



Object-based forest classification to facilitate landscape-scale conservation in the Mississippi Alluvial Valley



Michael Mitchell^{a,*}, R. Randy Wilson^b, Daniel J. Twedt^c, Anne E. Mini^d, J. Dale James^a

^a Ducks Unlimited, Southern Regional Office, 193 Business Park Drive, Suite E, Ridgeland, MS 39157, United States

^b U.S. Fish and Wildlife Service, 6578 Dogwood View Pkwy, Suite C, Jackson, MS 39213, United States

^c U.S. Geological Survey, Patuxent Wildlife Research Center, University of Memphis, 950 Getwell Road, Memphis, TN 38152, United States

^d Lower Mississippi Valley Joint Venture and American Bird Conservancy, 193 Business Park Drive, Suite E, Ridgeland, MS 39157, United States

ARTICLE INFO

Article history:

Received 27 July 2015

Received in revised form

21 January 2016

Accepted 25 January 2016

Available online 11 May 2016

ABSTRACT

The Mississippi Alluvial Valley is a floodplain along the southern extent of the Mississippi River extending from southern Missouri to the Gulf of Mexico. This area once encompassed nearly 10 million ha of floodplain forests, most of which has been converted to agriculture over the past two centuries. Conservation programs in this region revolve around protection of existing forest and reforestation of converted lands. Therefore, an accurate and up to date classification of forest cover is essential for conservation planning, including efforts that prioritize areas for conservation activities. We used object-based image analysis with Random Forest classification to quickly and accurately classify forest cover. We used Landsat band, band ratio, and band index statistics to identify and define similar objects as our training sets instead of selecting individual training points. This provided a single rule-set that was used to classify each of the 11 Landsat 5 Thematic Mapper scenes that encompassed the Mississippi Alluvial Valley. We classified $3,307,910 \pm 85,344$ ha (32% of this region) as forest. Our overall classification accuracy was 96.9% with Kappa statistic of 0.96. Because this method of forest classification is rapid and accurate, assessment of forest cover can be regularly updated and progress toward forest habitat goals identified in conservation plans can be periodically evaluated.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Before European settlement, approximately 9.7 million ha within the Mississippi Alluvial Valley were forested (King et al., 2006). By the 1930s the area of bottomland forest in this region had declined to about 4.2 million ha and suffered an additional 45% decline through the 1980s (Oswalt, 2013). Since then, conservation partners in the Mississippi Alluvial Valley have implemented strategic habitat conservation for wildlife via a landscape-scale approach to forest conservation and restoration (U.S. Fish and Wildlife Service, 2008). This approach relies upon: (1) the development of species-habitat models to define sustainable landscapes; (2) implementation of conservation actions in accordance with these landscape-designs; and (3) the ability to monitor and evaluate progress towards meeting conservation objectives. Specifically, the Lower Mississippi Valley Joint Venture partnership exists for the purpose of facilitating landscape-scale conservation and restoration of bottomland hardwood forest ecosystems with an emphasis on supporting healthy populations of avian species and other forest dependent wildlife species in this

region (Twedt et al., 1999; LMJV Operational Plan, 2013). Therefore, the ability to characterize the forest landscape (e.g., amount and spatial arrangement) is necessary to facilitate planning and evaluation of conservation goals.

To characterize the forest landscape of the Mississippi Alluvial Valley, conservation partners have historically used: (1) traditional analytical methods using 30 m resolution Landsat TM imagery and aerial photography in a supervised classification (Twedt and Loesch, 1999); (2) data from the U.S. Forest Service's Forest Inventory and Analysis Program (Bechtold and Patterson, 2005) in conjunction with aerial or satellite photography (Rudis and Birdsey, 1986; Oswalt, 2013); and (3) publicly available National Land Cover Data (NLCD; Fry et al. 2011).

