Detecting Surface Oil Using Unsupervised Learning Techniques on MODIS Satellite Data

Joshua Kidd
University of South Florida, jkidd@mail.usf.edu

Follow this and additional works at: http://scholarcommons.usf.edu/etd
Part of the American Studies Commons, and the Computer Sciences Commons

Scholar Commons Citation

This Thesis is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Detecting Surface Oil Using Unsupervised Learning Techniques on MODIS Satellite Data

by

Joshua Kidd

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Computer Science
Department of Computer Science and Engineering
College of Engineering
University of South Florida

Co-Major Professor: Lawrence Hall, Ph.D.
Co-Major Professor: Dmitry Goldgof, Ph.D.
Frank Muller-Karger, Ph.D.

Date of Approval:
March 26, 2012

Keywords: Fuzzy C-Means, Clustering, Texture, Sun Glint, Deepwater Horizon

Copyright © 2012, Joshua Kidd
# Table of Contents

List of Tables ...................................................................................................................................... ii

List of Figures ...................................................................................................................................... iii

Abstract ................................................................................................................................................ iv

Introduction ........................................................................................................................................... 1

Algorithms ............................................................................................................................................ 6
  Fuzzy C-Means Clustering ..................................................................................................................... 6
  Textures .................................................................................................................................................. 8

Data ....................................................................................................................................................... 10

Methodology ......................................................................................................................................... 12
  Generating the Level-2 De-glinted Data ............................................................................................... 12
  Features ................................................................................................................................................. 15

Results .................................................................................................................................................. 17
  Clustering with Water-leaving Radiance Features ............................................................................. 17
  Clustering in Stages Based on Observed Sun Glint .......................................................................... 19
  Clustering with Water-leaving Radiance and Texture Features ....................................................... 19

Discussion ............................................................................................................................................. 30

Conclusions .......................................................................................................................................... 35

References .............................................................................................................................................. 37
LIST OF TABLES

TABLE 1: Summary of Dataset..............................................................................................................10
TABLE 2: Clustering with Water-leaving Radiance Features .........................................................18
TABLE 3: Clustering in Stages Based on Observed Sun Glint......................................................20
TABLE 4: Clustering with Water-leaving Radiance and Texture Features .........................20
LIST OF FIGURES

FIGURE 1: MODIS Terra Satellite Image with NOAA Shape-file Layer Outlining the Surface Oil .................................................................2

FIGURE 2: Graphical Representation of the nLw 645 Product From April 29th ........13

FIGURE 3: Graphical Representation of the De-glinted Lw 645 Product From April 29th .................................................................15

FIGURE 4: Clustering Results for April 29th .................................................................21

FIGURE 5: Clustering Results for May 1st .................................................................22

FIGURE 6: Clustering Results for May 27th .................................................................23

FIGURE 7: Clustering Results for June 9th .................................................................24

FIGURE 8: Clustering Results for June 10th .................................................................25

FIGURE 9: Clustering Results for June 12th .................................................................26

FIGURE 10: Clustering Results for June 18th ...............................................................27

FIGURE 11: Clustering Results for June 25th ...............................................................28

FIGURE 12: The F-Measure For Each Experiment Across All Days .........................29
ABSTRACT

The release of crude oil or other petroleum based products into marine habitats can have a devastating impact on the environment as well as the local economies that rely on these waters for commercial fishing and tourism. The Deepwater Horizon catastrophe that started on April 20\textsuperscript{th} 2010 leaked an estimated 4.4 million barrels of crude oil into the Gulf of Mexico over a 3 month period threatening thousands of species and crippling the gulf coast. The National Oceanic and Atmospheric Administration (NOAA) used several satellite remote sensing technologies to manually track and predict the extent and location of oil on the surface of the gulf waters. This thesis proposes a methodology to automatically identify surface oil using an unsupervised clustering algorithm and compares the discovered regions of oil to the reports generated by NOAA during the incident. The fuzzy c-means clustering algorithm is used to partition the satellite image pixels into groups that represent either oil or not oil. A variety of MODIS data features and image analyzing techniques have been explored to produce the most accurate set of regions.
INTRODUCTION

