journal*, 24(1):43–69.
- Benz, U., P. Hoffman, G. Willhauck, I. Lingenfelder, and M. Heynen, 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information, *ISPRS Journal of Photogrammetry and Remote Sensing* 58(3-4):239–258.
- Burnett, C., and T. Blaschke, 2003. A multi-scale segmentation/object relationship modeling methodology for landscape analysis, *Ecological Modelling*, 168:233–249.
- Butera, M., 1983. Remote sensing of wetlands, *IEEE Transactions on Geoscience and Remote Sensing*, 21(3):383–392.
- Blaschke, T., 2003. Object-based contextual image classification built on image segmentation, *Proceedings of the 2003 IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, 27-28 October, Washington, D.C., pp. 113–119.
- Blaschke, T., 2010. Object based image analysis for remote sensing, *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1):2–16.
- Bruzzzone, L., and L. Carlin, 2006. A multilevel context-based system for classification of very high spatial resolution images, *IEEE Transactions on Geoscience and Remote Sensing*, 44(9):2587–2600.
- Castilla, G., G.J. Hay, and J.R. Ruiz-Gallardo, 2008. Size-constrained region merging (SCRM): An automated delineation tool for assisted photointerpretation, *Photogrammetric Engineering & Remote Sensing*, 74(4):409–419.
- Congalton, R.G., and K. Green, 2009. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Second edition, Boca Raton, Florida, CRC Press/Taylor and Francis.
- Corcoran, J.M., J.F. Knight, B. Brisco, S. Kaya, A. Cull, and K. Murnaghan, 2011. The integration of optical, topographic, and radar data for wetland mapping in northern Minnesota, *Canadian Journal of Remote Sensing*, 37(5):564–582.
- Cowardin, L.M., V. Carter, F.C. Golet, and E.T. LaRoe, 1974. *Classification of Wetlands and Deepwater Habitats of the United States*, U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C.
- Dahl, T.E., and C.E. Johnson, 1991. Status and trends of wetlands in the conterminous United States, mid-1970's to mid-1980, U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C. 28 p.
- Definiens Imaging, 2009. eCognition Imaging Developer, version 8, *ECognition User Guide*.
- Dechka, J.A., S.E. Franklin, M.D. Watmough, R.P. Bennett, and D.W. Ingstrup, 2002. Classification of wetland habitat and vegetation communities using multi-temporal Ikonos imagery in southern Saskatchewan, *Canadian Journal of Remote Sensing*, 28(5):679–685.
- Erskine, R.H., T.R. Green, J.A. Ramirez, and L.H. MacDonald, 2006. Comparison of grid-based algorithms for computing upslope contributing area, *Water Resources Research*, 42:W09416.
- Everitt, J.H., C. Yang, R.S. Fletcher, M.R. Davis, and D.L. Drawe, 2004. Using aerial color infrared photography and QuickBird satellite imagery for mapping wetland vegetation, *Geocarto International*, 19(4):15–22.
- Frazer, W.E., T.J. Monahan, D.C. Bowden, and F.A. Graybill, 1983. *Status and Trends of Wetlands and Deep-Water Habitats in the Conterminous United States, 1950's to 1970's*, Colorado State University, Fort Collins, Colorado.
- Federal Register, 1980. 40 CFR part 230: Section 404b (1), *Guidelines for Specification of Disposal Sites for Dredged or Fill Material*, 45(249):85352–85353.
- Federal Register, 1982. *Title 33: Navigation and Navigable Waters; Chapter II, Regulatory Programs of the Corps of Engineers*, 47(138):31810.
- Fournier, R.A., M. Grenier, A. Lavoie, and R. Helie, 2007. Towards a strategy to implement the Canadian Wetland Inventory using satellite remote sensing, *Canadian Journal of Remote Sensing*, 33(S1):S1–S16.

- Frohn, R., M. Reif, C. Lane, and B. Autrey, 2009. Satellite remote sensing of isolated wetlands using object-oriented classification of LANDSAT-7 data, *Wetlands*, 29(3):931–941.
- Galzki, J., J. Nelson, and D. Mulla. 2008. Identifying critical landscape areas for precision conservation in the Minnesota River Basin, *Proceedings of the 9th International Conference on Precision Agriculture* (R. Khosla, editor), Denver, Colorado.
- Grenier, M., A.M. Demers, S. Labrecque, M. Benoit, R.A. Fournier, and B. Drolet, 2007. An object-based method to map wetland using RADARSAT-1 and Landsat ETM images: Test case on two sites in Quebec, Canada, *Canadian Journal of Remote Sensing*, 33(1):S28–S45
- Gruber, S., and S. Peckham, 2008. Land-surface parameters and objects in hydrology, *Geomorphometry: Concepts, Software, Applications* (T. Hengl and H. Reuter, editors), Elsevier, Amsterdam, The Netherlands. pp. 171–194.
- Halabisky, M., L.M. Moskal, and S.A. Hall, 2011. Object-based classification of semi-arid wetlands, *Journal of Applied Remote Sensing*, 5:053511–1–053511–13.
- Hay, G.J., and G. Castilla, 2008. Geographic object-based image analysis (GEOBIA): A new name for a new discipline, *Object-based Image Analysis-Spatial Concepts for Knowledge-driven Remote Sensing Applications* (T. Blaschke, S. Lang, and G.J. Hay, editors), Springer-Verlag, Berlin, pp. 75–89.
- Hodgson, M.E., J.R. Jensen, H.E. Mackey, and M.C. Coulter, 1987. Remote sensing of wetland habitat: A wood stork example, *Photogrammetric Engineering & Remote Sensing*, 53(8):1075–1080.
- Huan, Y., and S. Zhang, 2008. Applications of high resolution satellite imagery for wetlands cover classification using object-oriented method, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(B7):521–526.
- Jensen J.R., D.J. Cowen, J.D. Althausen, S. Narumalani, and O. Weatherbee, 1993. The detection and prediction of sea level changes on coastal wetlands using satellite imagery and a geographic information system, *Geocarto International*, 4:87–98.
- Knight, J.F., B.T. Tolcser, J.M. Corcoran, and L.P. Rampi, 2013. The effects of data selection and thematic detail on the accuracy of high spatial resolution wetland classifications, *Photogrammetric Engineering & Remote Sensing*, 79(7):613–623.
- Laba, M., R. Downs, S. Smith, S. Welsh, C. Neider, S. White, M. Richmond, W. Philpot, and P. Baveye, 2008. Mapping invasive wetland plants in the Hudson River National Estuarine Research Reserve using QuickBird satellite imagery, *Remote Sensing of Environment*, 112:286–300.
- Land Management Information Center (LMIC), 2007. *Metadata for the National Wetlands Inventory, Minnesota*.
- Lunetta, R.S., and M.E. Balogh, 1999. Application of multi-temporal Landsat-5 TM imagery for wetland identification, *Photogrammetric Engineering & Remote Sensing*, 65(12):1303–1310.
- Maxa, M., and P. Bolstad, 2009. Mapping northern wetlands with high resolution satellite imagery and lidar, *Wetlands*, 29(1):248–260.
- Midwest Community Planning, LLC, 2012. Big Stone County Water Plan. URL: <http://www.bigstonecounty.org/environmental/water-planning/BigStoneCountyWaterPlan.pdf> (last date accessed: 04 March 2014).
- Millennium Ecosystem Assessment, 2005. *Ecosystems and Human Well-Being: Wetlands and Water Synthesis*, World Resources Institute, Washington, D.C.
- Minnesota Department of Natural Resources, 2006. *Tomorrow's Habitat for the Wild and Rare: An Action Plan for Minnesota Wildlife, Comprehensive Wildlife Conservation Strategy*, Division of Ecological Services, Minnesota Department of Natural Resources.
- Moffett, K., and S. Gorelick, 2013. Distinguishing wetland vegetation and channel features with object-based image segmentation, *International Journal of Remote Sensing*, 34(4):1332–1354.
- Moore, I.D., R.B. Grayson, and A.R. Ladson, 1991. Digital terrain modeling: a review of hydrological, geomorphological, and biological applications, *Hydrological Processes*, 5(1):3–30.
- MPCA, 2006. *A Comprehensive Wetland Assessment, Monitoring and Mapping Strategy for Minnesota*, Saint Paul, Minnesota, Minnesota Pollution Control Agency.
- Murphy, P.N.C., J. Ogilvie, K. Connor, and P.A. Arp, 2007. Mapping wetlands: A comparison of two different approaches for New Brunswick, Canada, *Wetlands*, 27(4):846–854.
- Myint, S.W., P. Gloyer, A. Brazel, S. Grossman-Clarke, and Q. Weng, 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery, *Remote Sensing of Environment*, 115(5):1145–1161.
- Nicollet County, 2008. Nicollet county local water management plan, URL: <http://www.co.nicollet.mn.us/> (last date accessed: 04 March 2014).
- O'Neil-Dunne, J.P.M., S.W. MacFaden, A.R. Royar, and K.C. Pelletier, 2012. An object-based system for LiDAR data fusion and feature extraction, *Geocarto International*, 10:1–16.
- Ozesmi, S.L., and M.E. Bauer, 2002. Satellite remote sensing of wetlands, *Wetlands Ecology and Management*, 10(5):381–402.
- Quinlan, J.R., 1990. Decision trees and decision-making, *IEEE Transactions on Systems, Man and Cybernetics*, 20(2):339–346.
- Rodhe, A., and J. Seibert, 1999. Wetland occurrence in relation to topography: A test of topographic indices as moisture indicators, *Agricultural and Forest Meteorology*, 98-99:325–340.
- Sader, S.A., D. Ahl, and W.S. Liou, 1995. Accuracy of Landsat-TM and GIS rule-based methods for forest wetland classification in Maine, *Remote Sensing of Environment*, 53(3):133–44.
- Seibert, J., and B. McGlynn, 2007. A new triangular multiple flow direction algorithm for computing upslope areas from gridded digital elevation models, *Water Resources Research*, 43(4):1–8.
- Sripada, R.P., R.W. Heiniger, J.G. White, and A.D. Meijer, 2006. Aerial color infrared photography for determining early in-season nitrogen requirements in corn, *Agronomy Journal*, 98(4):968–977.
- Stedman, S., and T.E. Dahl, 2008. *Status and Trends of Wetlands in the Coastal Watersheds of the Eastern United States 1998-2004*, National Oceanic and Atmospheric Administration, National Marine Fisheries Service and U.S. Department of the Interior, Fish and Wildlife Service, pp. 1–32.
- Story, M., and R.G. Congalton, 1986. Accuracy assessment: A user's perspective, *Photogrammetric Engineering & Remote Sensing*, 52(3):397–399.
- Turner, R.K., J.C.J.M. van den Bergh, T. Soderqvist, A. Barendregt, J. van der Straaten, E. Maltby, and E.C. van Ierland, 2000. Ecological-economic analysis of wetlands: scientific integration for management and policy, *Ecological Economics*, 35(1):7–23.
- Wilson, J.P., G. Aggett, Y.X. Deng, and C.S. Lam, 2008. Water in the landscape: A review of contemporary flow routing algorithms, *Advances in Digital Terrain Analysis* (Q. Zhou, B. Lees, and G. Tang, editors). Springer, Berlin, pp. 213–236.
- Wright, C., and A. Gallant, 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.
- Zhou, W., and A. Troy, 2008. An object-oriented approach for analyzing and characterizing urban landscape at the parcel level, *International Journal of Remote Sensing*, 29(11):3119–3135.

