Evaluation of the contribution of LiDAR data and postclassification procedures to object-based classification accuracy

Diane M. Styers
L. Monika Moskal
Jeffrey J Richardson
Meghan A. Halabisky
Evaluation of the contribution of LiDAR data and postclassification procedures to object-based classification accuracy

Diane M. Styers,¹ L. Monika Moskal,² Jeffrey J Richardson,² and Meghan A. Halabisky²

¹Western Carolina University, Department of Geosciences and Natural Resources, Stillwell Building 331, 246 Central Drive, Cullowhee, North Carolina 28723, United States
²University of Washington, School of Environmental and Forest Sciences, Remote Sensing and Geospatial Analysis Laboratory, Bloedel357/389, 4000 15th Avenue NE, Seattle, Washington 98195, United States

Abstract. Object-based image analysis (OBIA) is becoming an increasingly common method for producing land use/land cover (LULC) classifications in urban areas. In order to produce the most accurate LULC map, LiDAR data and postclassification procedures are often employed, but their relative contributions to accuracy are unclear. We examined the contribution of LiDAR data and postclassification procedures to increase classification accuracies over using imagery alone and assessed sources of error along an ecologically complex urban-to-rural gradient in Olympia, Washington. Overall classification accuracy and user’s and producer’s accuracies for individual classes were evaluated. The addition of LiDAR data to the OBIA classification resulted in an 8.34% increase in overall accuracy, while manual postclassification to the imagery + LiDAR classification improved accuracy only an additional 1%. Sources of error in this classification were largely due to edge effects, from which multiple different types of errors result. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.JRS.8.083529]

Keywords: accuracy; classification; land cover; LiDAR; object-based image analysis.

Paper 14184 received Mar. 27, 2014; revised manuscript received Aug. 12, 2014; accepted for publication Oct. 6, 2014; published online Nov. 6, 2014.

1 Introduction

Creators of LULC classifications have long utilized remote sensing data and methods due to the broad spatial coverage of aerial photographs and satellite imagery versus ground-based inventories. However, the methods used in their creation have dramatically changed over time as advancements in sensor technology have allowed for increased spectral, spatial, and temporal resolutions of base imagery. A variety of pixel-driven analysis methods have historically been the industry standard for creating LULC classifications and may still be adequate for use with moderate to low resolution imagery, in homogeneous landscapes, or in areas where high resolution imagery is costly or nonexistent.

Object-based image analysis (OBIA) methods have recently proven more appropriate for LULC mapping, especially with high resolution imagery and in heterogeneous landscapes such as urban areas where a greater level of classification detail is essential. This form of feature extraction allows the use of additional variables such as shape, texture, and contextual relationships to classify fine image features, and has been shown to improve classification accuracy over pixel-based methods. Application of the OBIA technique is no longer an exploratory research direction; rather it has been implemented operationally and continues to grow in popularity among urban forestry practitioners and land use planners. OBIA-driven
classifications allow users greater information extraction from these “maps” than just the distribution and percentage of various classes. These highly detailed maps could prove useful for monitoring urban forest stand health and extracting landscape-scale indicators of ecosystem health. Thus, an evaluation of the opportunities this approach provides is crucial.

Land use/land cover (LULC) classifications created using OBIA methods on high resolution imagery has not been without complications, especially in cases where either spectral or spatial resolution is still a major limiting factor of the imagery. For example, using imagery that has a high spatial resolution (<1 m) but a low spectral resolution (RGB only, or no IR bands) can make it difficult to differentiate different types of vegetation, such as trees, shrubs, and grasses, or between different impervious surfaces, such as roads and buildings. Even using imagery with one or more IR bands but collected at a time of year when all vegetation is photosynthetically active, it can be challenging to distinguish between different vegetation classes. In high-density urban areas, or landscapes having dramatic topography, separating these classes using high spectral/high spatial resolution imagery can be quite problematic due to shadowing. Thus, including remotely sensed data capable of capturing topography, canopy, and building heights in an LULC classification is a logical method for improving accuracies of imagery-driven data analysis.