Although these approaches provided useful and accurate estimates of forest area and distribution, advances in remote sensing software and analysis now permit classification of remotely sensed imagery that is more economical, efficient, and has improved accuracy compared to previously used pixel-based classification methods. For example, relatively inexpensive medium and high-resolution imagery is available that allows analysis at improved spatial and temporal scales. Geographic Information System (GIS) software has incorporated object-based image analyses as an alternative to pixel-based methods. Additionally, decision tree

* Corresponding author.

creation in software supports analysis and classification of landscapes with improved repeatability (Friedl and Brodley, 1997; Pal and Mather, 2003; Immitzer et al., 2012; Mellor et al., 2013). This combination of object-based image analysis and decision tree classification has the potential to facilitate and enhance landscape-scale conservation efforts through use of more effective, transparent, and repeatable analytical processes.

Object-based image analysis allows for segmentation, attribution, classification, and establishment of relationships among defined objects that are not possible in pixel-based analyses (Cufi et al., 2002). This image analysis method takes digital input (e.g., Landsat 5 Thematic Mapper [TM] imagery) in the form of spectral bands, as well as spectral indices created from these bands, and creates multiple sets of similar pixels (i.e., objects) that may be more meaningful and easier to analyze than individual pixels (Blaschke and Strobl, 2001). Each identified object assumes the attributes of all the pixels that comprise it, as well as contextual

information such as its relationship to surrounding objects.

When used in conjunction with classification and regression tree (CART) methods, object-based image analysis allows for the creation of dynamic decision tree rule-sets that can be applied to separate datasets (Breiman et al., 1984). This method uses training objects to predict the class of other objects. As a binary classification tree, a test and output decision is applied at each node within CART analysis until reaching a final prediction. The Random Forest algorithm is analogous with CART methodology but builds multiple decision trees and compares the outcome of all these trees to make a decision. Even so, Random Forest is fast, accurate, and is capable of forming as many trees as the user specifies without over fitting the data being analyzed (Breiman and Cutler, 2005). In many cases, Random Forest classification produces higher accuracies than other classification approaches and has been successfully used to separate imagery classes that are spectrally similar (Akar and Güngör 2012).