(Received 14 June 2013; accepted 27 September 2013; final version 15 January 2014)



Comparison of Flow Direction Algorithms in the Application of the CTI for Mapping Wetlands in Minnesota

Lian P. Rampi · Joseph F. Knight · Christian F. Lenhart

Received: 21 August 2013 / Accepted: 28 January 2014
© Society of Wetland Scientists 2014

Abstract Topography has been traditionally used as a surrogate to model spatial patterns of water distribution and variation of hydrological conditions. In this study, we investigated the use of light detection and ranging (lidar) data to derive two Single Flow Direction (SFD) and five Multiple Flow Direction (MFD) algorithms in the application of the compound topographic index (CTI) for mapping wetlands. The CTI is defined here as $\ln [(\alpha)/(\tan (\beta))]$, where α represents the local upslope contributing area and β represents the local slope gradient. We evaluated the following flow direction algorithms: D8, Rho8, DEMON, D- ∞ MD- ∞ , Mass Flux, and FD8 in three ecoregions in Minnesota. Numerous studies have found that MFD algorithms better represent the spatial distribution of water compared to SFD algorithms. CTI maps were compared to field collected and image interpreted reference data using traditional remote sensing accuracy estimators. Overall accuracy results for the majority of CTI based algorithms were in the range of 81–92 %, with low errors of wetland omission. The results of this study provide evidence that 1) wetlands can be accurately identified using a lidar derived CTI, and 2) MFD algorithms should be preferred over SFD algorithms in most cases for mapping wetlands.

Keywords Wetland mapping · Lidar · Flow direction algorithm · Compound topographic index

Introduction

Wetlands are distinctive ecosystems as a result of their hydrologic conditions, chemistry, and transitional bridge between terrestrial and aquatic life.

Wetlands benefit human society and nature in numerous ways. These include support of wildlife habitat, fishing activities and educational activities, protection of shorelines, reduction of negative effects of floods and drought, recharge of groundwater aquifers, cleansing of contaminated waters and climate regulation. The prairie pothole region of southern and western Minnesota, for example, is one of the critical waterfowl nesting and stopover points in the United States. Peatlands, which are abundant in northern Minnesota, have the ability to regulate climate change through carbon sequestration. Peatlands may hold up to 540 gigatons of carbon, representing in approximately 1.5 % of the total estimated global carbon storage (Bridgman et al. 2008; Anteau and Afton 2009; Charman 2009).

Despite their benefits, many wetlands have not been protected but instead have been drained and filled for agricultural or urban development. For example, the United States has lost about 53 % of the original wetlands since the mid-1800s. Those wetlands were converted to agricultural, urban and other commercial land uses (Dahl and Johnson 1991; Stedman and Dahl 2008). Similar change was seen in the state of Minnesota from the 1780s to the mid-1980s where about 42 % of the original wetlands were drained, ditched, filled and converted to other land uses (Dahl 2006). The vast majority of wetland loss occurred in the southern and western agricultural regions of the state while the northern forest region retains more than 90 % of its wetlands (Prince 2008).

Currently the most widely used quantitative source of wetland inventory in the majority of the United States, including Minnesota, is the National Wetlands Inventory (NWI).

L. P. Rampi (✉) · J. F. Knight · C. F. Lenhart
University of Minnesota, Saint Paul, MN, USA
e-mail: ortiz073@umn.edu

L. P. Rampi
Department of Forest Resources, 1530 Cleveland Avenue North,
Saint Paul, MN 55108, USA

However, many NWI maps are outdated, having been completed in the late 1980's, and many changes in the landscape have occurred. Furthermore, the NWI maps were created from aerial imagery (some black and white) collected from 1979 to 1988 (LMIC 2007). Thus, it is important and necessary to update wetland inventories with accurate locations of wetlands. An updated wetland inventory would greatly assist local and state government units in making better decisions for the preservation, protection and restoration of these valuable ecosystems.

The use of topography data provides a fast and cost-effective way to analyze watershed morphology, spatial distribution of soil moisture, and computation of terrain indices useful for improving river, lake, and wetland identification (Rodhe and Seibert 1999; Chaplot and Walter 2003; Sørensen et al. 2006; Corcoran et al. 2011). Digital Elevation Models (DEMs) are preferred to calculate terrain attributes because of the visual representation of these features and the easy computer implementation of algorithms to calculate terrain features (Guntner et al. 2004; Sørensen and Seibert 2007; Shoutis et al. 2010; Knight et al. 2013).

For example, flow direction algorithms can be calculated directly from DEMs to determine in which direction the outflow from a given cell will be distributed to one or more neighboring downslope cells. Flow direction algorithms are important for the calculation of topographic indices such as the Compound Topographic Index (CTI), also known as the Topographic Wetness Index (TWI). One of the valuable benefits of using indices such as the CTI is the ability to represent the distribution and flow of water (saturated vs. non-saturated areas) based only on topographic data (Moore et al. 1993; Guntner et al. 2004; Grabs et al. 2009). The CTI can identify parts of the landscape where sufficient wetness could allow the formation of wetlands. A potential issue with surface flow algorithms is that they do not detect wet areas that are not formed in topographic depressions such as groundwater discharge zones which often occur on slopes. These hydrologic settings may be more difficult to detect with flow direction algorithms in the application of the CTI for mapping rarer wetland types such as fens.

The CTI is based on the formula proposed by Beven and Kirkby (1979): $CTI = \ln[(\alpha)/(\tan(\beta))]$, where α represents the local upslope contributing area per unit contour draining through each cell, and β represents the local slope gradient. Upslope contributing areas are calculated using a flow direction algorithm; thus, the choice of flow direction algorithm is important because it influences the spatial pattern of the CTI values.

Flow direction algorithms are divided in two main groups based on how they distribute flow from one grid cell to another cell (Erskine et al. 2006; Gruber and Peckham 2008; Wilson et al. 2008). The first group consists of single flow direction (SFD) algorithms, which allow flow to pass to only one neighboring cell downslope. The following algorithms are examples

of the SFD group: the Deterministic D8 algorithm proposed by O'Callaghan and Mark (1984), and the random single direction algorithm Rho8 described by Fairfield and Leymarie (1991).

The second group consists of multiple flow direction (MFD) algorithms, which allow flow to pass to more than one neighbor cell downslope. This group is further subdivided into algorithms that allow flow to be distributed to a maximum of two, three, four, and eight neighbor cells downslope. Examples of algorithms that allow flow to be distributed to a maximum of two cells include the Digital Elevation Model Network (DEMON) proposed by Costa-Cabral and Burges (1994), and the Deterministic Infinite ($D\infty$) algorithm suggested by Tarboton (1997).

The Mass Flux (MF) algorithm proposed by Gruber and Peckham (2008) is an example of algorithms that allow flow to pass into a maximum of four neighbors cells. Examples of algorithms that allow flow to be distributed to a maximum of eight neighbor cells include the Triangular Multiple Flow direction algorithm ($MD\infty$) proposed by Seibert and McGlynn (2007), and the Divergent Flow algorithm (FD8) proposed by Freeman (1991). Studies related to hydrological applications across disciplines have used SFD algorithms such as the D-8 more often than MFD algorithms. Although several studies have confirmed that MFD algorithms can provide more accurate results in calculating the distribution and flow of water, the use of SFD algorithms continues (Wilson and Gallant 2000; Zhou and Liu 2002; Pan et al. 2004).

Numerous studies have shown differences between SFD and MFD algorithms for stream network applications and statistical distribution of primary and secondary terrain attributes (Tarboton 1997; Guntner et al. 2004). However, little research has been done to assess the accuracy of these types of algorithms using high resolution elevation data in the application of topographic derivatives such as CTI for identifying wetlands in the upper Midwest, U.S.A. In recent years, the acquisition of high resolution elevation data using Light Detection and Ranging (lidar) has increased.

