Labs 4-5: Change Detection of Flooded Areas in Mali

Julia Janicki

Introduction

The goal for this project is to map flooded areas of an area in Mali after a flooding event. The methods that were chosen to map flooded areas are common methods used for change detection, including density slicing and supervised classification, in this case the Support Vector Machine (SVM) algorithm is used.

Density slicing is a digital data interpretation method to enhance the information gathered from an individual brightness band and it is done by dividing the range of brightnesses in a single band into intervals, then assigning each interval to a color. SVM is a supervised classification algorithm, where the identity and location of some of the land cover types are known beforehand and sites are labeled as specific classes to create a training dataset. The sites are evaluated for separability and the supervised classification algorithm can determine the pattern of training class and each pixel in full dataset is compared to training patterns and labeled based on similarity.

The methods listed above were used to create classified maps that indicate areas that became flooded, areas that have stayed the same, and areas that involve other types of changes. An error matrix was created for each of the methods in order to obtain values including the total accuracy, producer's accuracy for each land cover class type, and consumer's accuracy for each land cover class type.

Study Area

The study area is an area in Mali not far from the Niger river that includes a river that branches out, and towards the top there is a bigger water body that could be a lake (See Figures 3 and 4). The majority of land are drier areas that could be savannahs, with vegetation scattered across the image, denser along the rivers. The same area in the later image includes many areas that were flooded, some areas that used to be water that became vegetated, and some shrubland areas where vegetation have grown in that used to be dry savannah.

Methods

Density Slicing

Density slicing is done by dividing the range of brightnesses in a single band into intervals, then assigning each interval to a color.
Since the goal is to map flooded areas, prior to density slicing, the images from the two time periods were first converted to NDWI images. The Normalized Difference Water Index (NDWI) uses Near Infrared (NIR) and the green channel and can be used to monitor changes in water content. NDWI was derived for the two images by using Band Algebra in ENVI with the following formula: \( \text{NDWI} = \frac{\text{Green} - \text{NIR}}{\text{Green} + \text{NIR}} \).

Image differencing is a technique that can be used for change detection. The NDWI image from the earlier date was subtracted from the NDWI image from the later date using Band Math in ENVI. Density slicing was then applied to the difference image to create a classified map by using Raster Color Slices in ENVI.

Various colors slices were created based on histogram values. A narrow slice for the values near 0 indicates no change (-0.01 to 0.01), the slice from -0.1 to -0.01 indicates slightly flooded areas, -0.2 to -0.1 indicates flooded areas, -0.4 to -0.2 indicates greatly flooded areas. On the other hand, 0.01 to 0.1 indicates areas that became drier, and 0.1 to 0.2 indicates area that became very dry.

The classes were combined to produce results that make the most sense for the purpose of this project, which is to map flooded areas. The final three classes are: Flooded, no change and drier areas.

**Support Vector Machine**

The second method used for change detection is the SVM supervised classification algorithm.

The SVM algorithm is derived from statistical learning theory and produces an optimal hyperplane by determining the location of decision boundaries that produce the optimal separation between classes. The optimal hyperplane maximizes the distance between itself and the planes representing the two classes. The support vectors are the data points that lie at the edge of each individual class hyperplane and are closest to the optimal hyperplane. SVMs can handle non-linear boundaries between classes by using kernel functions.

The images from the two periods were stacked to produce a stacked image before running the classification algorithm. A class scheme to produce training data was selected: land to water (flooded), vegetation to water (flooded), water to water (no change), land to land (no change), vegetation to vegetation (no change), water to land (drier), water to vegetation (drier) and land to vegetation and training data was created manually by drawing ROI polygons in ENVI.

The SVM classifier was applied to the stacked image with the following options: the kernal type set to radial basis, the gamma value set to the default 0.083, the penalty parameter set to the default 100, the pyramid levels set to the default 0, and the classification probability threshold set to the default 0.
After the SVM classification map was produced, some of the classes were combined in order to highlight the flooded areas, areas that did not change, and areas that changed otherwise. The final four classes are: flooded areas, no change, drier areas, and more vegetated areas.

Three iterations of the SVM algorithm were run, each time a more accurate training dataset was used by creating more refined ROIs for each of the classes to further improve the results until the results were acceptable.

Accuracy Assessment

The first step is to select a sample design for creating testing sites. In general, random points are better than non-random points since it would be governed by chance and each point has an equal likelihood of being represented. It is also better to choose a sample of certain number of points instead of comparing all the points, otherwise it will take too long. The next step is to obtain ground reference data. Finally, the last step is to produce the error matrix.