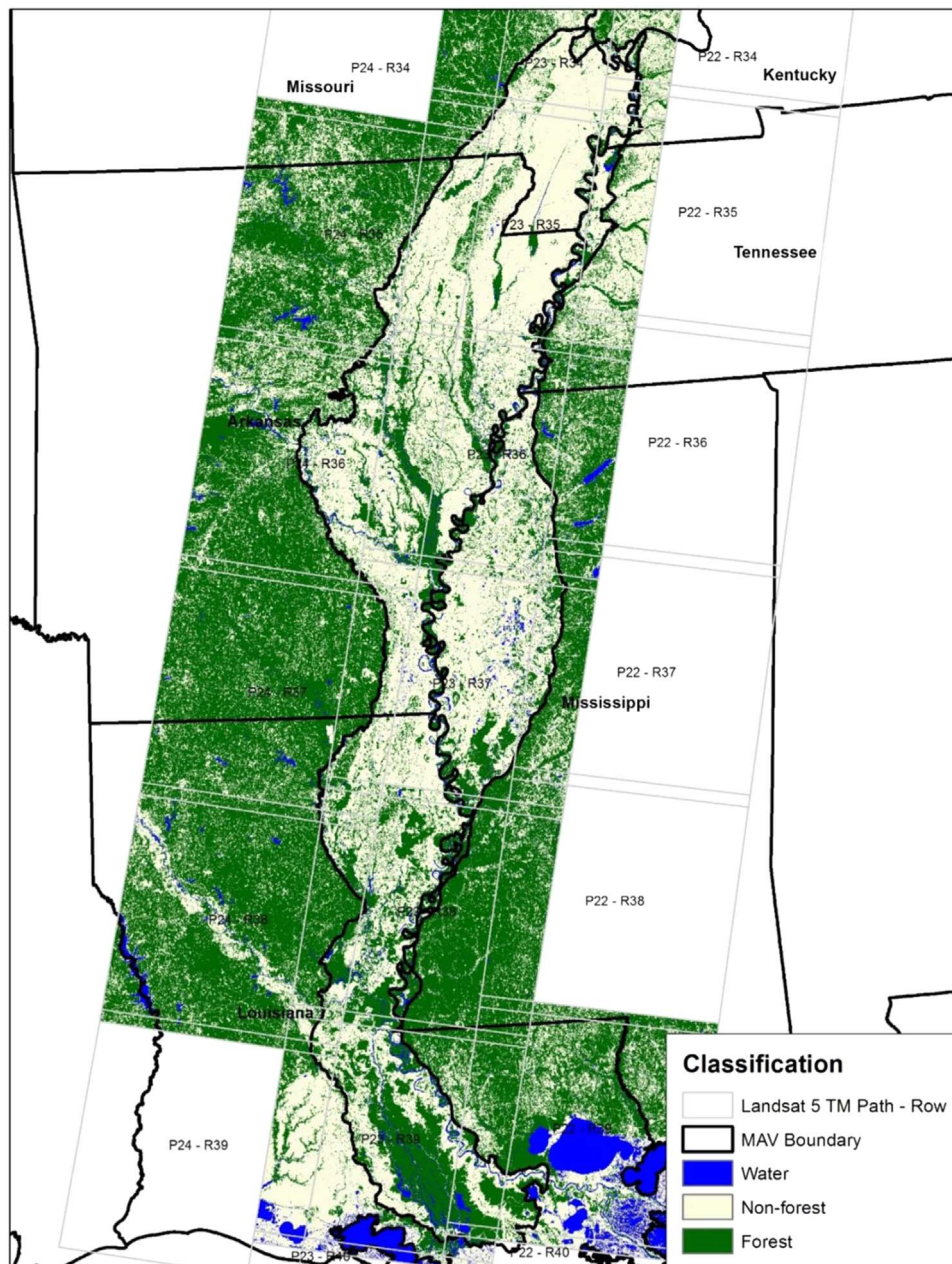


Fig. 1. Boundaries of the Mississippi Alluvial Valley and Landsat 5 Thematic Mapper (TM) scenes (path-rows) as well as the final classification.

Our objective was to develop a forest classification process, using object-based image analysis and Random Forest algorithms, which provides conservation partners an effective, accurate, and repeatable method to evaluate the areal extent and spatial distribution of forested habitat within the Mississippi Alluvial Valley.

2. Methods

2.1. Study area

The 10 million ha Mississippi Alluvial Valley is the floodplain adjacent to the Mississippi River from southern Illinois to southern Louisiana. We used boundaries defined as Bird Conservation Region 26 by the North American Bird Conservation Initiative (<http://www.nabci-us.org/bcrs.htm>), but refined it to better delineate the transition from alluvial floodplain and delta land to upland habitats (Fig. 1; <http://www.arcgis.com/home/item.html?id=c72185797b564b5995f44e9bc367163e>).

The Mississippi Alluvial Valley has experienced extensive deforestation over the past two centuries as a result of conversion to agriculture and, to a lesser extent, urbanization. Thus only about a quarter of the region remains forested and has been highly fragmented (Twedt and Loesch, 1999; Gardiner and Oliver, 2005; Oswalt, 2013). Extant forests are on more flood-prone soils with forest types predominated by sweetgum-Nuttall oak-willow oak, sugarberry-hackberry-elm-green ash, overcup oak-water hickory, and baldcypress-water tupelo (Oswalt, 2013).