On April 20th 2010, an explosion on the Deepwater Horizon oil-drilling rig initiated the release of an estimated 4.4 million barrels of crude oil into the Gulf of Mexico over an 86-day period. This catastrophic event quickly surpassed the Exxon Valdez incident to become the largest oil spill in US history with clean up and legal costs expected to exceed 40 billion dollars [1]. A disaster of this magnitude threatens thousands of species, including birds, dolphins, and turtles, and the economies of coastal towns that are populated by commercial fishermen and businesses that specifically cater to gulf coast tourism. In an effort to track the extent and location of this deadly pollutant, the National Oceanic and Atmospheric Administration (NOAA) released reports almost daily that estimated the shape and geographical coordinates of oil floating on the surface of the gulf waters [11]. These reports were manually generated using images captured from several different satellite-mounted remote sensing instruments, including traditional optical sensors that capture data in visible and near-infrared wavelengths, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) [12] and the Medium Resolution Imaging Spectrometer (MERIS) [13], and Synthetic-Aperture Radar (SAR) sensors that capture the reflection of transmitted microwaves, such as the COSMO SkyMed [14], RadarSat I [15], and Envisat ASAR [16]. In addition to providing a mechanism to track the extent of this crisis, the information released each day in these documents, coupled with wind speed and ocean current forecasts, assisted in modeling
projections of where the oil may be going as well as coordinating clean up efforts. The polygonal shapes, referred to in this thesis as shape-files, generated from NOAA’s data would also be used to populate the New York Times web application that tracked the oil spill in real time as the data became available [2] and the Environmental Response Management Application (ERMA) that was designed by NOAA and a group from the University of New Hampshire to assist in environmental emergencies [3]. Figure 1 shows a true color image that was generated using Terra MODIS data from April 29th with a NOAA shape-file layer outlining the surface oil. Many researchers have concluded that there is a definite need for an automated system capable of identifying and tracking oil spills [4, 5].

NASA’s MODIS instrument captures data at 36 different wavelengths, each referred to as a band, ranging from approximately 405 to 14,385 nm and a spatial
resolution of either 250 m, 500 m, or 1 km per side. This remote sensing device has been deployed on both the Terra and Aqua satellites as part of NASA’s Earth Observing System (EOS) initiative. Every day, the Terra satellite crosses over the Gulf of Mexico from north to south at approximately 12 pm EST and the Aqua satellite crosses over the exact same region traveling from south to north at approximately 2 pm EST. The MODIS sensor data captured each day is converted to ocean color products by NASA’s Ocean Biology Processing Group (OBPG) and publicly distributed on line from their site Ocean Color Web. OBPG makes MODIS products publicly available at various processing stages, or levels, such as the Level-1A version of the data containing the raw radiance counts recorded by the sensors, the Level-1B data composed of radiance counts with sensor calibration corrections, and the Level-2 data containing radiances with atmospheric corrections and custom products specific to ocean color applications such as chlorophyll-a [6]. Due to the data’s temporal coverage and accessibility on line, the Terra and Aqua MODIS satellites represent excellent candidates for monitoring surface oil in the Gulf of Mexico.

Many researchers have experimented with using MODIS products to identify oil spills in various bodies of water. Hu et al. [4] demonstrated the use of radiance images generated from the 250 m – 500 m resolution Level-1B bands to discover patches of surface oil in Lake Maracaibo, Venezuela. The researchers were able to intensify the contrast between the surrounding water and contaminated regions by applying atmospheric corrections specific to that body of water. This observation suggests that the atmospherically corrected Level-2 MODIS data should perform better than the Level-1B version of the data. Easson et al. [7] attempted to find a relationship between a variety of
standard MODIS Level-2 ocean products, such as sea surface temperature (SST), chlorophyll concentration, and remote sensing reflectance, and surface oil found by a research vessel in the Gulf of Mexico during the time period from May 5 to May 15, 2010. Unfortunately, the researchers concluded that these standard products, at a resolution of 1 km, were not as effective as other methodologies used for identifying surface oil such as true color images containing sun glint. After reviewing the MODIS data from this time period, cloud coverage significantly reduced the size of this dataset and the presence of sun glint may have caused many of the standard ocean product calculations to be incorrect and saturated. Shi et al. [5] applied the fuzzy c-means (FCM) clustering algorithm to 250 m resolution Level-1B MODIS data with additional texture features in an attempt to identify surface oil in the Bohai sea on April 3, 2005. Although the researchers were able to generate a cluster that seemed to correspond to the oil slick when texture features were included in the dataset, they were unsuccessful in distinguishing the potential oil patches from the coastal shoreline.