A random sample design was chosen given that the study area does not have a lot of variability in specific areas. Sixty-four testing pixels were generated randomly from a noise image for the two images to create the testing data. The testing data was then labeled based on comparing the two original images by visualizing both true-color and false-color as well as using Google Earth as a reference. The test data was labeled into four classes: flooded areas, no change, drier areas, and more vegetated areas.

Since it is difficult to identify specific areas that became vegetated in the density slicing method given that the method only detects the absolute change in pixel value between the two images, only the flooded areas, no change, drier areas subset of the test data with 54 training pixels was used to conduct accuracy assessment for the classified image resulting from the density slicing method.. Testing data for all four classes listed above were used for the classified map resulting from the SVM algorithm.

An error matrix was produced for each of the methods by comparing the randomly generated “ground truth” test pixels to the land cover type classes produced in the final images by using the “Confusion Matrix Using Ground Truth ROIs” function in ENVI, including total accuracy, user’s accuracy and producer’s accuracy. Total accuracy is calculated by dividing the number of correctly identified pixels by the total pixels, user’s accuracy for each class is calculated by the diagonal element of that particular class by the total pixels identified as that class, and producer’s accuracy for each class is calculated by the diagonal element of that particular class by the total pixels that should have been identified as that class.

User’s accuracy is associated with the error of commission, and it is when an area is included in a category to which it does not truly belong. Producer’s accuracy associated with the error of omission, and it is when an area is excluded in a category to which it does truly belong. In other words, a user would find that X% of the time an area visited on the ground that the classified map calls class A will actually be class A. Producer can claim that Y% of the time that an area is class A was identified as such.
Results

The result land cover classification of the flooded areas in an area of Mali can be seen in Figure 1. for the density slicing method and Figure 2. for the Support Vector Machine algorithm. At the end there were 3-4 land cover type classes: flooded areas, no change, drier areas, and more vegetated areas (SVM only).

The error matrix for the map resulting from the density slicing method can be seen in Table 1. The overall accuracy for the image is 61.11% (33/54) and the kappa coefficient is 0.4290.

The error matrix for the map resulting from the SVM algorithm can be seen in Table 2. The overall accuracy for the image is 76.67% (46/60) and the kappa coefficient is 0.6729.

The density slicing method has a lower overall accuracy. The flooded areas have low user’s accuracy at 53.4% but high producer’s accuracy at 91.7%, which means that users would find that only about half of the time when they visit the area in this image that what the classified map identifies as flooded areas will actually be flooded areas, and that almost all of the time that an area should be flooded areas is actually classified as flooded areas. The drier areas have both average user’s accuracy at 63.2% and producer’s accuracy at 66.7%, which means that users would find that about slightly more than half of the time when they visit the area in this image that what the classified map identifies as drier areas will actually be drier areas, and about slightly more than half of the time that an area should be drier areas is actually classified as drier areas. The no change areas have higher user’s accuracy at 71.4% but low producer’s accuracy at 41.7%, which means that users would find that a lot of the time when they visit the area in this image that what the classified map identifies as flooded areas will actually be flooded areas, and that almost all of the time that an area should be flooded areas is actually classified as flooded areas.

The SVM method has a higher overall accuracy. The flooded area have lower user’s accuracy at 61.1% but high producer’s accuracy at 91.7%, which means that users would find that slightly over half of the time when they visit the area in this image that what the classified map identifies as flooded areas will actually be flooded areas, and that almost all of the time that an area should be flooded areas is actually classified as flooded areas. The drier areas class is opposite as it has high user’s accuracy at 81.8 % and low producer’s accuracy at 50%, which which means that users would find that most of the time when they visit the area in this image that what the classified map identifies as drier areas will actually be drier areas, and that only half of the time that an area should be drier areas is actually classified as drier areas. The no change areas have high user’s and producer’s accuracies at 87.5%, which means that users would find that almost all of the time when they visit the area in this image that what the classified map identifies as no change areas will actually be no change areas, and that almost all of the time that an area should be no
change areas is actually classified as no change areas. Finally, the more vegetated areas have high user's accuracy at 71.4% and high producer's accuracy at 83.3% and which means that users would find that a lot of the time when they visit the area in this image that what the classified map identifies as more vegetated areas will actually be more vegetated areas, and that almost all of the time that an area should be more vegetated areas is actually classified as more vegetated areas.