2.2. Data and band normalization

We obtained 11 Landsat 5 TM scenes (<http://earthexplorer.usgs.gov>) from October or November 2011 that had < 10% cloud cover. Images were selected to capture the landscape after most crops have been harvested yet trees still had leaves. These images were

normalized to account for differences in sun illumination, geometry, and atmospheric effects following procedures outlined by Chander et al. (2009) and Thenkabail (2009). These normalized images were clipped in ArcGIS 10.1 (Environmental Systems Research Institute [ESRI], Redlands, CA) with the Worldwide Reference System - Landsat descending shapefile (<http://landsat.gsfc.nasa.gov>) to remove corrupt edges of imagery where all Landsat sensors did not overlap.

2.3. Calculate band ratios and transformations

We calculated spectral band ratios (TM2/TM3, TM7/TM2, TM3/TM4, TM3/TM5, TM4/TM3, TM4/TM5, TM5/TM7) and band transformations (Land Water Mask [LWM], Normalized Difference Vegetation Index [NDVI], Green Normalized Difference Vegetation Index [GNDVI], Specific Leaf Area Vegetation Index [SLAVI], Normalized Difference Water Index [NDWI], and Tasseled Cap: Brightness, Greenness, and Wetness) from Landsat 5 TM imagery (Fig. 2).

Band ratios (i.e., band division) when applied to multi-spectral imagery tends to negate environmental effects on spectral values, as well as enhances differences in reflectance of certain materials (Richards and Xiuping, 2006). Band transformations, such as Tasseled Cap, were used to highlight particular features (e.g., vegetation health, vegetation moisture, or soil.

moisture) and properties by combining multiple spectral bands (Crist and Kauth, 1986).

2.4. Segment, train, and classify imagery

Developed areas were identified and delineated, before conducting object-based image analysis, using a morphological open tool (ERDAS Imagine 2014 Version 2013; Hexagon Geospatial, Norcross, GA) with a 5×5 pixel windows on 2006 National Land Cover Data (NLCD; “developed” classes 21, 22, 23, and 24) to

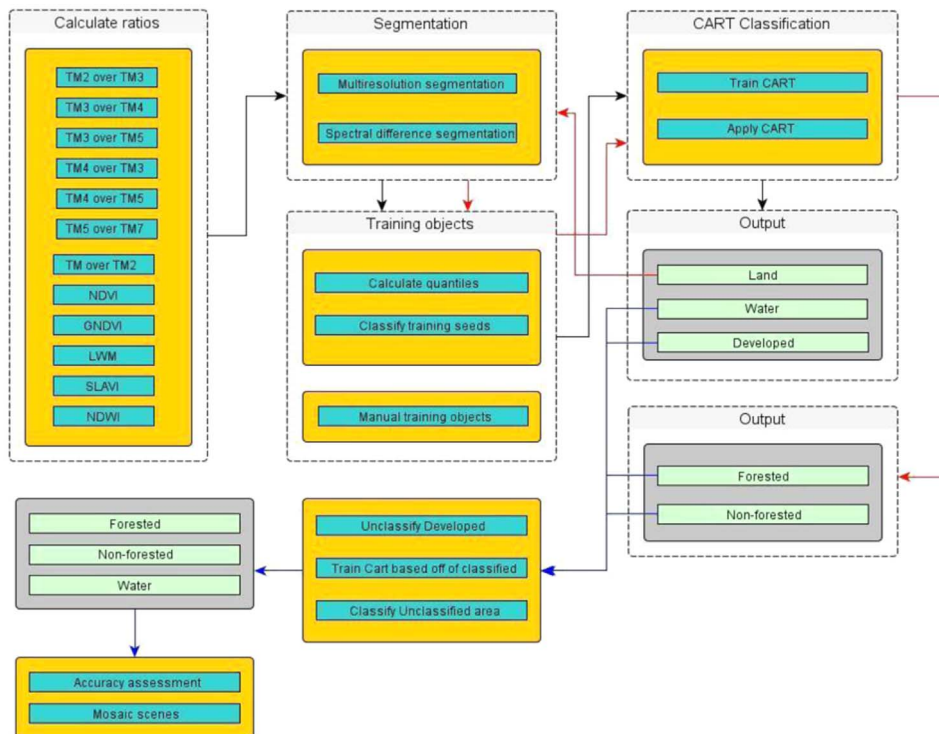


Fig. 2. Flowchart showing the classification process. Start at 'Calculate ratios' and follow the black arrows. Upon reaching 'Output', continue through step 2 by following red arrows. Again, upon reaching 'Output', follow blue arrows through completion. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