In this thesis, I have applied the FCM clustering algorithm to the same MODIS data sets used by NOAA to generate their surface oil tracking reports during the Deepwater Horizon crisis in 2010. The objective is to determine whether the FCM clustering algorithm can accurately predict the presence of surface oil using MODIS Level-2 based features by comparing the clustering results to the shape-files generated by NOAA. The remaining sections are organized as follows. The algorithms section of the paper gives an overview of the fuzzy c-means clustering algorithm and the algorithm required to generate the entropy texture feature. The Data section provides an overview of the MODIS data and the preprocessing required to calculate the Level-2 products. The
methodology section describes the approach to estimate the de-glinted Level-2 water-leaving radiances and the techniques used to select features and mask areas in the image such as land, clouds, and shallow waters. The results section describes the 3 experiments conducted and presents the results as tables and figures. The discussion section walks the reader through each of the days in the dataset and compares the results obtained for each of the experiments. The conclusion section concentrates on methodologies to improve the results and suggests several ideas moving forward.
ALGORITHMS

Fuzzy C-Means Clustering

Constructing a model that can be used to predict the presence of oil typically requires definitive ground truth, or location-tagged data collected on site that can be used to classify pixels of a satellite image found to contain surface oil. In the absence of categorized satellite data to train a model, unsupervised learning techniques can often be used to discover hidden structure in the dataset. Clustering is a very popular unsupervised learning approach that assigns multidimensional data points to exactly one group, or cluster, out of c possible clusters based on some similarity measure. The clustering algorithm used in this thesis, fuzzy c-means (FCM), is a variation of this traditional approach that allows for the assignment of data points to multiple clusters using a membership function to capture how similar a data point is to each cluster [8]. This fuzzy partitioning of pixels provides flexibility for those data points that may actually be a mixture or hybrid of oil and water.

The FCM algorithm represents its clusters using a real numbered \( c \times f \) matrix named \( C \), where \( c \) is the number of clusters to generate and \( f \) is the total number of features in the dataset. Each row in the \( C \) matrix correlates to a cluster’s centroid, which is a point in feature space that represents the center of the cluster. Each data point’s degree of membership within each cluster is tracked in a real numbered \( n \times c \) matrix named \( U \), where \( n \) represents the total number of data points in the dataset and \( c \)
represents the number of clusters to generate. The degree of membership value must be a number between 0 and 1 and all membership values for each data point must sum to 1:

\[
\sum_{i=1}^{c} U[j][i] = 1 \quad \text{for all } j
\]

A greater membership value indicates a higher membership for the data point in a cluster.

The algorithm begins by initializing the cluster centroids to random values within each feature set’s range. Next, the U matrix fields are calculated using the following equation:

\[
U[j][i] = \left( \sum_{k=1}^{c} \left( \frac{d_{ji}}{d_{jk}} \right)^{2/(m-1)} \right)^{-1} \quad \text{for } 1 \leq j \leq n; 1 \leq i \leq c
\]

where \(d_{ji}\) and \(d_{jk}\) represent the distance between the current data point and the specified cluster’s centroid and \(m\) is the weighted exponent. A variety of distance metrics can be used to quantify the difference between a data point and a centroid [17], this implementation uses Euclidean distance. With an initialized membership matrix, \(U\), each cluster center, \(C[i]\), can now be updated using the following equation:

\[
C[i] = \frac{\sum_{j=1}^{n} \sum_{k=1}^{f} (U[j][i])^{m} X[j][k]}{\sum_{j=1}^{n} (U[j][i])^{m}}
\]

where \(X\) represents the set of data points in feature space and \(m\) is the weighted exponent. Notice that the function of the weighted exponent \(m\) is to determine how much influence the membership matrix has on the centroid calculation. The next iteration of the membership matrix is calculated using the new centroid values and compared to the membership matrix of the previous iteration to determine whether the FCM algorithm can halt execution. This routine of calculating cluster centroids and updating the membership
matrix is repeated until the change in membership matrix drops below a configured threshold value known as the stopping criteria:

$$\sum_{j=1}^{n} \sum_{i=1}^{c} (U_t[j][i] - U_{t-1}[j][i])^2 < \varepsilon$$

where \( t \) represents the current algorithm iteration and \( \varepsilon \) represents the configured stopping criteria. To improve the algorithm’s efficiency, this cost function is calculated at the same time as the membership matrix, \( U \).

**Textures**

The output from a variety of image processing techniques can be used as features to identify regions of interest in satellite data. Texture is an image characteristic that can be calculated from a single band of MODIS data and is used to capture the spatial relationship between pixel intensities [9]. A few researchers have reported good results when using entropy textures, derived from gray-tone spatial-dependence matrices, as features to identify surface oil [5, 10].