Discussion

The result produced by the density slicing method is not very successful because it only detects the differences between the two NDWI images. If an area is still water but perhaps deeper then the index may be big enough to fall into the category of flooded even though it had already been water. It also does not distinguish newly formed vegetation, since the density slicing is based on differences in water content. Since vegetation has a higher water content than savanna the areas that became vegetated were also classified as flooded due to the increase in water content that gives it a more negative value. A lot of what should have been no change areas have been misidentified as drier or flooded areas. Most of the actually flooded areas were correctly identified but many what should have been drier or no change areas were classified as flooded, so there should be less flooded areas than what is shown in the classified image.

The result produced by the SVM algorithm is more successful since finer classes can be created as training data, then the resulting classes can be combined to create the final classes more meaningful for the purpose of this project. Most of the no change areas are correctly identified. Most of the actually flooded areas were correctly identified but many what should have been drier or no change or more vegetated areas were classified as flooded, so there should be less flooded areas than what is shown in the classified image. Most of the classes have improved in both user's and producer's accuracies in the SVM image except for drier areas, which has a lower producer's accuracy. A more accurate training dataset could potentially be used with more refined polygons drawn for each of the classes or use individual pixels to further improve the results.

Conclusions sum up results and findings in a few sentences

Flooded areas were identified in an area in Mali by using change detection methods to identify changed areas between two images from different time periods. The methods used were density slicing and the SVM supervised classification algorithm, and an error matrix was produced for each method to compute the overall accuracy and user's and producer's accuracies for each class. The SVM map produced a better overall result and better user's and producer's accuracies for all classes except for drier areas.
Figure 1. The final classified map that displays flooded areas by using density slicing for mapping change detection in an area in Mali at path 197 and row 50. The acquisition date of the two images are 2013 June 19 and 2013 August 22, the resolution of the image is 30 m, and the image was acquired by Landsat 8 OLI, with bands 1 to 5 plus 7.
Figure 2. The final classified map that displays flooded areas by using a SVM classifier for mapping change detection in an area in Mali at path 197 and row 50. The acquisition date of the two images are 2013 June 19 and 2013 August 22, the resolution of the image is 30 m, and the image was acquired by Landsat 8 OLI, with bands 1 to 5 plus 7.
Figure 3. The true color image (band 4 in the red channel, band 3 in the green channel, band 2 in the blue channel) of an area in Mali at path 197 and row 50. The acquisition date of the image is 2013 June 19, the resolution of the image is 30 m, and the image was acquired by Landsat 8 OLI, with bands 1 to 5 plus 7.
Figure 4. The true color image (band 4 in the red channel, band 3 in the green channel, band 2 in the blue channel) of an area in Mali at path 197 and row 50. The acquisition date of the image is 2013 August 22, the resolution of the image is 30 m, and the image was acquired by Landsat 8 OLI, with bands 1 to 5 plus 7.
### Table 1
Error matrix for the final classified map based on density slicing of the subtraction image of the NDWI image from the earlier date from the NDWI image from the later date, displaying the flooded areas as well as no change areas and drier areas. See figure 1 for associated classified image.

<table>
<thead>
<tr>
<th>Class</th>
<th>Flooded Areas</th>
<th>Drier Areas</th>
<th>No Change</th>
<th>Total</th>
<th>User's Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flooded Areas</td>
<td>11</td>
<td>2</td>
<td>8</td>
<td>21</td>
<td>0.534</td>
</tr>
<tr>
<td>Drier Areas</td>
<td>1</td>
<td>12</td>
<td>6</td>
<td>19</td>
<td>0.632</td>
</tr>
<tr>
<td>No Change</td>
<td>0</td>
<td>4</td>
<td>10</td>
<td>14</td>
<td>0.714</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>18</td>
<td>24</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Producer's Accuracy</td>
<td>0.917</td>
<td>0.667</td>
<td>0.417</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2
Error matrix for the final classified map based on SVM classification of the image from the earlier date and the image from the later date, displaying the flooded areas as well as no change areas, drier areas and more vegetated areas. See figure 2 for associated classified image.

<table>
<thead>
<tr>
<th>Class</th>
<th>Flooded Areas</th>
<th>More Vegetated Areas</th>
<th>Drier Areas</th>
<th>No Change</th>
<th>Total</th>
<th>User's Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flooded Areas</td>
<td>11</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>18</td>
<td>0.611</td>
</tr>
<tr>
<td>More Vegetated Areas</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>7</td>
<td>0.714</td>
</tr>
<tr>
<td>Drier Areas</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td>11</td>
<td>0.818</td>
</tr>
<tr>
<td>No Change</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>21</td>
<td>24</td>
<td>0.875</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>6</td>
<td>18</td>
<td>24</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Producer's Accuracy</td>
<td>0.917</td>
<td>0.833</td>
<td>0.5</td>
<td>0.875</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>