Lidar is an active remote sensing technology that uses laser light to produce accurate land elevation data. Numerous studies have confirmed the importance of lidar data to improve the process of mapping wetlands (Jenkins and Frazier 2010; Knight et al. 2013; Lang et al. 2013). Lang and McCarty (2009) mapped forested wetlands using lidar intensity and obtained a high overall accuracy of 96.3 %. They compared their lidar intensity results to NIR photointerpretation of wetlands, which had an overall accuracy of 70 % for the same area. Antonarakis et al. (2008) also achieved high overall accuracy results of 95–99 % for mapping open water features using a combination of lidar intensity and lidar derived terrain attributes. Thus, the goal of this paper was to assess the suitability of a selection of two Single Flow Direction (SFD) and five Multiple Flow Direction (MFD) algorithms for use in creating several CTIs from lidar data for wetland mapping in three ecoregions in the state of Minnesota, U.S.

Study Areas Description

This study was conducted in three study areas within three different ecoregions in the state of Minnesota (Fig. 1). The first study area is located in the Northern Glaciated Plains ecoregion and consists of five watersheds of a 12-digit-level Hydrologic Unit Code (i.e., HUC-12). The five watersheds include Big Stone Lake, Big Stone Lake State Park, Barry Lake, Fish Creek, and Salmonson Point, all within Big Stone County. The total area of the five watersheds together is 293 km² with primarily loamy soils and a mixture of well and poorly drained soils. Land use within these watersheds is predominantly agricultural with grain crops, including corn and soybeans. The elevation of these watersheds ranges from 290 to 364 m above sea level.

The average annual precipitation in this area is 640 mm with 360 mm occurring in the growing season of May to September. These watersheds are part of the prairie pot-hole region in Minnesota, characterized by numerous small depressional wetlands known as prairie potholes. Wetlands in this ecoregion are of vital importance for waterfowl habitat, storage of surface water, groundwater

recharge and discharge, and reduction in the risk of downstream flooding (Winter and Rosenberry 1995; LaBaugh et al. 1998).

The second study area is located in the Central Hardwood Forest ecoregion and contains five watersheds of a 12-digit level Hydrologic Unit Code (i.e., HUC-12). The five watersheds include Upper Lake Minnetonka, Riley Creek, Purgatory Creek, Lower Lake Minnetonka and the City of Shakopee-Minnesota River. These watersheds are located within Hennepin and Carver counties. The total area of the five watersheds is 69 km², with fine to moderately coarse texture and well drained soils. Land use is dominated by urban development including medium density residential, with some areas for commercial growth and open space. The elevation across these watersheds is 209–332 m above sea level. The average annual precipitation is 762 mm while during the growing season (May to September) it is 508 mm. The majority of the wetlands types in these watersheds are shallow marshes and wet meadows (City of Chanhassen Surface Water Management Plan 2006).

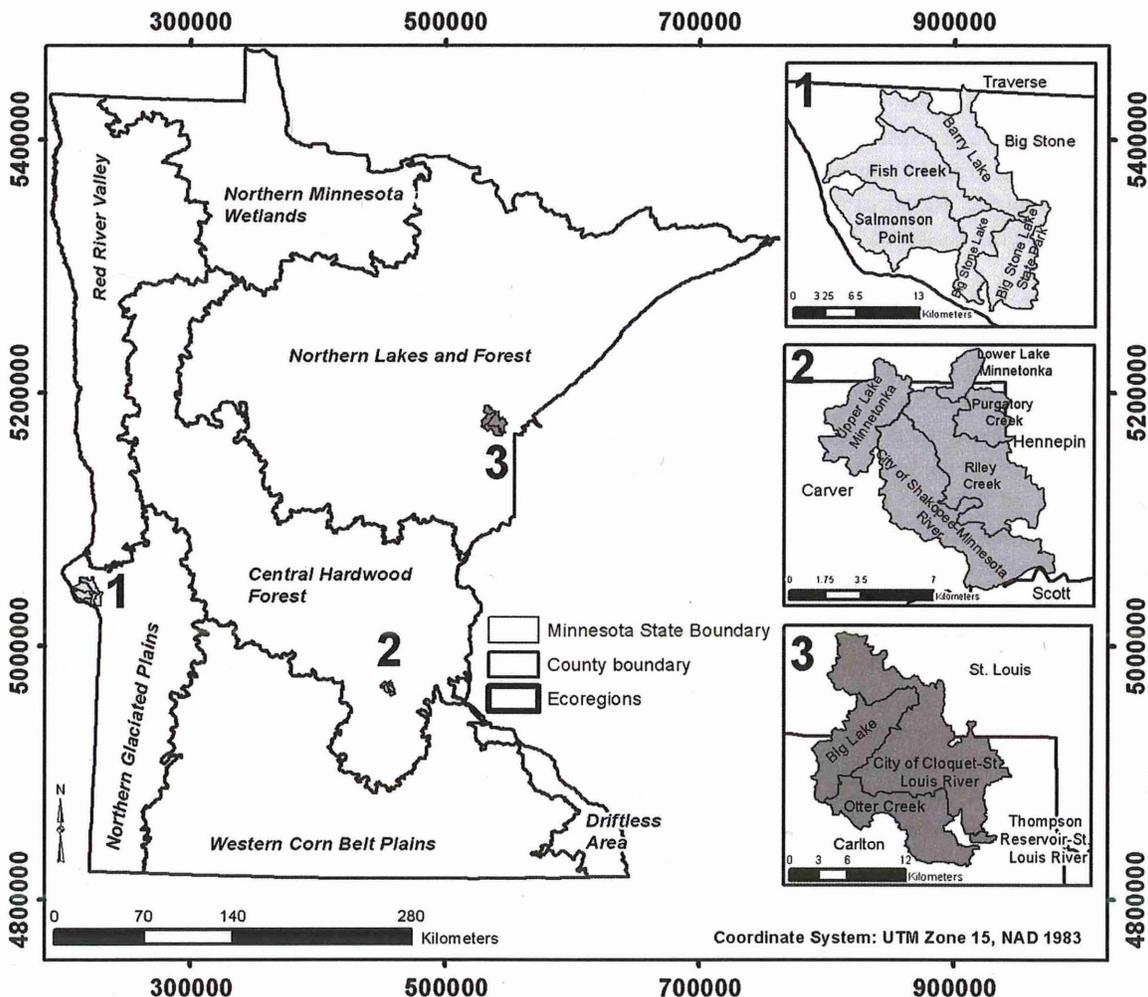


Fig. 1 Three study areas located in three different ecoregions in the state of Minnesota, U.S.A

The third study area is located within the Northern Lakes and Forest ecoregion and includes four watersheds of a 12-digit level Hydrologic Unit Code (i.e., HUC-12). The four watersheds include Big Lake, the City of Cloquet-St. Louis River, Otter Creek and the Thompson Reservoir-St. Louis River. These watersheds are located between St. Louis and Carlton counties. The total area of the four watersheds together is 265 km² with poorly drained soils and near-surface water tables. The main land use in these watersheds is mixed forested land dominated by conifer forest, mixed hardwood-conifer forest and conifer bogs and swamps. The elevation in these areas ranges between 307 and 436 m above sea level. The average annual precipitation is 710 mm and during the growing season (May to September) the average precipitation is 440 mm. Wetlands types in these watersheds are primarily forested wetlands covered by coniferous and tall shrubby vegetation (Minnesota Department of Natural Resources 2010).

Lidar Data

We used a 3 m lidar DEM for each study area to compute seven different flow direction algorithms. The 3 m lidar DEM for the Northern Glaciated Plains study area was obtained from the International Water Institute (IWI) lidar download portal. The DEM was created by interpolating the bare earth point LAS files using the 'Raster to ASCII' command in the Environmental Systems Research Institute (ESRI) ArcGIS software.

Collection of the lidar data used to create the DEM occurred during the spring of 2010 (leaf-off conditions) by Fugro Horizons Inc. with an average post spacing of 1.35 m. The lidar data horizontal accuracy was of ± 1 m (95 % confidence level), with a vertical accuracy RMSE of 15.0 cm.

The 3 m lidar DEM for the Central Hardwood Forest study area was downloaded from the Minnesota Geospatial Information Office (MnGeo). This lidar DEM was produced by the Minnesota DNR by extracting bare earth points from the point cloud data. The DEM was hydro flattened using the edge of the water breaklines. Collection of the lidar point cloud data took place between Nov 11 and Nov 17, 2011 by Fugro Horizons Inc. with an average post spacing of 1.5 m. The horizontal accuracy for these data was of ± 0.6 m (95 % confidence level), and a vertical accuracy RMSE of 5 cm.

The 3 m lidar DEM for the Northern Lakes and Forest study area was also acquired from the Minnesota Geospatial Information Office (MnGeo). The 3 m DEM was produced by the Minnesota DNR by extracting bare earth points from the point cloud data. The DEM was also hydro flattened using the edge of the water breaklines. Acquisition of the lidar data took place between May 3 and May 5, 2011 by Woolpert Inc. with an average post spacing of 1.5 m. The horizontal accuracy of

the lidar data was ± 1.2 m (95 % confidence level), with a vertical accuracy RMSE of 5 cm.

Analysis Methods

This section is composed of three subsections: The first describes the pre-processing steps applied to the lidar DEMs. The second describes the steps and software used to calculate each of the lidar derived terrain attributes required for the CTI calculation CTIs. The third explains the accuracy assessment procedures used to assess the results for each study area.

Lidar DEM Pre-Processing

Each lidar DEM was subset to a shapefile watershed boundary that was obtained from the Minnesota Department of Natural Resources (DNR). Sinks or pits that did not have a surface water outlet were moderately filled to avoid irregularities that could interfere with correct hydrologic flow (trapping flow). We used the tool *fill sinks XXL* implemented in the free open source software System for Automated Geoscientific Analysis (SAGA) v. 2.1.0. We chose this tool because it offers the option to fill sinks fully or partially by keeping a minimum slope gradient along the flow path.