The first step in calculating a texture from a MODIS band is to generate the collection of gray-tone spatial-dependence (GTSD) matrices. For this exercise, we have quantized the MODIS data so that each intensity value is an integer between 0 and 255. In this implementation, a single GTSD will be a 256 x 256 integer matrix named \( G \) and capture the number of times that each pair of intensities appear in sequence. For example, if the current pixel being processed has a value of 25 and its neighboring pixel has a value of 30 then the GTSD cell \( G[25][30] \) will be incremented by 1. A total of 4 GTSD matrices will be created to handle the 0°, 45°, 90°, and 135° directions,
respectively. The algorithm used to calculate the texture image from these matrices requires them to be symmetrical; therefore the frequency of pixel intensity relationships will be calculated for both neighbors in each direction. For example, the 0° direction GTSD will require that the intensities for pixels $I[x][y]$ and $I[x+1][y]$ be evaluated as well as the intensities for pixels $I[x][y]$ and $I[x-1][y]$. After the frequencies have been captured the matrices are normalized so that the cell values are real numbers between 0 and 1.

Haralick et al. [8] describes 14 different textures that can be derived from the GTSD matrices discussed in the last paragraph. For the experiments in this thesis we will use the following equation to convert each GTSD matrix cell into an entropy value:

$$E[i][j] = -G[i][j] \cdot \log(G[i][j])$$

Similar to the methodology used to create the GTSD matrices, the entropy values of each pixel in the image can be looked up in the appropriate matrix using the intensity of the current pixel and its neighbor. The entropy texture value for each pixel is calculated by sliding a 7 x 7 window over the image and summing the entropy values for all pixels in the neighborhood. Border pixels that do not have a 7 x 7 neighborhood or other pixels that have a neighborhood containing an invalid pixel will be excluded from the texture calculation. Excluded pixels are removed from the dataset before clustering and flagged as invalid in the visual results. Texture images are created for each direction and averaged to establish a final entropy texture image for the specified MODIS band. Texture features will be created for each MODIS product used in these experiments.
DATA

I was able to identify 17 NOAA shape-files that were manually generated using either Terra or Aqua MODIS data [12] captured between April 20th and July 15th, 2010. From this set, over half had to be discarded due to excessive cloud coverage or shapes that were clearly incomplete or invalid. The Level-1A versions of the 8 remaining days were downloaded from the Ocean Color Web’s Level-1 and Level-2 MODIS data browser. Table 1 shows a summary of the dataset.

<table>
<thead>
<tr>
<th>Date</th>
<th>Source</th>
<th>Oil Pixels (Inside Shape)</th>
<th>Non-Oil Pixels (Outside Shape)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr-29</td>
<td>Terra</td>
<td>21,410</td>
<td>48,725</td>
</tr>
<tr>
<td>May-01</td>
<td>Terra</td>
<td>28,142</td>
<td>55,613</td>
</tr>
<tr>
<td>May-27</td>
<td>Aqua</td>
<td>108,961</td>
<td>757,607</td>
</tr>
<tr>
<td>Jun-09</td>
<td>Terra</td>
<td>73,507</td>
<td>168,548</td>
</tr>
<tr>
<td>Jun-10</td>
<td>Aqua</td>
<td>75,988</td>
<td>222,365</td>
</tr>
<tr>
<td>Jun-12</td>
<td>Aqua</td>
<td>88,445</td>
<td>296,422</td>
</tr>
<tr>
<td>Jun-18</td>
<td>Terra</td>
<td>76,878</td>
<td>66,966</td>
</tr>
<tr>
<td>Jun-25</td>
<td>Terra</td>
<td>63,895</td>
<td>293,496</td>
</tr>
</tbody>
</table>

In addition to a browsing tool that provides a mechanism for users to search and download MODIS data by date and area, NASA’s OBPG also provides access to their satellite image analysis software package named SeaDAS [18] that allows researchers to process, display, and analyze data captured from several different satellite sensors including MODIS. SeaDAS scripting functionality was used in this experiment to
convert the raw Level-1A radiance counts of each pixel into specific geo-located, geophysical Level-2 products scaled to the finest resolution available for MODIS data, 250 meters [6]. To standardize each day in the dataset, the resulting Level-2 data files were cropped to create sub-scenes with longitude boundaries of -83.0 to -95.0 degrees and latitude boundaries of 25.0 to 32.0 degrees, transformed from their native geographical projection to an equidistant cylindrical projection, and scaled to a consistent width and height. The boundaries were manually chosen to encompass the entire region affected by the Deepwater Horizon incident.
METHODOLOGY