Light detection and ranging (LiDAR) datasets have repeatedly been used to produce accurate terrain models and height measurements of forest canopy, buildings, and other landscape structures through active sensing of three-dimensional (3-D) objects on the ground surface. Early applications of LiDAR data included topographic models for high resolution bare ground digital elevation models, vegetation height and cover estimates, and forest structural characteristics including leaf area index and aboveground biomass. Although OBIA methods were originally developed for use with imagery, new research has begun to combine LiDAR data with high resolution imagery for mapping urban street trees obscured by building shadows, ground impervious surfaces, floodplains, canopy cover, forest stand structure, and transportation and utility corridors. Further, the addition of LiDAR data to an OBIA classification algorithm has been shown to improve classification results, especially for distinguishing between trees, shrubs, and grasses, and also between buildings and ground impervious surfaces. As LiDAR technologies continue to improve, allowing for faster and more cost-effective data acquisition, it is expected that the availability of public-domain LiDAR data will improve, and thus, the need to understand how these data can enhance OBIA-driven mapping efforts is essential.

Often a certain amount of manual postclassification takes place during the classification process, no matter which method is used. In the case of OBIA, postclassification typically involves reclassifying segments classified in error to the correct class based on the analyst’s knowledge of the region and level of image interpretation skills. However, we currently do not know how much manual postclassification contributes to classification accuracies. Image interpretation skills are an integral part of the OBIA method, so it is reasonable to suggest that taking a certain amount of time to utilize these expert skills to address the most obvious errors should increase the producer's accuracy of the classes as well as the overall classification accuracy. This operator-machine connection is the key technological concept behind OBIA, and it is this ability to emulate the human concept of context that drives the object segmentation process. Utilizing the strengths of the OBIA method and the expertise of humans, the classification should result in a detailed, accurate, high quality dataset useful for a variety of applications.

Based on what we have learned from this and other OBIA-based projects in other study areas, we would like to suggest a methodology framework that others, especially new users, might consider using. We believe a comprehensive classification can be created by a trained individual using freely available high spatial resolution imagery, LiDAR data, and remote sensing software packages. While others have shown the applicability of nonpublicly available data, or data flown for a specific mission, we demonstrate that these data can be used for other (postmission) purposes. None of the data used in this study were specifically flown for this project, or even for the purpose of LULC mapping. The overall goal of this project was to test a commonly used OBIA method for creating an LULC classification using imagery and LiDAR (I + L) data, examine overall and class accuracies, and determine any sources of
error observed in the resulting product. Based on the above goal, the specific objectives of this study were to

1. evaluate how much LiDAR data contributes to classification accuracies,
2. assess how much impact manual postclassification has on LULC product accuracies, and
3. propose an object-based methodology framework for classifying public-domain remotely sensed data for postmission applications in urban environments.

2 Methods

2.1 Study Area

The study area is a ∼6.54-hectare transect along a rural to urban gradient, located within the city of Olympia, Washington, USA (Fig. 1). The city of Olympia is located at the southern end of Puget Sound and is experiencing a moderate population growth with a population that is expected to double in 30 to 40 years. Puget Sound provides numerous ecosystem services to the region’s residents and Olympia’s urban tree canopy, detectable through LULC mapping, plays a vital role.

2.2 Remote Sensing Datasets

In this work, we used 16-bit, four-band, color infrared imagery acquired in March of 2008 with a pixel resolution of 0.12 m (0.4 feet). On March 2 to 3, 2008 discrete-point LiDAR data were also acquired over the study area with the Optech 3100 sensor at a survey altitude of 899.16 m (2950 feet). Accuracy of the LiDAR data conforms to Puget Sound LiDAR Consortium standards with an absolute vertical accuracy of 0.03 m RMSE (0.11 feet RMSE). The data point cloud was collected at 4 pulse returns with a density of 4.66 first returns/m². Figure 2 shows the various images representing the remote sensing datasets used in this study.

Fig. 1 Map of study area: South Puget Sound, Olympia, Washington, USA.
We generated a 1-m per-pixel spatial resolution LiDAR data-based digital terrain model (DTM) using the freely available FUSION software developed by the USDA Forest Service, Pacific Northwest Research Station. To accomplish this, we used the groundfilter algorithm within FUSION to filter the likely ground points from the LiDAR point cloud, and then created a smoothed raster from that point cloud at a 1-m resolution. We then applied the ground model to produce the same resolution standardized canopy height metrics which, in conjunction with the ground model, were the key LiDAR-based variables in our OBIA-driven classification.

### 2.3 Land Use/Land Cover Classification

#### 2.3.1 Developing a basic classification

We originally developed an OBIA algorithm for a separate LULC classification project in Seattle, Washington and wanted to see if it would work in different locations within the same ecoregion (i.e., urbanized Pacific Northwest coastal areas) with limited modification. The LULC classes we selected are prominent throughout the greater Pacific Northwest coastal region and although certain classes are under-represented in this specific study area (e.g., intertidal and bare ground), we wanted to keep these classes in the algorithm for future OBIA analyses in different study areas across the Pacific Northwest coast.