Otherwise, if no minimum slope gradient value was specified, all the sinks would be filled to the spill elevation which would create completely flat areas. Due to the high resolution of our lidar DEMs we avoided filling surface depressions completely by preserving a minimum slope gradient of 0.001 between cells. The resultant sink-moderately-filled DEM for each study area was used to compute the required terrain attributes for calculation of the upslope contributing areas.

Derived Terrain Surfaces

The following flow direction algorithms were implemented in different software packages for the computation of seven upslope contributing areas: The D8, Rho8 and DEMON algorithms were implemented using the SAGA software; the FD8 and MD- ∞ algorithms were implemented using Whitebox Geospatial Analysis Tools v. 1.0.7 open source software; the Mass Flux algorithm was implemented using the River Tools v. 3.0.3, GIS software; the D ∞ algorithm was implemented using the Terrain Analysis Using Digital Elevation Models (TAUDEM) v. 5.0 toolbox in ArcGIS 9.3.1; and the seven upslope contributing areas for each study area were used to calculate the seven CTIs in ArcGIS v. 10.1.

We computed a slope grid in degrees from the partially pre-filled DEM using the spatial analyst tool in ArcGIS v. 10.1 and then converted to radians using the raster calculator. The method used in ArcGIS to compute the slope is the average

maximum technique, where the maximum rate of change in value from a cell to its neighbors is calculated using a 3×3 cell neighborhood around the center grid cell (Burrough and McDonell 1998).

We modified the resultant slope by adding a minimum value of 0.0001 to avoid division by zero for CTI calculations. The raster calculator in ArcGIS v. 10.1 was used to modify the slope and impose the minimum value. Finally, we calculated all the CTIs based on the formula proposed by Beven and Kirkby (1979): $CTI = \ln[(\alpha)/(\tan(\beta))]$. The CTI computations were carried out in ArcGIS v. 10.1 using the raster calculator from the Spatial Analyst toolbox.

Accuracy Assessment

We evaluated the CTI results for each study area based on traditional accuracy assessment methods, including error matrices, overall accuracy, producer's accuracy, user's accuracy, and kappa statistic (\hat{k}) for upland and wetland classes. We also executed a significance test of error matrices known as the Z Pair-Wise statistical test described by Congalton and Green (2009). This Z-test was used to determine whether there was a statistically significant difference between the various CTIs at an alpha level of 0.05. The Z-test was also performed between every CTI and the NWI wetland map using the same classification scheme (upland/wetland).

We thresholded the CTI results into two classes: uplands and wetlands. The threshold values were determined through a series of trial-and-error experiments, where several CTIs across the three different ecoregions were assessed against field data collection and photointerpretation reference points. Results indicated that the most common value for separating upland from wetlands using a 3 m lidar CTI was always the value closest to the mean value of the entire range.

The CTIs and NWI were assessed against a set of independent randomly generated sample points for each study area.

These reference data used for the Northern Glaciated Plain and Central Hardwood Forest study areas were collected from a few sources that included: randomly generated field sites visited by trained field crews in the summers of 2009 and 2010, plots generated by the MN Department of Natural Resources Wetland Status and Trend Monitoring Program (WSTMP) using centroids from polygons of 2006 and 2008 updates, and randomly generated points using photointerpretation by our experienced analyst.

The reference data used for the Central Hardwood Forest study area was developed by the City of Chanhassen using a combination of photo-interpretation and field data collection during the fall of 2004, and the growing season of 2005. The field data collected for the three study areas contained the following information: Plant type and percent coverage, land-cover/land-use type, UTM coordinates, 5–6 photos per site, and the Cowardin wetland type (Cowardin et al. 1974). Upland types included crop fields, other agriculture, forests, grasslands, urban areas, construction areas, bare areas, and others. We used 2000 reference data points for the Northern Glaciated Plains study area, 9,994 for the Central Hardwood Forest study area and 2,000 for the Northern Lakes and Forest study area.

Results

Accuracy assessment results and significance tests of the three study areas are summarized in Tables 1, 2, 3, 4, 5, and 6. Maps of the seven CTIs and NWI wetland/upland classification are displayed in Figs. 2, 3, 4, 5, 6, and 7. Overall accuracy results for the majority of CTIs across the three study areas were in the range of 81–92 % with low errors of wetland omission.

Wetlands larger than 0.20 ha (0.5 acres) throughout the three study areas were identified by all the algorithms, with producer's and user's accuracies in the range of 67–97 % and 65–98 %, respectively.

Table 1 Accuracy estimators of the seven CTIs algorithms and the NWI for the Northern Glaciated Plains study area (Classification scheme: wetland/upland)

CTI algorithm	Threshold used	Overall accuracy	Wetland user's accuracy	Upland user's accuracy	Wetland producer's accuracy	Upland producer's accuracy	Overall kappa
D8	6.0	92	87	98	97	87	0.84
Rho8	6.7	71	70	72	67	75	0.42
DEMON	8.1	92	87	97	96	88	0.84
D-∞	7.2	92	87	97	97	87	0.84
MD-∞	6.1	92	87	97	97	87	0.83
Mass Flux	6.1	91	98	85	85	98	0.82
FD8	11.0	86	87	85	82	89	0.71
NWI	1	88	87	89	87	88	0.76

Total # points used for the accuracy assessment: 2000

Table 2 Significance test (Z-test) for comparing the seven algorithms and the NWI for the Northern Glaciated Plains study area (Classification scheme: wetland/upland)

CTI type	Kappa1 vs. Kappa2	Z-value
D8 vs. Rho8	0.84 vs. 0.42	17.6*
D8 vs. FD8	0.84 vs. 0.71	6.7*
D8 vs. NWI	0.84 vs. 0.76	4.5*
Rho8 vs. DEMON	0.42 vs. 0.83	17.5*
Rho8 vs. D-∞	0.42 vs. 0.84	17.6*
Rho8 vs. MD-∞	0.42 vs. 0.83	17.2*
Rho8 vs. Mass Flux	0.42 vs. 0.82	16.6*
Rho8 vs. FD8	0.42 vs. 0.71	10.9*
Rho8 vs. NWI	0.42 vs. 0.76	13.2*
DEMON vs. FD8	0.83 vs. 0.71	6.5*
DEMON vs. NWI	0.83 vs. 0.76	4.3*
D-∞ vs. FD8	0.84 vs. 0.71	6.6*
D-∞ vs. NWI	0.84 vs. 0.76	4.4*
MD-∞ vs. FD8	0.83 vs. 0.71	6.2*
MD-∞ vs. NWI	0.83 vs. 0.76	4.0*
Mass Flux vs. FD8	0.82 vs. 0.71	5.6*
Mass Flux vs. NWI	0.82 vs. 0.76	3.4*
Fd8 vs. NWI	0.71 vs. 0.76	2.24*

*A Z-value over 1.96 indicates that there is a significant difference at the 95 % confidence level

Also, a comparison assessment of the seven CTIs and the original NWI was performed for each study area, using the same two classes (wetland/upland). The comparison assessment was done using the kappa-statistic (Z-test) proposed by Congalton and Green (2009). The majority of the CTIs based flow direction algorithms derived from lidar data for identifying wetlands; produced higher accuracy results compared to the NWI results that were in the range of 75–88 % for overall accuracy, 73–97 % for user's accuracy and 71–87 % for producer's accuracy across the three study areas.

Results for the Northern Glaciated Plains Study Area

Detailed accuracy assessment results of the seven CTIs algorithms and NWI results of two classes (wetland/upland) are reported in Table 1. The overall accuracies for the CTIs evaluated in this area were in the range of 71–92 %, with overall kappa scores in the range of 0.42–0.84. Producer's and user's accuracies for the CTI's were in the range of 67–97 % and 70–98 % respectively.

The majority of CTIs, with the exception of the CTI Rho8, showed low errors of commission and omissions for the wetland class. The NWI accuracy assessment results were lower than the majority of CTIs for predicting wetland locations in this study area. Table 2 displays only the significance test (Z-test) results of those CTI and NWI results that were found to be statistically different at a 95 % confidence level. These Z-test results revealed that the CTI FD8, CTI Rho8 and NWI maps were significantly different compared to every CTI evaluated.

This statistical difference for the CTI FD8, CTI Rho8 and NWI suggests that the other algorithms are more suitable for identifying wetland occurrences in this ecoregion.

A visual comparison of the seven CTI algorithms and NWI polygons for a small portion of the Northern Glaciated Plains study area are presented on Fig. 2.

This qualitative comparison revealed more details of the differences between the algorithms and the original NWI polygons for representing flow water distribution in wetlands in that area. Figure 3 illustrates a Color-Infrared (CIR) map and a CTI map for this entire study area. Overall, the D8, D-∞, and Mass Flux CTIs were the only algorithms for this study area that showed excellent agreement with the reference data in the visual and quantitative assessment, with the highest overall accuracy results in the range of 91–92 %.

Results for the Central Hardwood Forest Study Area

Accuracy assessment results of the seven CTIs algorithms and NWI results of two classes (wetland/upland) for this study area are presented in Table 3.