Generating the Level-2 De-glinted Data

After processing the raw MODIS data for each day in the dataset, the SeaDAS graphical user interface was used to manually explore a variety of Level-2 products. It was discovered that none of the ocean color Level-2 standard products could be used as features in the clustering algorithm because a majority of the pixels were flagged as invalid. To demonstrate the severity of this issue, Figure 2 shows a graphical representation of the $L_w$ 645 product, the water-leaving radiance measured at a wavelength of 645 nm, from April 29th. The pixels in the image that have been set to a color of black indicate saturated water-leaving radiance values. This observation is common for pixels that represent land or clouds, but the large region of black pixels in the center of the image is believed to be caused by the existence of sun glint, the specular reflection of sunlight from the surface of the gulf, interfering with the Level-2 product calculations. The surface oil resulting from the Deepwater Horizon spill is located in this saturated region of pixels.

An ideal feature to capture the presence of oil floating on top of the gulf would be the Level-2 product that represents the radiance counts reflected back to the sensor from the surface of the water, known as the water-leaving radiance or $L_w$. Therefore it is necessary to calculate an estimation of the water-leaving radiance values at each wavelength that is not saturated by sun glint. The total radiance value measured by each
sensor, \( L_{\text{nnn}} \), where the last 3 characters represent the sensor’s wavelength, is composed of radiances reflected from various sources along the path to the surface of the water plus the water-leaving radiance \( L_{w \text{nnn}} \):

\[
L_{t \text{nnn}} = [L_r \text{nnn} + L_a \text{nnn} + L_g \text{nnn} + L_f \text{nnn}] + L_{w \text{nnn}}
\]

The path radiances, grouped by square brackets in the equation, include radiances caused by Rayleigh scattering of molecules in the atmosphere (\( L_r \)), light scattering from aerosols in the atmosphere (\( L_a \)), sun glint (\( L_g \)), and light scattering from foam on the surface of the gulf (\( L_f \)). For the estimation of \( L_w \text{nnn} \), where \( \text{nnn} \) represents all wavelengths in the visible spectrum, we will select a band from the infrared range, in this case 859 nm, and assume that its water-leaving radiance value is zero:

\[
L_{w859} = 0
\]
We will also assume that the path radiances caused by aerosols, sun glint, and foam are equal across all wavelengths. This assumption allows us to formulate the following equation:

\[
[L_a \! nnn + L_g \! nnn + L_f \! nnn] = [L_a \! 859 + L_g \! 859 + L_f \! 859]
\]

Therefore the equation below can be used to estimate the de-glinted water-leaving radiance values for all sensors with wavelengths in the visible spectrum:

\[
L_{w \! nnn} = (L_t \! nnn - L_r \! nnn) - (L_t \! 859 - L_r \! 859)
\]

Figure 3 shows a visualization of the de-glinted L_w 645 product values from April 29th.

Even after the de-glinted Level-2 MODIS products have been calculated, some of the water-leaving radiance values can still be multiple magnitudes larger than the rest of the data. To eliminate these extreme data points, pixels that had a value greater than 900 were flagged as saturated and removed from the dataset before the execution of the clustering algorithm. The remaining estimated water-leaving radiance values are normalized to be a value between 0 and 1 using the following equation:

\[
n_{L_{w \! nnn}[i,j]} = \frac{L_{w \! nnn}[i,j] - \min(L_{w \! nnn})}{\max(L_{w \! nnn}) - \min(L_{w \! nnn})}
\]

where \(L_{w \! nnn}[i, j]\) is the current pixel’s value and \(\max(L_{w \! nnn})\) and \(\min(L_{w \! nnn})\) represent the maximum and minimum water-leaving radiance values for the specified wavelength and day [20]. These normalized de-glinted Level-2 MODIS estimations, \(nL_{w \! nnn}\), will be the basis for the features used in the clustering algorithm.
Features

The de-glinted water-leaving radiance estimations were calculated for the 645 nm band, which has a native resolution of 250 m, the 469 and 555 nm bands, which both have a native resolution of 500 m, and the 412, 438, and 488 nm bands, which all have a native resolution of 1 km. It is important to note that since the data was processed at the finest resolution of 250 m, SeaDAS was required to estimate pixel values for the lower resolution bands, 500 m and 1 km, by using interpolation. After experimenting with various subsets of these bands and reviewing them individually using the SeaDAS GUI, it was discovered that the bands with a native resolution of 1 km often contained strange sensor-related patterns and an excessive number of saturated pixels. Therefore it was concluded that only the 469, 555, and 645 m bands would be used as features in this research. Coincidentally, these bands also make up the red, blue, and green channels of FIGURE 3: Graphical Representation of the De-glinted L\textsubscript{w} 645 Product From April 29th.
the true color images generated by SeaDAS. With the features selected, additional processing was required.