Two different LULC classification approaches were used in this study—one approach used imagery alone (IA) while the second used I + L data. We used Trimble’s eCognition®
software program to create a segmentation and basic classification for each of the two datasets (IA and I+L). In both cases, the classification process began with a multiresolution segmentation of the four-band aerial imagery, using a scale parameter of 700, a shape parameter of 0.1, a compactness parameter of 0.5, and layer weights of blue = 1, green = 1, red = 2, and near-infrared = 2 (Fig. 3). These specific parameters were selected based on skilled visual interpretation of the image segmentations that resulted from using different scale parameters. Although we do recognize the inherent subjectivity in segmentation scale selection and the resulting image object size, an in-depth discussion on that topic is beyond the scope of our study. Several other published studies have focused on segmentation scale selection optimization and have documented how widely these parameters can vary depending on the object(s) of interest.

We used a combination of logical values, the feature space optimization tool, and two-dimensional (2-D) feature space plots in eCognition®, as well as expert interpretation to select threshold values. LiDAR data thresholds were assigned based on logical height cutoff points for distinguishing between tall and short objects. The mean brightness of each object, computed as the mean of the four bands, was used to classify buildings and shadows at each brightness extreme (i.e., light and dark). Normalized difference vegetation index (NDVI) thresholds were based on logical breakpoints between photosynthetically active tree vegetation, senescent grass, and nonphotosynthetic impervious surfaces. Feature space optimization and 2-D feature space plots allowed us to find appropriate thresholds using brightness for rooftops, darkness for shadows, darkness in the near-infrared for water, darkness in near-infrared for bare ground, and brightness in the blue band for impervious surfaces. Finally, the classification was further improved by implementing several hierarchical rules to remove small misclassified objects enclosed by other objects (e.g., cars on freeways).

2.3.2 Classification using imagery alone

At the end of the processes described above, we had a basic classification based solely on imagery; however, it contained multiple objects misclassified as shadows, did not discriminate well between buildings and ground impervious surfaces, and misclassified intertidal areas as ground impervious. Several extra rules were written into the algorithm to address these errors.

---

**Fig. 3** Decision tree for the initial basic classifications for both imagery alone (IA) and imagery and LiDAR.

**Fig. 4** Decision tree for the IA classification, including parameter thresholds (additional rules to initial basic classification shown in bold italic).

---
The rule relative border to shadows (>0.15) was used to identify and classify buildings incorrectly assigned to the ground impervious class. This worked well for us because buildings cast shadows, but ground impervious surfaces do not. Next, building shadows were classified by finding the object with the largest classified border, and classifying the shadow by that object’s class. A second rule for buildings was then run based on finding ground impervious objects with a large rectangular fit (>0.85) and a small length:width ratio (<5.0). This located objects that had a general square shape, and helped to find buildings that did not have large relative borders to shadows (such as those surrounded by trees). Intertidal areas misclassified as ground impervious surfaces were reclassified based on their relative border to water (>0.15). Finally, bare ground was located by performing a quad tree segmentation on the ground impervious surfaces layer and classifying objects with an NDVI value greater than −0.05 as bare ground. A final set of filtering algorithms filtered out any remaining small misclassified objects.

2.3.3 Classification using imagery and LiDAR data

We also needed to apply additional rules to an I + L classification to compensate for objects classified in error (Fig. 5). The median difference between the maximum height model and the ground model was computed for all objects. This value represented a good approximation of the height of each object above ground, and a threshold of 1.8 m (6 feet) was used to separate tall objects such as buildings and trees, from short objects such as grass and ground impervious surfaces. Shadows were first classified into temporary classes based on median difference. Using an NDVI threshold (−0.4) tall temporary objects were classified as trees or buildings, and short temporary objects as grass or ground impervious. A chessboard segmentation was then run on the grass and tree classes, followed by a reclassification based on median height. The median height value was better at explaining differences between these two classes because irrigated lawns created confusion when using NDVI alone. Median height was used to reclassify ground impervious surfaces and buildings. Intertidal areas misclassified as ground impervious were reclassified based on their relative border to water (>0.15). As in the IA classification, bare ground was located using a quad tree segmentation and an NDVI threshold (−0.05) of the ground impervious class. Lastly, a set of filtering algorithms removed small misclassified objects.