Table 3 Accuracy estimators of the seven CTIs algorithms and the NWI for the Central Hardwood Forest study area (Classification scheme: wetland/upland)

CTI algorithm	Threshold used	Overall accuracy	Wetland user's accuracy	Upland user's accuracy	Wetland producer's accuracy	Upland producer's accuracy	Overall kappa
D8	6.1	88	88	89	89	87	0.77
Rho8	5.5	72	71	74	75	70	0.45
DEMON	7.3	85	85	85	85	85	0.71
D-∞	5.4	86	82	92	93	79	0.72
MD-∞	5.1	87	87	87	87	87	0.74
Mass Flux	5.0	85	84	87	87	84	0.71
FD8	5.6	70	70	70	71	70	0.41
NWI	1	85	97	77	71	98	0.70

Total # points used for the accuracy assessment: 9994

Table 4 Significance test (Z-test) for comparing the seven algorithms and the NWI for the Central Hardwood Forest study area (Classification scheme: wetland/upland)

CTI type	Kappa1 vs. Kappa2	Z-value
D8 vs. Rho8	0.77 vs. 0.45	28.7*
D8 vs. DEMON	0.77 vs. 0.71	6.21*
D8 vs. D-∞	0.77 vs. 0.72	4.7*
D8 vs. MD-∞	0.77 vs. 0.74	2.9*
D8 vs. Mass Flux	0.77 vs. 0.71	6.0*
D8 vs. FD8	0.77 vs. 0.41	31.9*
D8 vs. NWI	0.77 vs. 0.70	7.6
Rho8 vs. DEMON	0.45 vs. 0.71	22.5*
Rho8 vs. D-∞	0.45 vs. 0.72	24.15*
Rho8 vs. MD-∞	0.45 vs. 0.74	25.7*
Rho8 vs. Mass Flux	0.45 vs. 0.71	22.7*
Rho8 vs. FD8	0.45 vs. 0.41	3.18*
Rho8 vs. NWI	0.45 vs. 0.70	21.6*
DEMON vs. MD-∞	0.71 vs. 0.74	3.2*
DEMON vs. FD8	0.71 vs. 0.41	25.7*
D-∞ vs. FD8	0.72 vs. 0.41	27.4*
D-∞ vs. NWI	0.72 vs. 0.70	2.8*
MD-∞ vs. Mass Flux	0.74 vs. 0.71	3.0*
MD-∞ vs. FD8	0.74 vs. 0.41	28.9*
MD-∞ vs. NWI	0.74 vs. 0.70	4.5*
Mass Flux vs. FD8	0.71 vs. 0.41	25.9*
Fd8 vs. NWI	0.41 vs. 0.70	24.8*

*A Z-value over 1.96 indicates that there is a significant difference at the 95 % confidence level

Overall accuracy percentages for the CTIs assessed in this study area were in the range of 70–88 %, with overall kappa scores in the range of 0.41–0.77. Producer's and user's accuracies for the CTI's were in the range of 71–93 % and 70–88 %, respectively. The majority of CTI algorithms excluding the Rho8 and FD8 showed low errors of commission and omissions for the wetland class. NWI producer's accuracy

was relatively low compared to the majority of CTIs, which resulted in higher rates of wetland omission in this area. Table 4 displays the significance test (Z-test) results of those CTIs and NWI maps that were found to be significantly different at a 95 % confidence level.

The CTI FD8, CTI Rho8 and CTI D8 were found to be statistically significant compared to the rest of the CTI and NWI results. A detailed visual comparison of the seven algorithms, wetland polygons created by the City of Chanhassen, and NWI polygons for a small portion of this study area is presented in Fig. 4. This visual comparison exposes many differences between the polygons created by the City of Chanhassen, the NWI polygons and the straight flow water patterns of the single flow direction algorithms. A map of the CTI and CIR image for the complete study area is shown in Fig. 5.

In general, out of all the algorithms tested, the D-∞ and MD-∞ CTIs indicated excellent agreement with the reference data in the visual and quantitative assessment for this study area. These CTIs had high overall accuracy results in the range of 86–87 %, with low errors of wetland omissions and commission.

Results for the Northern Lakes and Forest Study Area

Table 5 shows accuracy assessment results for the two classes (wetland/upland) for this study area. Overall accuracy results for the CTI's based algorithms evaluated in this study area were in the range of 69–82 % with kappa scores between 0.38 and 0.64. Producer's and user's accuracies for the CTI's were in the range of 80–86 % and 65–81 %, respectively. NWI accuracy assessment estimators were lower compared to the majority of the CTI algorithms for this area. Lower accuracy assessment results of the NWI revealed the inaccuracy of the polygons in this forested area for identifying wetlands, particularly forested wetlands.

Table 6 displays significance tests (Z-tests) for only CTI algorithms that were found to be statistically different at a

Table 5 Accuracy estimators of the seven CTIs algorithms and the NWI for the Northern Lakes and Forest study area (Classification scheme: wetland/upland)

CTI algorithm	Threshold used	Overall accuracy	Wetland user's accuracy	Upland user's accuracy	Wetland producer's accuracy	Upland producer's accuracy	Overall kappa
D8	5.2	82	80	84	86	78	0.64
Rho8	6.1	69	65	77	84	54	0.38
DEMON	7.1	75	73	77	80	70	0.50
D-∞	7.0	81	81	81	81	81	0.61
MD-∞	5.5	82	80	83	84	80	0.63
Mass Flux	6.0	81	80	81	82	79	0.61
FD8	5.8	81	79	83	83	78	0.61
NWI	1	75	73	78	80	70	0.50

Total # points used for the accuracy assessment: 2000

Table 6 Significance test (Z-test) for comparing the seven algorithms and the NWI for the Northern Lakes and Forest study area (Classification scheme: wetland/upland)

type	Kappa1 vs. Kappa2	Z-value
D8 vs. Rho8	0.64 vs. 0.38	9.78*
D8 vs. DEMON	0.64 vs. 0.50	5.43*
D8 vs. NWI	0.64 vs. 0.50	5.24*
Rho8 vs. DEMON	0.38 vs. 0.50	4.18*
Rho8 vs. D-∞	0.38 vs. 0.61	8.84*
Rho8 vs. MD-∞	0.38 vs. 0.63	9.43*
Rho8 vs. Mass Flux	0.38 vs. 0.61	8.59*
Rho8 vs. FD8	0.38 vs. 0.61	8.80*
Rho8 vs. NWI	0.38 vs. 0.50	4.37*
DEMON vs. D-∞	0.50 vs. 0.61	4.52*
DEMON vs. MD-∞	0.50 vs. 0.63	5.10*
DEMON vs. Mass Flux	0.50 vs. 0.61	4.28*
DEMON vs. FD8	0.50 vs. 0.61	4.48*
D-∞ vs. NWI	0.61 vs. 0.50	4.34*
MD-∞ vs. NWI	0.63 vs. 0.50	4.91*
Mass Flux vs. NWI	0.61 vs. 0.50	4.10*
Fd8 vs. NWI	0.61 vs. 0.50	4.30*

*A Z-value over 1.96 indicates that there is a significant difference at the 95 % confidence level

95 % confidence level. The CTI FD8, CTI Rho8 and CTI D8 were found to be statistically significant different compared to the rest of the CTI and NWI results for mapping wetlands. Visual comparisons of the seven CTI algorithms and NWI polygons for a small portion of this study area are shown in Fig. 6. This visual comparison revealed the differences between the algorithms and NWI polygons for predicting forested wetlands. Figure 7 shows two maps: the CTI and CIR image, for the whole study area. In general, the D-∞, MD-∞, and Mass Flux CTIs were the only algorithms that had excellent agreement with the reference data in the visual and quantitative assessment for this study area. These three algorithms had the highest overall accuracy results in the range of 81–82 %, with relatively low errors of wetland omissions and commission.

Discussion

We compared and evaluated seven CTI based algorithms derived from lidar DEMs for identifying wetlands across three different ecoregions in Minnesota. The computation of the CTI offered a practical and fast method to identify wetlands greater than 0.20 ha. All CTI based maps showed a relatively high overall percentage of agreement with the reference data for wetland and upland classes (69–92 %). Results of this

study demonstrate that lidar derived CTIs can significantly improve the accuracy of wetlands classification compared to the NWI across different ecoregions in Minnesota.

Although a direct comparison of the NWI and our CTI results may be not fair because of the differences in data types and techniques used to create these two wetland maps; the CTI-based approach developed here provides an alternative efficient and accurate method to update wetland maps. Available updated wetland maps would be valuable for many governmental and non-governmental entities that currently only used NWI maps as a tool and resource to monitor and take decisions regarding wetlands.

Our results showed the importance of choosing the correct flow direction algorithm for identifying wetlands location visually and quantitatively. Visual comparison of the seven CTI algorithms in the three study areas revealed noticeable differences that are partially seen in the quantitative accuracy assessment analysis for some algorithms.