In addition to the standard Level-2 products calculated by SeaDAS, a 32-bit set of processing flags is also generated for each pixel to capture defined conditions such as detecting whether atmospheric correction calculations failed, the presence of extremely high sun glint, or turbid water. These processing flags, referred to as l2_flags, were used in this experiment to exclude pixels from the clustering algorithm if the pixel was over land, was in shallow water, or contained cloud contamination. While the LAND and COASTZ flags did an excellent job of excluding pixels over land or in shallow water along the shore, experimentation with the cloud contamination flag, named CLDICE, indicated that the algorithm produced many false positives. The SeaDAS developer’s forum, accessible through OBPG’s Ocean Color Web, suggested using the processing flags labeled SSTWARN and SSTFAIL to identify pixels containing clouds. Although it appears that not all pixels containing clouds were identified using these l2_flags, the methodology seemed to exclude a majority of pixels containing cloud contamination in the satellite images.
RESULTS

Clustering with Water-leaving Radiance Features

In the first experiment, the de-glinted water-leaving radiances, $L_w 469$, $L_w 555$, and $L_w 645$, were used as features in the FCM clustering algorithm. First the region of interest for each day was identified by executing a function scripted in Python that was capable of finding the rectangular envelope containing the polygonal shape-file with a predefined padding. Clustering was then performed on the resulting region of pixels using 3, 4, and 5 centroids. The clustering results for 4 and 5 centroids are reported in Table 2. The results for 3 centroids were excluded from this thesis because the overall accuracy was very low compared to the results produced with 4 and 5 centroids. For this experiment, it was assumed that 2 clusters represented oil when 4 centroids were used and 3 clusters represented oil when 5 centroids were used. When reviewing the data in SeaDAS, it was observed that the pixels containing oil had a higher water-leaving radiance than the pixels not containing oil. Therefore, oil clusters were automatically chosen by sorting the centroids using the average of their water-leaving radiance feature values and labeling the top 1, 2, or 3 clusters as oil when clustering with 3, 4, or 5 centroids, respectively. The table reports the accuracy of each iteration by specifying the percent of true positives, false positives, and an overall accuracy metric known as the F-measure (FM) [19]:
TABLE 2: Clustering with Water-leaving Radiance Features

<table>
<thead>
<tr>
<th></th>
<th>4 Centroids</th>
<th></th>
<th>5 Centroids</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True Positive</td>
<td>False Positive</td>
<td>F-Measure</td>
<td>True Positive</td>
</tr>
<tr>
<td><strong>Apr-29</strong></td>
<td>53.61%</td>
<td>0.11%</td>
<td>0.697</td>
<td>61.65%</td>
</tr>
<tr>
<td><strong>May-01</strong></td>
<td>50.00%</td>
<td>25.53%</td>
<td>0.499</td>
<td>53.33%</td>
</tr>
<tr>
<td><strong>May-27</strong></td>
<td>27.06%</td>
<td>47.18%</td>
<td>0.119</td>
<td>32.71%</td>
</tr>
<tr>
<td><strong>Jun-09</strong></td>
<td>34.57%</td>
<td>19.81%</td>
<td>0.384</td>
<td>46.95%</td>
</tr>
<tr>
<td><strong>Jun-10</strong></td>
<td>47.91%</td>
<td>21.42%</td>
<td>0.455</td>
<td>62.66%</td>
</tr>
<tr>
<td><strong>Jun-12</strong></td>
<td>30.32%</td>
<td>20.90%</td>
<td>0.303</td>
<td>37.41%</td>
</tr>
<tr>
<td><strong>Jun-18</strong></td>
<td>47.03%</td>
<td>1.91%</td>
<td>0.633</td>
<td>58.55%</td>
</tr>
<tr>
<td><strong>Jun-25</strong></td>
<td>54.19%</td>
<td>3.18%</td>
<td>0.642</td>
<td>65.87%</td>
</tr>
</tbody>
</table>