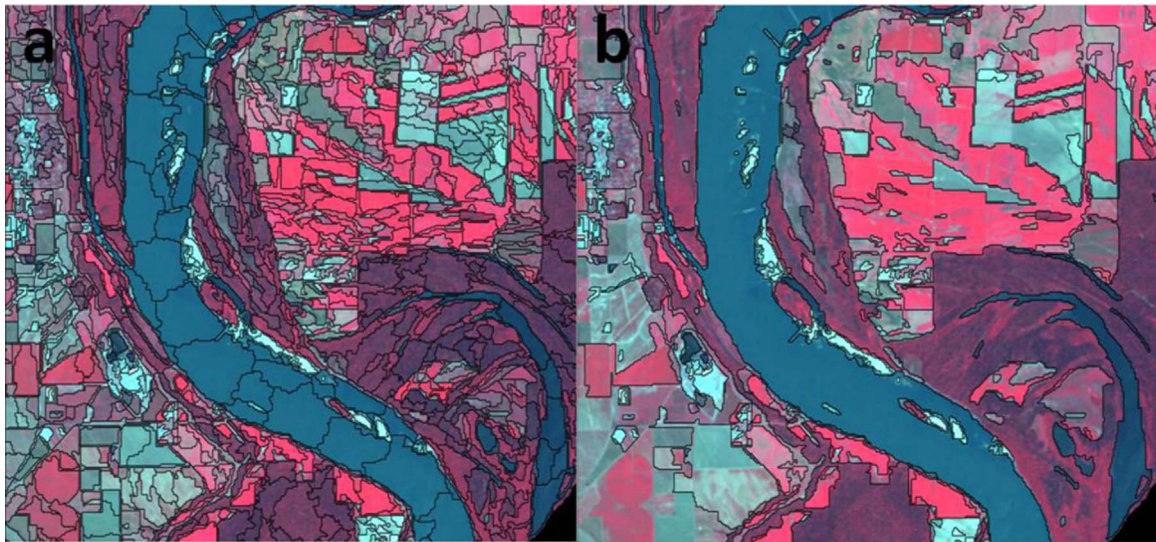


Fig. 3. Two-step segmentation process in which (a) all objects are initially segmented, and (b) dissolved based on spectral difference.

include large developed areas (cities) while omitting small developed areas and thin linear features (e.g., roads). These developed areas were masked and were not included in the initial training and classification of TM imagery. We supplemented the resultant developed areas by manually delineating any developed areas observed on 2013 high resolution imagery (ESRI World Imagery) that were not classified as developed on 2006 NLCD.

The initial step (multi resolution segmentation) of a two step segmentation process used a bottom-up, pairwise regional merging technique in eCognition to segment undeveloped areas into objects. This approach extracted features that were characterized not only by their spectral signatures but also by their shape. The resultant output effectively delineated agricultural fields and forest patches but each field or forest patch consisted of many small objects. Therefore, a second segmentation process, also using eCognition, was used to merge these small objects based on their spectral similarities, into contiguous and identifiable features (Fig. 3).

We classified land and water training sets used in object-based image analysis from normalized TM spectral bands, TM band ratios, and TM band transformations. To classify our training water dataset we selected all objects that had Land Water Mask and near-infrared mean band values within the lower 1% of their respective indices values across the entire scene. As objects representing open water were relatively homogenous within these spectral index bands, this lowest percentile index value adequately characterized open water. Conversely, to adequately characterize the land data training set we selected all objects that had mean Land Water Mask values in the upper 60% of scene values and mean NDVI in the upper 20% of scene values.