2.3.4 Manual postclassification of land use/land cover classifications

Efforts to improve the classifications included systematic manual reclassification; the analyst deployed the postclassification in a gridded pattern from west to east and north to south to achieve this process. Postclassification for the IA classification took longer than for the
I + L classification, as it contained more errors. The IA postclassification took approximately 12 h to complete, while the I + L postclassification took only a third of the time (∼4 h). The greatest time sink for manual postclassification in both datasets was correcting confusion between driveways and houses. The segmentation included small sections of driveway with houses, forcing us to manually reclassify driveways as ground impervious rather than as buildings. There was not a good way to differentiate between ground impervious and buildings in the IA classification, and it was time consuming to postclassify areas where there were residential developments. It was also difficult to separate grass and trees in the IA classification, so we spent a good bit of time manually reclassifying irrigated lawns as grass rather than trees.

Although it might be relatively easy to adjust the algorithm to better distinguish between driveways and houses using LiDAR data (i.e., making the segmentation size smaller and relying on height differences), it would be challenging to do for the IA classification. For this imagery, there is not enough spectral information to differentiate trees from grass and buildings from ground impervious surfaces. Texture might have been able to help discriminate the trees, but it is so memory intensive (i.e., it would have taken weeks to run the algorithm) it was not used in this project. Further, the amount of time it would have taken the texture algorithm to run could likely equal or exceed the amount of time it took to manually reclassify objects, and improvement in classification accuracy is not as certain. Our approach to test the postclassification impacts on the OBIA-driven LULC classification results was fairly simple; however, it provides a general overview of how postclassification affects accuracy.

2.4 Accuracy Assessment

An accuracy assessment was conducted for the four different classifications: IA and I + L, each with and without postclassification. Accuracy assessment methods were based on the guidelines of Foody2,56 and Congalton and Green.57 First, the area of each class was quantified using polygons from the I + L segmentation algorithm (without manual postclassification) for all tiles in study area. These were used to determine the number of points that should fall in each class. Buffering was not used as we did not want to bias our analysis toward large objects which tend to be classified correctly by OBIA methods. We then generated a total of 300 stratified random points, which were distributed proportionally by area for each class, for visual accuracy assessment of the classifications by an expert image analyst. Each point was then visually classified by a trained image analyst using 12 cm color infrared imagery and a canopy/building surface model created using the LiDAR data. The classes used are described in Table 1.

<table>
<thead>
<tr>
<th>Classes</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bare ground</td>
<td>Point is over a pervious surface that does not contain vegetation</td>
</tr>
<tr>
<td>Buildings</td>
<td>Point is over an impervious surface that is built up vertically above the ground at least 3.048 m (10 feet) tall</td>
</tr>
<tr>
<td>Grass</td>
<td>Point within low vegetation. Can be irrigated lawns, nonirrigated lawns, or scrubby areas that do not contain woody plants greater than 0.91 m (3 feet) tall</td>
</tr>
<tr>
<td>Impervious</td>
<td>Point is over an impervious surface that is not built up vertically</td>
</tr>
<tr>
<td>Intertidal</td>
<td>Point within muddy or rocky area directly adjacent to water. It is assumed that these areas would be submerged under high tide</td>
</tr>
<tr>
<td>Shrubs</td>
<td>Point within vegetation that is less than 4.57 m (15 feet) tall and greater than 0.91 m (3 feet) tall</td>
</tr>
<tr>
<td>Trees</td>
<td>Point within vegetation that is greater than 4.57 m (15 feet) tall</td>
</tr>
<tr>
<td>Water</td>
<td>Point within ocean, lake, or pond. Must be liquid water. Vegetation floating on water is still classified as water</td>
</tr>
</tbody>
</table>

Table 1 Land use/land cover class descriptions.
3 Results and Discussion

3.1 Accuracy for Imagery Alone Classification

The OBIA method produced sufficient results using the aerial IA; overall accuracy was 71.3% (khat = 0.59). The method was effective at extracting trees/shrubs, and to a lesser degree, ground impervious surfaces (e.g., roads, driveways, paved parking lots). With the exception of water, tree canopy was classified most effectively, and had a producer’s accuracy of 92.7% and a user’s accuracy of 77.4% (khat = 0.59). Since the producer’s accuracy was greater than the user’s accuracy, these results suggest that we have captured the finest amount of detail in our classification these data will allow. This also indicates that the classification algorithm has overclassified tree canopy, and points were generally misclassified as grass. The ground impervious class had a producer’s accuracy of 66% and user’s accuracy of 75.6% (khat = 0.71), indicating we may have been able to optimize the algorithm further to achieve a higher producer’s accuracy. Ground impervious surface points were misclassified across all classes except water. Producer’s accuracy for the remaining classes was near or below 50%, excluding water.