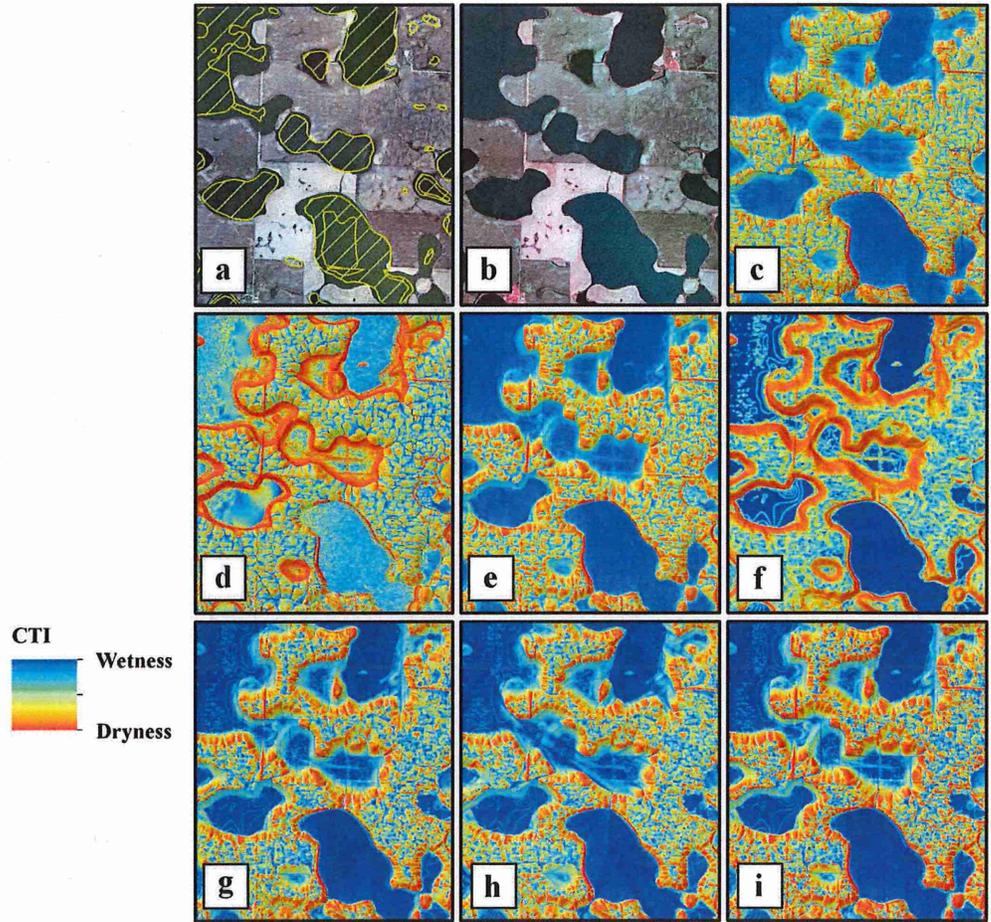
We speculate that the quantitative accuracy assessment analysis did not show strong differences for all the algorithms because of the type of reference data used to assess these algorithms: points instead of polygon reference data types. For example, the D8 SFD algorithm exhibits similar quantitative accuracy results compared to three of the MFD algorithms (D-∞, MD-∞, Mass Flux) in the three study areas; nevertheless, the qualitative visual analysis exposes major difference related to unrealistic parallel flow patterns of the SFD algorithms (D8 and Rho8) for differentiating wetlands from uplands.

Similarities and differences between the two groups of algorithms are also highlighted in the way each of these algorithms tends to distribute the flow and accumulation of water in wetlands and uplands across the three study areas.

The Northern Glaciated Plains study area exhibited similarities in the way the majority of the CTI based algorithms represented water flow and accumulation for wetland mapping. For example, the D8, D-∞, and Mass Flux CTIs showed parallel flow patterns and similarly high accuracy assessment results. Low topography relief and presence of more concave hillslopes in this study area were the two main factors that favored greater flow convergence for the majority of wetlands located in this study area. These factors may explain the similarities in performance of the majority of flow direction algorithms in this area. Additionally, this study area had the highest overall accuracy, user's and producer's accuracy results compared to the other two study areas.

High accuracy results can be explained primarily because of the type of wetlands found in this study area, known as prairie pothole wetlands or depressional wetlands (LaBaugh et al. 1998). The majority of flow accumulation that contributes to the hydrology of these wetlands tends to occur in these topographic depressions that can be identified efficiently using high resolution elevation data. As a result, the CTI method tested in this

Fig. 2 Visual comparison of **a** the NWI polygons, **b** CIR aerial imagery 2011, **c** D8 CTI, **d** Rho8 CTI, **e** DEMON CTI, **f** FD8 CTI, **g** D-∞ CTI, **h** MD-∞ CTI, **i** Mass Flux CTI for the Northern Glaciated Plains study area. Higher CTI values represent water accumulation (potential wetland location) and lower CTI values represent dryness



study is an efficient mapping technique to identify these wetlands because of the topographic nature of this index.

For the Central Hardwood Forest study area marked visual differences between the SFD and MFD algorithms were

Fig. 3 **a** CIR aerial imagery 2011 map, and **b** CTI map for the Northern Glaciated Plains study area

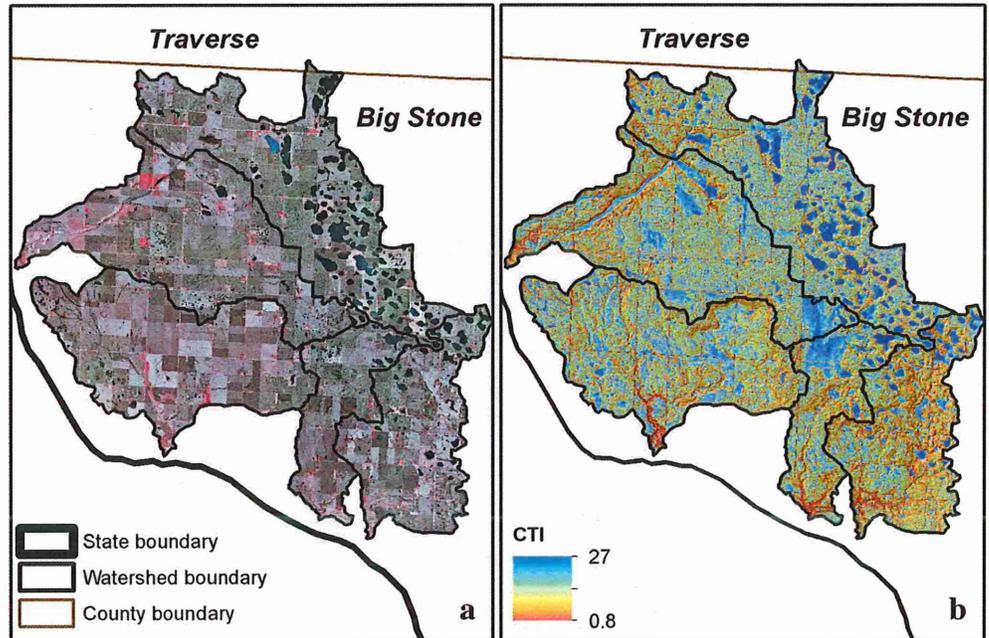
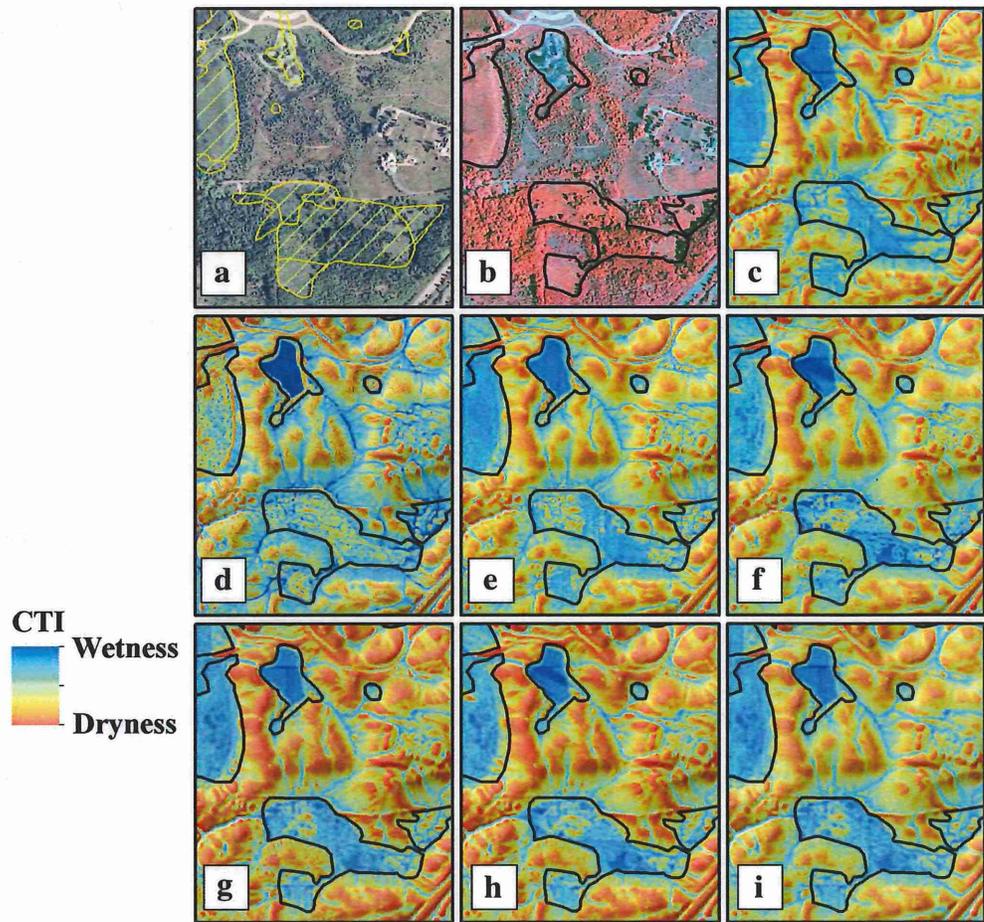


Fig. 4 Visual comparison of **a** the NWI polygons, **b** CIR aerial imagery 2008 and wetland polygons created by the City of Chanhassen, **c** D8 CTI, **d** Rho8 CTI, **e** DEMON CTI, **f** FD8 CTI, **g** D-∞ CTI, **h** MD-∞ CTI, **i** Mass Flux CTI for the Central Hardwood Forest study area. Higher CTI values represent water accumulation (potential wetland location) and lower CTI values represent dryness



displayed in this study. For example, parallel flow patterns were very evident on the D8, Rho8 and DEMON CTIs. The

Rho8 showed the lowest accuracy assessment results for classifying wetlands and uplands.