Each scene was manually inspected to verify that training objects were accurately representing their respective class. Training objects that represented an incorrect class were removed from the training set. On inspection, training objects were manually added based on the analyst's interpretation of the need for additional objects within the training set. For example, the boundaries of open water (typically marsh) were not classified in either training set and therefore representative objects within these areas were manually added to either the land or water training dataset. Training objects were used as input into the Random Forest classifier to classify all non-training objects. This resulted in a binary classification (land or water) for all non-developed areas within each Landsat TM scene.

2.5. Segment and classify imagery (forest vs. non-forest)

The objects classified as land were again segmented to separate and classify them as either forest or non-forest using the same procedures we used to classify land and water. Thus, for areas classified as land, all spectral band, band ratio, and band indices object data were used to train the Random Forest classifier. Mean values for SLAVI, TM7/TM2, Near Infrared, and Tasseled cap brightness were used to separate the classes and create training objects. Upon completion of the classification (forest, non-forest, and water) we used all classified objects as a training dataset to inform the Random Forest classifier to classify developed areas which were heretofore masked from classification.

Because recently planted and regenerating forests have spectral characteristics similar to grasslands or agricultural lands, we incorporated spatially explicit data obtained from local, state, and federal agencies that depicted areas of forest restoration. These data included lands enrolled in the U.S. Department of Agriculture's Wetland Reserve Program (WRP) and Conservation Reserve Program (CRP). All of these areas were classified as forest, regardless of the designation assigned through our classification process.

2.6. Accuracy assessment

Classified output images were merged and accuracy of forest, non-forested, and water classification was assessed within the boundaries of the Mississippi Alluvial Valley using ERDAS Imagine (Version 2013; Hexagon Geospatial, Norcross, GA). We evaluated 500 random points (160 forest, 240 non-forest land, and 100 water) by visually comparing classified value with habitat observed on Google Earth imagery (Google Earth 7.1, October 2013). Accuracy assessments produced an error matrix from which we calculated overall accuracy, producer's accuracy (the probability that a land-cover of an area is classified as such: errors of omission), user's accuracy (the probability that the classification actually matches the true land-cover type: errors of commission) and Cohen's Kappa statistic. The proportion of forest area was adjusted for accuracies in classification following recommendation of Olofsson et al. (2013).

Table 1

Classification (rows) and habitat observed on Google Earth imagery (Google Earth 7.1, October 2013; columns) for 500 random points distributed among objects classified as forest (160 points), non-forest land (240 points) and water (100 points).

	Forest	Land	Water	Classified	Classified area (ha)
Forest	157	3	0	160	3,141,847
Land	8	231	1	240	6,749,188
Water	0	1	99	100	397,370
Observed	165	235	100	500	10,288,405

3. Results

Our overall classification accuracy was 96.9% (Table 1). The accuracy assessments for forest, non-forest land, and water were similar with respective user's accuracies of 98.1%, 96.3%, and 99.0%. Respective producer's accuracies were 93.2%, 99.0%, and 93.3% and Kappa statistics were 0.97, 0.93, and 0.99. After adjusting the area classified as forest to account for classification error, our results indicated 3,307,910 ± 85,344 ha ($\bar{x} \pm SE$) within the Mississippi Alluvial Valley as forest, which included forest within developed areas (Fig. 1). The reforestation dataset that was incorporated into the forest classification totaled 551,307 ha, of which 159,322 ha would have otherwise been designated as non-forest within our classification.

4. Discussion

We found object-based image analysis combined with Random Forest classification was an effective method of classifying forest across a large landscape and was accurate based on comparison to known imagery. This methodology provides conservation planners and land managers an effective tool to evaluate conservation efforts. Furthermore, utilizing decision trees provide a transparent, repeatable and easily interpreted process to facilitate regular updates and assessments of forest area and their spatial distribution.