3.2 Accuracy for Imagery Alone Classification with Postclassification

The overall classification accuracy slightly improved to 74.7% (khat = 0.56). Again, the method appeared to be effective at extracting tree canopy and ground impervious surfaces. Producer’s accuracy was greater than user’s accuracy for the tree canopy and building classes. Although producer’s accuracy decreased by <1% for tree/shrubs, there was an increase in producer’s accuracy for the grass class, improving from 36% to 42%. Manual postclassification also increased producer’s accuracy for ground impervious (from 66% to 72.3%), buildings (from 52% to 68%), and water (from 96% to 100%).

3.3 Accuracy for Imagery and LiDAR Classification

The addition of LiDAR data to the LULC classification process improved the overall classification accuracy to 79.7% (khat = 0.77), and enabled us to separate the tree/shrub class into two distinct tree and shrub classes. Interestingly, this method appeared to be more effective at extracting grass and buildings than trees and ground impervious surfaces, as was the case with the IA classification algorithm. The grass class had a producer’s accuracy of 90%, but a user’s accuracy of 64.3%, as some of these points still appear as trees, shrubs, or even bare ground. Producer’s and user’s accuracies for the building class were equal at 88%, and demonstrate the improvement in accuracy the LiDAR data provided for this class. For both tree and ground impervious classes, user’s accuracy was greater than producer’s accuracy, which was 92.7% versus 84.17% and 88.10% versus 78.7%, respectively. Although separated from the tree class, the shrub class had a low producer’s accuracy of 52.9% (user’s accuracy 50%; khat = 0.47).

3.4 Accuracy for Imagery and LiDAR Classification with Postclassification

The small amount of time taken to postclassify the I + L classification (~4 h) resulted in a 1% improvement of overall classification accuracy from 79.7% to 80.7%. Since the LiDAR data appeared to have cleared up most of the points classified in error on the IA classification, there were fewer points of error in the I + L classification. Manual postclassification slightly improved the producer’s accuracy of the ground impervious and water classes from 78.7% and 96% to 83% and 100%, respectively. The manual postclassification also had the benefit of resolving large, obvious errors that comprised a small percentage of the landscape, but were especially glaring when viewing the final map, thus improving the overall quality of the product.

3.5 Comparisons Between the Four Classifications

As is evident in Table 2, the overall accuracy increased with each classification iteration. Two of our main objectives were to determine the contribution of LiDAR data to classification
Manual postclassification was more time intensive for the IA classification; \( \sim 12 \) h of work resulted in a 3.3\% increase in overall classification accuracy. Most of this improvement was reflected in increases in the building and ground impervious class accuracies. When LiDAR data were added to the classification, overall accuracy increased an additional 5\%. Substantial increases in producer’s accuracy are observed in the grass, building, and ground impervious classes. The addition of LiDAR data enabled us to divide the tree/shrub class into separate tree and shrub classes, which resulted in a higher level of classification detail but lower class accuracies. We assume this is in part due to the confusion between low tree crown edges and actual shrubs, but also to the inability of the image interpreter to determine object height (i.e., trees versus shrubs) on the 2-D imagery used as ground truth. Although producer’s accuracy in the tree class decreased in the I + L classification, user’s accuracy significantly increased.

Manual postclassification of the I + L classification took \( \sim 4 \) h and improved the overall accuracy 1\%. The overall difference from the IA classification (without manual postclassification) and the I + LiDAR data classification (with manual postclassification) was 9.4\%, improving from 71.3\% to 80.7\% (Table 2). Manual postclassification did not significantly increase the accuracy of our classifications; however, we believe any amount of increase in accuracy is desirable, and based on the short amount of time it takes to do, it is worth the investment. Not only is the accuracy of the map improved but a more visually pleasing product is created as well. For example, in the absence of manual postclassification, we would have shown the State Capitol Building as covered and surrounded by water, which in actuality is merely a shadow cast by the Capitol rotunda (see Fig. 6).