Fig. 5 **a** CIR aerial imagery 2008 map, and **b** CTI map for the Central Hardwood Forest study area

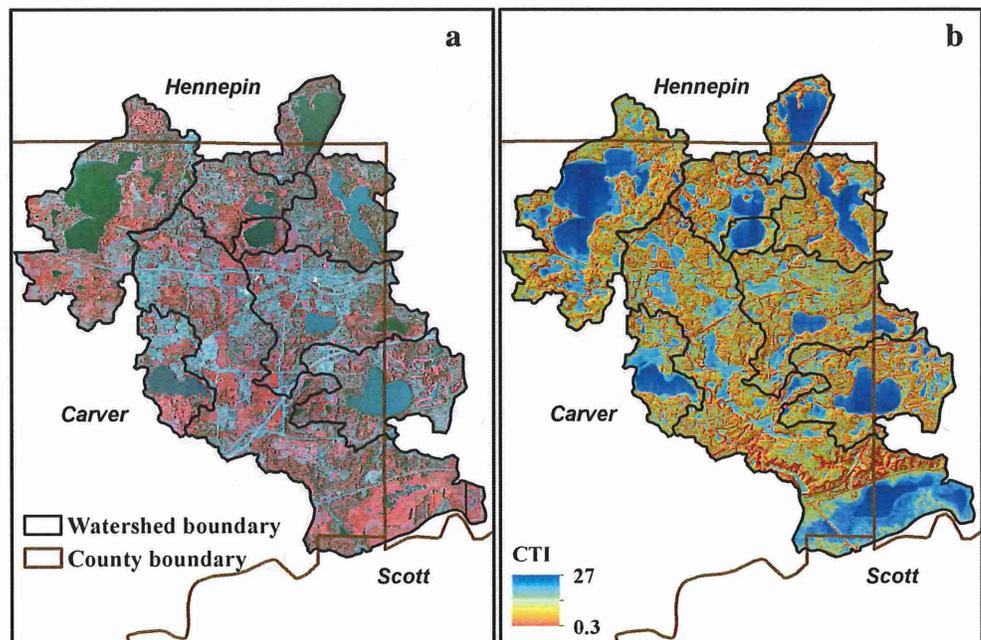
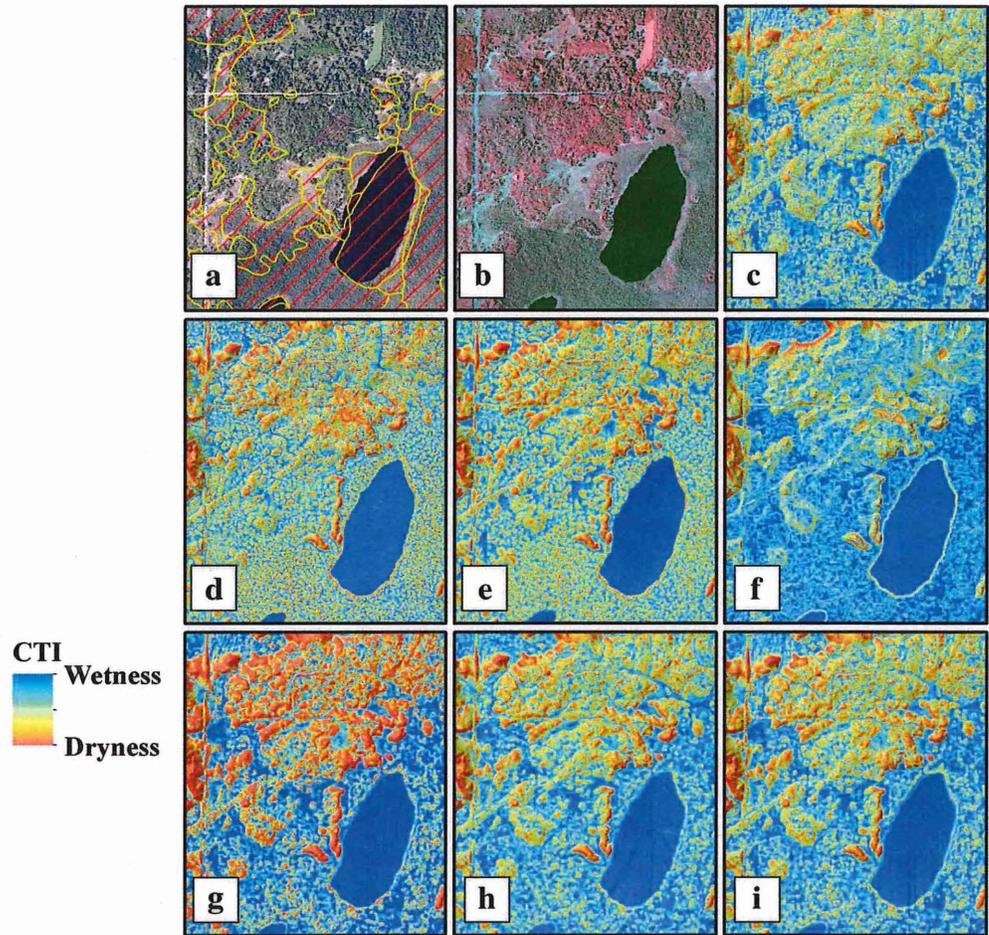


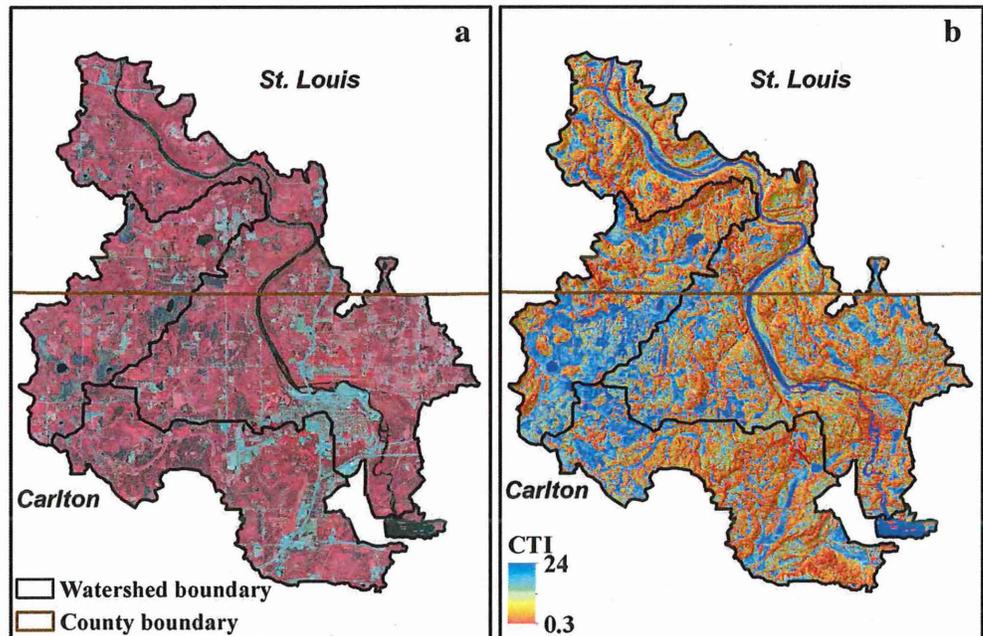
Fig. 6 Visual comparison of **a** the NWI polygons, **b** CIR aerial imagery 2009, **c** D8 CTI, **d** Rho8 CTI, **e** DEMON CTI, **f** FD8 CTI, **g** D-∞ CTI, **h** MD-∞ CTI, **i** Mass Flux CTI for the Northern Lakes and Forest study area. Higher CTI values represent water accumulation (potential wetland location) and lower CTI values represent dryness



This study area had the second relatively high overall accuracy, user's and producer's accuracy results compared to the other two study areas. Visually and statistically the best

algorithms for mapping wetlands in this area were the D-∞ and MD-∞ CTIs. Marked differences between algorithms in this study area can be attributed to the presence of medium to

Fig. 7 **a** CIR aerial imagery 2009 map, and **b** CTI map for the Northern Lakes and Forest study area



high topography relief and more convex hillslopes near or in the type of existing wetlands in this area.

The majority of existing wetlands include open water, shallow and deep marshes, and unconsolidated bottom (Knight et al. 2013). Thus, CTIs based on MFD algorithms were more suitable than SFD algorithm for this area to represent realistic patterns of wetlands areas and greater flow of divergence distribution of water.

The Northern Lakes and Forest study area study area had the lowest overall accuracy, user's accuracy, and producer's accuracy results for all the CTIs maps compared to the other two study areas. This can be explained because of the wetlands types located in this area which includes calcareous fens, sedge meadows, hardwood wetlands, coniferous swamps, and coniferous bogs. The majority of these existing wetlands in this area are groundwater-fed wetlands, and generally high in the landscape.

For example, fens wetlands are groundwater discharge wetlands that occur along topographic or geologic breaks or where groundwater aquifers are exposed near the surface. Thus, these types of wetlands are less sensitive to topography influence and inundation events as they are located at an elevation above floodplain.

Nevertheless, of all the CTI's based algorithms, the MFD algorithms performed better at visually separating uplands from wetlands. The D- ∞ , MD- ∞ , and Mass Flux CTIs had the highest accuracy results for separating wetlands from uplands. The three MFD algorithms mentioned above allowed for a more divergent and smoother distribution of water in very pronounced convex-steep hillslopes near or close to these wetlands.

Lang et al. (2013) reinforces our results regarding the accuracy and preferences for MFD over SFD algorithms for identifying wetland locations. The Lang et al. (2013) results indicate that the FD8 CTI multiple flow direction algorithms derived from lidar data performed better than other non-distributed flow direction algorithms including the D8 for identifying locations of forested wetlands in the Coastal Plain of Maryland.