Under the auspices of programs such as CRP, WRP, and other conservation programs, there has been extensive, successful forest restoration throughout the Mississippi Alluvial Valley (Stanturf et al., 2000; King et al., 2006). In our classification, we attempted to capture reforestation without introducing errors of commission by including agriculture that was spectrally similar to reforestation. Even so, we found it difficult to separate newly reforested land from agricultural land. Forest restoration sites may take > 5 years before they can be spectrally distinguished from agriculture and another 10 years before they can be separated from scrub-shrub habitats when visualizing these areas using moderate resolution imagery. Therefore, to ensure young forest habitats are classified as forest, we recommend continued reliance on spatially defined reforestation databases that are incorporated as forest post-classification.

Using object-based image analysis and a Random Forest classification allowed us to classify multiple Landsat scenes with no modification to the rule-set; however, refinement of our rule-set may improve future iterations of our analyses and allow for national scalability. The rule-set we created utilizes only a few tools and data types, but through further testing and implementation this process may become increasingly automated and accurate. Our rule-set could also be expanded to incorporate other data (LiDAR, elevation, hydrology) to enhance classification results and further increase accuracy but this would come at the expense of greater processing time. Since we are working with such a large area and were only interested in three classes (forest, non-forest

and water) we were successful using imagery from one point in time (October–November 2011).

More complicated forest classifications, such as forest type, likely would require multi-temporal data to incorporate greater spectral variability within each class. Even so, the object-based forest classification methods could be used for this more complex classification. Adding additional data will increase calculation time and, depending on desired classification, may not significantly increase classification accuracy.

Although including data used in training a classifier within an accuracy assessment may bias estimated accuracy, we chose to run the assessment on the entire classification because our training dataset was not defined from known sample points. Rather our training set consisted of more broadly defined objects that were created through our rule-set.

Regular assessment of land cover and habitat types provide conservation partners the ability to understand the effects of conservation actions. Within the Mississippi Alluvial Valley, period assessment of bottomland forest cover is essential for evaluated progress towards habitat conservation goals. Thus, accurate classification of forest cover is required for conservation planning, including efforts that prioritize areas for reforestation to increase forest cover area. Using our object-based forest classification methodology, conservation partners will be able to quickly classify forest on the landscape allowing for the ability to track gains and losses in forest cover over time and at more timely intervals than the 5–10 year assessments currently available via National Land Cover classifications.

Acknowledgments

We thank the Lower Mississippi Valley Joint Venture, U.S. Fish and Wildlife Service, U.S. Geological Survey, American Bird Conservancy, and Ducks Unlimited for their financial and logistical support. Reference to commercial products or trade names does not imply U.S. Government endorsement.

References

- Akar, Ö., Güngör, O., 2012. Classification of multispectral images using random forest algorithm. *J. Geod. Geoinf.* 1 (2), 105–112.
- Bechtold, W.A., Patterson, P.L., 2005. The Enhanced Forest Inventory And Analysis Program - National Sampling Design and Estimation Procedures. General Technical Report SRS-80. U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC, p. 85.
- Blaschke, T., Strobl, J., 2001. What's wrong with pixels? some recent developments interfacing remote sensing and GIS. *GIS Z. Geoinf. Syst.* 6, 12–17.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. *Classification and Regression Trees*. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA.
- Breiman, L., Cutler, A., 2005. Random Forests. https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm (accessed 2.10.14).
- Chander, G., Markham, B.L., Helder, D.L., 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sens. Environ.* 113, 893–903.
- Crist, E.P., Kauth, R.J., 1986. The Tasseled Cap de-mystified. *Photogramm. Eng. Remote Sens.* 52 (1), 81–86.
- Cufi, X., Munoz, X., Friexenet, J., Marti, J., 2002. A review of image segmentation techniques integrating region and boundary information. *Adv. Imaging Electron Phys.* 120, 1–39.
- Friedl, M.A., Brodley, C.E., 1997. Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* 61, 399–409.
- Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N., Wickham, J., 2011. Completion of the 2006 national land cover database for the conterminous United States. *PE&RS* 77 (9), 858–864.
- Gardiner, E.S., Oliver, J.M., 2005. Restoration of bottomland hardwood forests in the Lower Mississippi Alluvial Valley, U.S.A. In: Stanturf, J.A., Madsen, P. (Eds.), *Restoration of Boreal and Temperate Forests*. CRC Press, Boca Raton, FL, pp. 235–251.
- Immitzer, M., Atzberger, C., Koukal, T., 2012. Tree species classification with random forest using very high spatial resolution 8-Band Worldview-2 Satellite Data.