### 3.6 Error Attribution

#### 3.6.1 Error in the imagery alone classification

Examination of each individual erroneous point helped us determine specific sources of error in the classifications. In the IA classification, there were discrepancies for 85 of the 300 total accuracy assessment points, most of which fell into four categories. The main source of misclassification, accounting for half of our error (43 points), was due to our inability to discriminate...
buildings versus ground impervious surfaces and trees versus grass using spectral values alone. Irrigated lawns, having high NDVI values which make it difficult to separate them from trees without height information, accounted for 27 of these erroneous points. Several of these points (as well as a few others) were confounded by errors related to spectral thresholds, where objects were brighter or darker in specific bands than they were expected to be. One example is buildings with asphalt roofs that had a similar brightness values to ground impervious surfaces and thus, were included in the ground impervious class. An additional 13 points were located on objects that fell in shadows, resulting in inaccurate classification of those objects. Eleven points were located along the edge, or borderline, of two classes; in these cases, the points have about an equal chance of being assigned to either class by both the image analyst and the image interpreter (i.e., the point could belong to either class, depending on individual opinion).

Several class-specific sources of error are also worth noting. The bare ground class had the highest error (0% accuracy), but that appears to be a consistent outcome in LULC mapping. A secondary source of class error is confusion between the water and intertidal classes. It would be relatively easy to include the intertidal areas in the water class, but since these areas are unique ecosystems high in biodiversity we wanted to keep them in a separate class. On an interesting note, one point classified as grass was actually a building that had a “green roof.” This could be an important feature to remember as the green roofs and living buildings gain popularity, especially in urban areas. Finally, 7 of the 85 error points were due to differences in image interpretation between the image analyst that created the classification and the image interpreter that conducted the accuracy assessment; 5 of these were water versus intertidal.

3.6.2 Error in the imagery and LiDAR classification

There were discrepancies for 61 of the 300 points in the I + L classification. A majority of these points also fit into four categories of error, but these are different types of errors than those observed in the IA classification. As expected, the addition of LiDAR data substantially helped differentiate spectrally similar objects of varying heights. However, the main sources of error were still due to confusion between trees, shrubs, and grass, and it would be more obvious if the shrub class did not have errors of its own. This error resulted from a combination of two very different issues: (1) the difficulty associated with assessing an object’s specific height based on a 2-D image and (2) the shrub class existing at the edge of two height classes and its relative narrow range in height values. The first of these represents an unavoidable incongruity associated with using 2-D imagery to ground truth a 3-D dataset. The second issue deals with edge effects, from which multiple different types of errors result.

We anticipated that adding LiDAR data to the image classification would help in distinguishing between trees, shrubs, and grasses, and buildings and ground impervious surfaces. The inclusion of LiDAR data did indeed help in the classification process, added a greater level of detail,
and improved the classification accuracies, but was not error-free. Several of these errors (6 of the 61 points) were directly due to edge effects associated with LiDAR data and the OBIA technique. An additional 28 “borderline” points could also be attributed to edge effects.

There are several ways in which edges affect how objects are classified (or misclassified) and contribute to error in the LULC classification. One of these deals with issues related to the combination of spatial resolution of the imagery used and the object’s size. If spatial resolution is low, pixel resolution is larger than an object’s edge. In other words, you will have mixed pixels within an object, no matter its size. For example, a single pixel (or object) may contain both the edge of a building and the edge of a sidewalk, leaving the algorithm to select whether that pixel (or object) is either a building or ground impervious surface. Spectral complexity, a second source contributing to edge errors, is an additional complication of using high spatial resolution imagery. Consider the roof of a downtown building: it is usually composed of different materials such as metal flashing and asphalt, and can contain objects like HVAC vents and communications equipment, may have planters, outdoor furniture, or parking for cars, and each of these objects and the building itself casts shadows. Although we can recognize the single, large object as a building based on its shape and context, the computer has a more difficult time “seeing” all of these smaller objects as multiple components of a single building object. A third issue we encountered related to object edges is when the outer edge of a tree canopy reaches near or touches the ground, it is classified as shrub or grass due to the height of the LiDAR data points returned from that location. We suggest that including some sort of crown segmentation could help with this problem. We also observed instances where there was a mismatch between datasets due to the radial distortion inherent in orthophotos (but not in LiDAR data) resulting in the canopy being slightly off. This problem has been well-documented in the literature and it has been suggested that LiDAR-based orthorectification can minimize these misalignment issues.