Our significance Z-test results for the three study areas confirmed the significant differences between the SFD algorithms and MFD, particularly for the Rho8 CTI, across the three study areas. CTI based algorithms (D8, D- ∞ , MD- ∞ , Mass Flux D, and FD8) wetland/upland classification maps in general were significant improvements over the NWI map for two of our study areas. However, for the Central Hardwood Forest study area, the CTI based algorithms (D8, D- ∞ , MD- ∞ , and Mass Flux D) outperformed only the NWI. NWI results in this area had high errors of omission because of rapid urban development over the past 6 years.

Our research demonstrated the outdated nature of many NWI maps in Minnesota. Still, many of these maps are used by governmental and non-governmental policymakers

for wetland management and policy development for lack of better data. Improved mapping accuracy will be greatly beneficial for policymakers developing local or regional wetland inventories, restoration or mitigation plans and other policies.

Conclusions

Lidar derived CTIs enable a fast, efficient, and more accurate method to estimate current wetland location compared to NWI maps. Our results provide evidence that different wetland types in varied ecoregions can be identified accurately using lidar derived terrain indices. In general, the seven CTI based algorithms were able to predict wetland locations across different ecoregions. However, there were statistically and visually significant differences in their performance.

Our visual comparison results revealed that CTIs based on MFD algorithms are generally better than CTIs based on SFD algorithms for separating wetlands from uplands. Based on our results, we suggest the use of the following algorithms: D- ∞ , MD- ∞ or Mass Flux in the application of the CTI for mapping wetlands in areas similar to the ones evaluated in this study. The MFD algorithms represented the distribution and accumulation of water (wetness) in wetlands in a more visually accurate form compared to SFD algorithms.

Further research is encouraged to investigate the effect of different DEM resolutions and use of the CTI combined with other ancillary data such as optical data for mapping wetlands. The combination of the CTI and other ancillary data could potentially help to identify wetlands located at an elevation above floodplain level where elevation information alone is not as influential as it is for depressional wetlands. For example, organic flat wetlands and groundwater discharge-fed wetlands that occur along slopes including some types of fens may require additional tools to map with greater accuracy.

Additional research is also needed to address evaluate numerically the visual differences seen in this study from the different flow direction algorithms. One possible approach could be a wet area-polygon based assessment, that would extract and measure the amount of CTI wet areas found only in wetland references polygons.

Finally, the use of NWI maps continues across different parts of the country because these maps are the most accessible information available. Many of these NWI maps need to be updated. Remote sensing techniques including those based on the CTI offer a fast, cost-effective and reliable method to quickly identify wetland location and update such maps.

Acknowledgments This research was funded by the Minnesota Environment and Natural Resources Trust (ENRTF), the Minnesota Department of Natural Resources (MNDNR), and the United States Fish and Wildlife Services (USFWS: Award 30181AJ194).

References

- Anteau MJ, Afton AD (2009) Wetland use and feeding by lesser scaup during spring migration across the upper Midwest, USA. *Wetlands* 29:704–712
- Antonarakis AS, Richards KS, Brasington J (2008) Object-based land cover classification using airborne Lidar. *Remote Sensing of Environment* 112:2988–2998
- Beven KJ, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Journal* 24:43–69
- Bridgham SD, Pastor J, Dewey B, Weltzin JF, Updegraff K (2008) Rapid carbon response of peatlands to climate change. *Ecology* 89:3041–3048
- Burrough PA, McDonell RA (1998) Principles of geographical information systems. Oxford University Press, New York, 190 pp
- Chaplot V, Walter C (2003) Subsurface topography to enhance the prediction of the spatial distribution of soil wetness. *Hydrological Processes* 17:2567–2580
- Charman DJ (2009) Peat and peatlands. Elsevier Inc, 541–548
- City of Chanhassen Surface Water Management Plan (2006) In: The second generation surface water management plan - Chanhassen, Minnesota: <http://www.ci.chanhassen.mn.us/serv/cip/swmp/wetlandsmanagement.htm>. Accessed 25 May 2013
- Congalton RG, Green K (2009) Assessing the accuracy of remotely sensed data: principles and practices, 2nd edn. CRC Press/Taylor and Francis, Boca Raton
- Corcoran JM, Knight JF, Brisco B, Kaya S, Cull A, Murnaghan K (2011) The integration of optical, topographic, and radar data for wetland mapping in northern Minnesota. *Canadian Journal of Remote Sensing* 37(5):564–582
- Costa-Cabral M, Burges SJ (1994) Digital elevation model networks (DEMON): a model of flow over hillslopes for computation of contributing and dispersal areas. *Water Resources Research* 30:1681–1692
- Cowardin, L.M., V. Carter, F.C. Golet, and E.T. LaRoe, 1974. Classification of wetlands and deepwater habitats of the United States, U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C.
- Dahl TE (2006) Status and trends of wetlands in the conterminous United States 1998 to 2004. U.S. Department of the Interior; Fish and Wildlife Service, Washington, D.C., p 112
- Dahl TE, Johnson CE (1991) Status and trends of wetlands in the conterminous United States, mid-1970's to mid-1980's. U.S. Fish and Wildlife Service, Washington, DC, p 28
- Erskine RH, Green TR, Ramirez JA, MacDonald LH (2006) Comparison of grid-based algorithms for computing upslope contributing area. *Water Resources Research* 42, W09416
- Fairfield J, Leymarie P (1991) Drainage networks from grid digital elevation models. *Water Resources Research* 27:709–717
- Freeman GT (1991) Calculating catchment area with divergent flow based on a regular grid. *Computers and Geosciences* 17:413–422
- 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* 373:15–23
- Gruber S, Peckham S (2008) Land-surface parameters and objects in hydrology. In: Hengl T, Reuter HI (eds) *Geomorphometry: concepts, software, applications*. Elsevier, Amsterdam, pp 171–194
- Guntner A, Seibert J, Uhlenbrook S (2004) Modeling spatial patterns of saturated areas: an evaluation of different terrain indices. *Water Resources Research* 40, W05114
- Jenkins RB, Frazier PS (2010) High-resolution remote sensing of upland swamp boundaries and vegetation for baseline mapping and monitoring. *Wetlands* 30:531–540
- Knight JF, Tolcser BT, Corcoran JM, Rampi LP (2013) The effects of data selection and thematic detail on the accuracy of high spatial resolution wetland classifications. *Photogrammetric Engineering and Remote Sensing* 79:613–623
- LaBaugh JW, Winter TC, Rosenberry DO (1998) Hydrologic functions of prairie wetlands. *Great Plains Research: A Journal of Natural and Social Sciences* 8:17–37
- Land Management Information Center (LMIC) (2007) Metadata for the National Wetlands Inventory, Minnesota
- Lang MW, McCarty GW (2009) Lidar intensity for improved detection of inundation below the forest canopy. *Wetlands* 29:1166–1178
- Lang MW, McCarty GW, Oesterling R, Yeo I (2013) Topographic metrics for improved mapping of forested wetlands. *Wetlands* 33:141–155
- Minnesota Department of Administration (AdminMN) Office of geographic and demographic analysis state demographic center, 2010 census: Minnesota city profiles. <http://www.demography.state.mn.us/CityProfiles2010/index.html>. Accessed 20 May 2013
- Moore ID, Gessler PE, Nielsen GA, Peterson GA (1993) Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal* 57:443–452
- O'Callaghan JF, Mark DM (1984) The extraction of drainage networks from digital elevation data. *Computer Vision, Graphic and Image Processing* 28:328–344
- Pan F, Peters- Lidar CD, Sale MJ, King AW (2004) A comparison of geographical information system-based algorithms for computing the TOPMODEL topographic index. *Water Resources Research* 40: 1–11
- Prince H (2008) *Wetlands of the American Midwest: a historical geography of changing attitudes*. Chicago: University of Chicago Press
- Rodhe A, Seibert J (1999) Wetland occurrence in relation to topography - a test of topographic indices as moisture indicators. *Agricultural and Forest Meteorology* 98–99:325–340
- Seibert J, McGlynn B (2007) A new triangular multiple flow direction algorithm for computing upslope areas from gridded digital elevation models. *Water Resources Research* 43:1–8
- Shoutis L, Dunca TP, McGlynn B (2010) Terrain-based predictive modeling of Riparian vegetation in Northern Rocky Mountain watershed. *Wetlands* 30:621–633
- Sørensen R, Seibert J (2007) Effects of DEM resolution on the calculation of topographical indices: TWI and its components. *Journal of Hydrology* 347:79–89
- Sørensen R, Zinko U, Seibert J (2006) On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrology and Earth System Sciences* 10:101–112
- Stedman S, Dahl TE (2008) Status and trends of wetlands in the coastal watersheds of the Eastern United States 1998 o 2004. National Oceanic and Atmospheric Administration, National Marine Fisheries Service and U.S. Department of the Interior, Fish and Wildlife Service, 32 pages
- Tarboton DG (1997) A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research* 33:309–319
- Wilson JP, Gallant JC (2000) Secondary topographic attributes. In: Wilson JP, Gallant JC (eds) *Terrain analysis: principles and applications*. Wiley, New York, pp 87–131
- Wilson JP, Aggett G, Deng YX, Lam CS (2008) Water in the landscape: a review of contemporary flow routing algorithms. In: Zhou Q, Lees B, Tang G (eds) *Advances in digital terrain analysis*. Springer, Berlin, pp 213–236
- Winter TC, Rosenberry DO (1995) The interaction of ground water with prairie pothole wetlands in the Cottonwood Lake Area, eastcentral North Dakota, 1979–1990. *Wetlands* 15:193–211
- Zhou Q, Liu X (2002) Error assessment of grid-based flow routing algorithms used in hydrological models. *International Journal of Geographical Information Science* 16:819–842