- Remote Sens. 4, 2661–2693.
- King, S.L., Twedt, D.J., Wilson, R.R., 2006. The role of the Wetland Reserve Program in conservation efforts in the Mississippi River Alluvial Valley. *Wildl. Soc. Bull.* 34, 914–920.
- Lower Mississippi Valley Joint Venture Management Board, 2013. Lower Mississippi Valley Joint Venture Operational Plan. Vicksburg, Mississippi.
- Mellor, A., Haywood, A., Stone, C., Jones, S., 2013. The performance of random forests in an operational setting for large area sclerophyll forest classification. *Remote Sens.* 5, 2838–2856.
- Olofsson, P., Foody, G.M., Stehman, S.V., Woodcock, C.E., 2013. Making better use of accuracy data in land change studies: estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* 129, 122–131.
- Oswalt, S.N., 2013. Forest Resources of the Lower Mississippi Alluvial Valley. General Technical Report SRS-GTR-177. USDA Forest Service, Southern Research Station, Asheville, NC, p. 29.
- Pal, M., Mather, P.M., 2003. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sens. Environ.* 86, 554–565.
- Richards, J.A., Xiuping, J., 2006. *Remote Sensing Digital Image Analysis: An Introduction*, 4th ed. Springer, Germany.
- Rudis, V.A., Birdsey, R.A., 1986. Forest resource trends and current conditions in the lower Mississippi Valley. Resource Bulletin SO-116. U.S. Department of Agriculture Forest Service, Southern Forest Experiment, New Orleans, LA, pp. 7.
- Stanturf, J.A., Gardiner, E.S., Hamel, P.B., Devall, M.S., Leininger, T.D., Warren Jr., M.L., 2000. Restoring bottomland hardwood ecosystems in the Lower Mississippi Alluvial Valley. *J. For.* 98 (8), 10–16.
- Thenkabail, P.S., 2009. Satellite Sensor Data Normalization Issues. U.S. Geological Survey. Lecture at Boston University, Boston, USA. October 27–29, 2009. <http://www.pancroma.com/downloads/General%20Landsat.pdf> (accessed 28.03.14).
- Twedt, D.J., Pashley, D., Hunter, W.C., Mueller, A.J., Brown, C., Ford, R.P., 1999. Partners in Flight Bird Conservation Plan for the Mississippi Alluvial Valley (Physiographic Area # 05). http://www.partnersinflight.org/bcps/plan/MAV_plan.html (accessed 28.03.14).
- Twedt, D.J., Loesch, C.R., 1999. Forest area and distribution in the Mississippi Alluvial Valley: Implications for breeding bird conservation. *J. Biogeogr.* 26, 1215–1224.
- U.S. Fish and Wildlife Service, 2008. Strategic Habitat Conservation Handbook: A Guide to Implementing the Technical Elements of Strategic Habitat Conservation (Version 1.0). Report from the National Technical Assistance Team, 11 February, 2008. <http://www.fws.gov/landscape-conservation/pdf/SHCHandbook.pdf> (accessed 2.10.14).