3.7 Other Issues

Other than those issues described above, the only other error apparent in our I + L classification is due to the presence of artifacts at the LiDAR data seams, or tile boundaries (see vertical artifact in Fig. [1]). This can be resolved by a few different tiling and stitching processes, depending on the size of the full dataset. One option is to stitch all of the tiles together to create one, large seamless tile. However, for large study areas, the stitched file can be over a terabyte and taxes the analytical capabilities of most desktop workstations. Another, more reasonable, approach is to make the LiDAR data tiles larger than the imagery tiles and merge overlapping areas to reduce edge effects and make the tiles appear seamless. We did not perform any tiling and stitching processes on these datasets because both the LiDAR data and imagery were tiled for us by the vendor. It was not until the classification processes were complete that this error became apparent. However, since this was a methods-based study, we did not attempt to reconcile these errors in the present classifications but rather have taken note of this as an issue to resolve in future studies.

3.8 Toward an Object-Based Methodology Framework

Based on the strengths and limitations we have recognized through the completion of this and other OBIA-based projects in other study areas, we would like to suggest a methodology framework for OBIA-based LULC classification development in urban areas. By following this framework, some of the errors present in this case study can be avoided. It is our aim to share what we have learned from these experiments and suggest ways to improve upon the methods we employed herein. We hope this information can serve as a starting point for others attempting an OBIA-based LULC classification for the first time, no matter their budget or experience level.

When planning to create an OBIA-based LULC classification, first check to see what remote sensing datasets already exist for your study area. OBIA methods can be used on free, publicly available, high spatial resolution imagery such as U.S. National Agriculture Imagery Program (NAIP) imagery, which has a national extent and has been successfully classified to achieve detailed LULC maps for urban planning, management, and scientific research.
imagery can be accessed through the United States Department of Agriculture (USDA) Geospatial Data Gateway. OBIA methods can also be used on high spatial resolution satellites with global coverage, such as Quickbird or GeoEye, although these data are not typically free of cost. Additionally, LiDAR data consortia, devoted to develop public-domain, high resolution LiDAR datasets would facilitate these efforts as well. These datasets can be accessed through the United States Geological Survey (USGS), the lead federal agency coordinating efforts toward a National LiDAR dataset.

It has been shown that LiDAR data is especially useful for classifying areas falling within shadows and distinguishing between objects of similar spectral reflectance but different heights. However, LiDAR data may not always be available. In this case, some of the issues inherent to using IA can be solved through the use of ancillary datasets. Many municipal agencies have previously invested a lot of time and money into creating parcel datasets, building footprints, transportation routes, utility corridors, flood zones, and other geospatial data layers. In many locales these datasets are freely available to the public, and when added to an LULC classification they not only help create visually pleasing products, but can also be utilized in class discrimination which can save the overall amount of time spent on classification development. It could be as simple as masking out a water class by using a coastline or open water body boundary file, or a utilizing a building footprint’s layer to help discriminate buildings from ground impervious surfaces such as roads, driveways, or patios.

Likely, a good bit of pre-processing will need to be completed prior to beginning the segmentation and classification processes. Industry standard pre-processing techniques include atmospheric and radiometric correction, georeferencing, and mosaicking images, as well as creating LiDAR-based DTMs, canopy height models, and other raster files for height categories of interest. Based on our experience, other pre-processing might include tiling and stitching imagery and LiDAR datasets to reduce edge effects and creating a tree crown segmentation layer to minimize misclassification of tree canopy edges as shrubs or grass.

The software you choose to use to create your classification will be the next decision. There are several software packages and add-on modules currently available to perform the OBIA method. Examples of these include the freely available SPRING software developed by Brazil’s National Institute for Space Research (INPE) and the industry-leading proprietary eCognition® software, a product of Trimble®. Although anyone can learn to use either of these or other OBIA-based software programs, there can be a greater learning curve with proprietary programs which typically have many more algorithm development options. For those interested in starting an OBIA project, we have created a guided analytical training workshop which is freely available online to anyone interested.

Postclassification is generally a part of the LULC classification creation process, no matter which method is used, especially when obvious errors might render the final product unaccept-able for the end user. For OBIA-based projects, postclassification usually involves reclassifying segments classified in error to the correct class based on the analyst’s expert skills and knowledge of the region. It should only take a short amount of time to manually postclassify OBIA-based classifications, making the increase in accuracy worth the time. It is certainly appropriate to capitalize on the intimate connection between operator and computer that is fundamental to the OBIA method, especially considering this procedure has no major implications to accuracy but does improve the quality of your classification.