\[ FM = \frac{2 \times Precision}{(Precision + Recall)} \]

\[ Precision = \frac{TP}{(TP + FP)} \]

\[ Recall = \frac{TP}{(TP + FN)} \]

where the TP variable represents the number of true positives, FP represents the number of false positives, and FN represents the number of false negatives. The images labeled as (A) in Figures 4 through 11 provide a graphical representation of the results for each day when clustering with 5 centroids. The shades of green represent pixels identified as oil, the shades of blue represent pixels identified as not oil, and the black color represents invalid pixels that were excluded from the clustering calculation. The red line outlines the region identified as oil in the NOAA shape-file.
**Clustering in Stages Based on Observed Sun Glint**

Based on the results from the previous experiment and the observation that sun glint in the satellite image creates at least 2 distinct regions, the data was divided along an estimation of this sun glint line and clustering was performed individually on each new set of pixels. The location of the sun glint was estimated manually using true color images generated from the MODIS data. The goal of this experiment is to improve surface oil detection accuracy by clustering these smaller and less diverse regions in 2 separate stages. If this simple technique dramatically improves the overall accuracy of the clustering results, a more formal methodology could then be developed to calculate the position of the sun glint line using the position and angle of the satellite and sun. Clustering these smaller regions with 3 and 4 centroids seemed to perform a little better than with 5 centroids. The results for this experiment have been summarized in Table 3. This technique was not used on the data from June 18th, resulting in a blank row in Table 3, since the true color image did not contain an obvious sun glint line. Similar to the way that results were depicted in the last experiment, the images labeled as (B) in Figures 4 through 9 and Figure 11 represent the clustering results graphically.

**Clustering with Water-leaving Radiance and Texture Features**

The final experiment involved clustering each day in the dataset using the 3 water-leaving radiance features from the first experiment and 3 additional texture features created from the de-glinted radiance bands. The results obtained when clustering with 3 and 4 centroids are shown in Table 4. For this experiment, it was assumed that 1 cluster
TABLE 3: Clustering in Stages Based on Observed Sun Glint

<table>
<thead>
<tr>
<th></th>
<th>3 Centroids</th>
<th></th>
<th>4 Centroids</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True Positive</td>
<td>False Positive</td>
<td>F-Measure</td>
<td>True Positive</td>
</tr>
<tr>
<td>Apr-29</td>
<td>19.29%</td>
<td>0.00%</td>
<td>0.323</td>
<td>55.39%</td>
</tr>
<tr>
<td>May-01</td>
<td>16.74%</td>
<td>0.62%</td>
<td>0.284</td>
<td>42.98%</td>
</tr>
<tr>
<td>May-27</td>
<td>17.50%</td>
<td>9.78%</td>
<td>0.189</td>
<td>30.76%</td>
</tr>
<tr>
<td>Jun-09</td>
<td>25.98%</td>
<td>7.53%</td>
<td>0.363</td>
<td>50.11%</td>
</tr>
<tr>
<td>Jun-10</td>
<td>20.55%</td>
<td>18.01%</td>
<td>0.237</td>
<td>50.19%</td>
</tr>
<tr>
<td>Jun-12</td>
<td>27.37%</td>
<td>8.69%</td>
<td>0.350</td>
<td>41.30%</td>
</tr>
<tr>
<td>Jun-18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun-25</td>
<td>31.97%</td>
<td>9.42%</td>
<td>0.365</td>
<td>51.48%</td>
</tr>
</tbody>
</table>

represented oil when 3 centroids were used and 2 clusters represented oil when 4
centroids were used. The images labeled as (C) in Figures 4 through 11 provide a
graphical representation of the results for each day when clustering with 3 centroids.

Figure 12 contains a chart that summarizes the F-measure values for each experiment
across all days in the data set.