Finally, an accuracy assessment is requisite to any LULC classification project. The authors have found the methods suggested in the book by Congalton and Green, collected from highly regarded sources in the remote sensing community, to be of the highest standards. There are many ways to conduct an accuracy assessment and these will vary according to the location and scope of the project and the type and resolution of the datasets used. Although no industry standard has been set, 85% seems to be the most universally acceptable accuracy level even though there is nothing specifically significant about this number and this threshold continues to be debated. Further, methods for conducting a statistically sound accuracy assessment for OBIA-based classifications are still currently being developed and refined, as disparate accuracy levels can result from using points versus polygons (i.e., objects) uncertainties about object boundary locations on the ground and object segment area and location goodness of fit. Last,
it is also worth noting that ground truthing your accuracy assessment in urban areas can be quite problematic since a large proportion of urban lands are private property, making accessibility to randomly generated ground control points a huge issue. In these cases, it is often necessary to substitute a ground truth survey with very high spatial resolution imagery.

4 Conclusions

This study quantifies the additional benefit of LiDAR data and manual postclassification important for initial planning of LULC mapping in urban areas. Data acquisition costs are an important factor when planning an LULC classification and jurisdictions may have trouble justifying the high cost of acquiring LiDAR data. Combining discrete-point high density aerial LiDAR data and OBIA-based classification methods have been shown to improve LULC classification accuracy, and is especially suitable for distinguishing between trees, shrubs, and grasses, and buildings and ground impervious surfaces. Results from this study corroborate those findings and indicate LiDAR data does indeed contribute to improved overall and class accuracies. When LiDAR data were added to our LULC classification, overall accuracy increased 5%. Substantial increases in producer’s accuracy were observed in the grass, building, and ground impervious classes. The addition of LiDAR data enabled us to divide trees and shrubs into separate classes, which resulted in a higher level of detail.

Manual postclassification also improved overall accuracy and improved visual appeal of the final product. Image interpretation is integral to the OBIA method, promoting an intimate connection between operator and machine. It is thus appropriate to capitalize on these strengths to correct obvious postclassification errors, since this procedure does not have major implications for accuracy but does improve the quality of the classification. It would take much more time and effort to train the algorithm to perform as well.

Acknowledgments

Funding for this research was provided through a Joint Venture Agreement between the University of Washington (UW) Remote Sensing and Geospatial Analysis Laboratory and the USDA Forest Service Pacific Northwest Research Station, as well as the UW Precision Forestry Cooperative through Corkery Family Chair Fellowships. We also thank the City of Olympia for providing the 2008 Olympia imagery and LiDAR datasets, flown by Watershed Sciences. Finally, we kindly thank the anonymous reviewers who helped us to improve this paper.

References


**Diane M. Styers** is an assistant professor in the Department of Geosciences and Natural Resources at Western Carolina University. She received her MA in geography from Georgia State University in 2005, PhD in forestry from Auburn University in 2008, and completed a postdoc in the University of Washington Remote Sensing and Geospatial Analysis Laboratory in 2011. Her research involves using remote sensing to analyze changes in ecosystem structure and function, particularly in response to disturbances.

**L. Monika Moskal** is an associate professor in the School of Environmental and Forest Sciences at University of Washington and principal investigator in the Remote Sensing and Geospatial
Analysis Laboratory. She received her MS and PhD in geography from University of Calgary (2000) and University of Kansas (2005). Her research involves developing tools to analyze remotely sensed data by exploiting the spatial, temporal, and spectral capabilities of data and using those to investigate landscape structure.

Jeffrey J Richardson is a postdoc in the Remote Sensing and Geospatial Analysis Laboratory at the University of Washington. He received his BS in creative writing and biology from Kalamazoo College in 2003, his MS in remote sensing from the University of Washington in 2008, and his PhD in interdisciplinary science from the University of Washington in 2011. His research involves solving problems related to forests, energy, and food with remote sensing and geospatial analysis.

Meghan A. Halabisky is a PhD student in the Remote Sensing and Geospatial Analysis Laboratory at the University of Washington, where she is advised by Dr. L. Monika Moskal. She completed concurrent degrees, a master of science and a master of public affairs, in 2010 at the University of Washington. Her research interests include development of high resolution remote sensing techniques for spatiotemporal analysis of wetland dynamics, landscape change, and climate change impacts.