TABLE 4: Clustering with Water-leaving Radiance and Texture Features

<table>
<thead>
<tr>
<th></th>
<th>3 Centroids</th>
<th></th>
<th>4 Centroids</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True Positive</td>
<td>False Positive</td>
<td>F-Measure</td>
<td>True Positive</td>
</tr>
<tr>
<td>Apr-29</td>
<td>70.35%</td>
<td>0.44%</td>
<td>0.821</td>
<td>89.79%</td>
</tr>
<tr>
<td>May-01</td>
<td>47.48%</td>
<td>3.16%</td>
<td>0.618</td>
<td>73.00%</td>
</tr>
<tr>
<td>May-27</td>
<td>67.59%</td>
<td>28.64%</td>
<td>0.373</td>
<td>84.10%</td>
</tr>
<tr>
<td>Jun-09</td>
<td>67.24%</td>
<td>44.27%</td>
<td>0.471</td>
<td>85.40%</td>
</tr>
<tr>
<td>Jun-10</td>
<td>74.16%</td>
<td>28.60%</td>
<td>0.572</td>
<td>95.35%</td>
</tr>
<tr>
<td>Jun-12</td>
<td>61.77%</td>
<td>24.81%</td>
<td>0.481</td>
<td>80.33%</td>
</tr>
<tr>
<td>Jun-18</td>
<td>64.71%</td>
<td>6.06%</td>
<td>0.763</td>
<td>84.82%</td>
</tr>
<tr>
<td>Jun-25</td>
<td>77.18%</td>
<td>22.30%</td>
<td>0.543</td>
<td>84.63%</td>
</tr>
</tbody>
</table>
FIGURE 4: Clustering Results for April 29th. Image (A) shows the results obtained when using the de-glinted water-leaving features and 5 centroids. Image (B) shows the results when data is separated along glint line and clustered separately with 4 centroids. Image (C) shows the results when the de-glinted bands and textures are used as features with 3 centroids.
FIGURE 5: Clustering Results for May 1st. Image (A) shows the results obtained when using the deglinted water-leaving features and 5 centroids. Image (B) shows the results when data is separated along glint line and clustered separately with 4 centroids. Image (C) shows the results when the deglinted bands and textures are used as features with 4 centroids.
FIGURE 6: Clustering Results for May 27th. Image (A) shows the results obtained when using the de-glinted water-leaving features and 5 centroids. Image (B) shows the results when data is separated along glint line and clustered separately with 4 centroids. Image (C) shows the results when the de-glinted bands and textures are used as features with 3 centroids.
FIGURE 7: Clustering Results for June 9th. Image (A) shows the results obtained when using the de-glinted water-leaving features and 5 centroids. Image (B) shows the results when data is separated along glint line and clustered separately with 4 centroids. Image (C) shows the results when the de-glinted bands and textures are used as features with 3 centroids.
FIGURE 8: Clustering Results for June 10th. Image (A) shows the results obtained when using the de-glinted water-leaving features and 5 centroids. Image (B) shows the results when data is separated along glint line and clustered separately with 4 centroids. Image (C) shows the results when the de-glinted bands and textures are used as features with 3 centroids.
FIGURE 9: Clustering Results for June 12th. Image (A) shows the results obtained when using the de-glinted water-leaving features and 5 centroids. Image (B) shows the results when data is separated along glint line and clustered separately with 4 centroids. Image (C) shows the results when the de-glinted bands and textures are used as features with 3 centroids.
FIGURE 10: Clustering Results for June 18th. Image (A) shows the results obtained when using the de-glinted water-leaving features and 5 centroids. Image (B) shows the results when the de-glinted bands and textures are used as features with 3 centroids.
FIGURE 11: Clustering Results for June 25th. Image (A) shows the results obtained when using the de-glinted water-leaving features and 5 centroids. Image (B) shows the results when data is separated along glint line and clustered separately with 4 centroids. Image (C) shows the results when the de-glinted bands and textures are used as features with 3 centroids.
FIGURE 12: The F-Measure For Each Experiment Across All Days.
DISCUSSION

The first day in the dataset, April 29th 2010, is only 9 days after the explosion of the Deepwater Horizon oil-drilling platform and has the smallest region of interest encountered in these experiments. Image (A) in Figure 4 shows the results obtained when clustering with just the 3 water-leaving radiance bands as features, a methodology identified as Experiment 1 in this section, and a centroid count of 5. This graphical result shows that the algorithm does a very good job of capturing the oil pixels in the center of the region outlined by the NOAA shape-file. The corresponding results in Table 2 indicate that clustering with 5 centroids resulted in 61.65% true positives and 0.22% false positives for an overall F-Measure of 0.760. In this case, one of the clusters generated by this routine captures the darker, presumably thicker, region of oil and represents it as dark green in the graphic. The two additional oil regions, represented as lighter shades of green, seem to capture the thinner areas of the surface oil. This will be a common observation with the results generated by the clustering methodology used in the first experiment. In future experiments with this data, instead of hardcoding the number of clusters that represent oil and automatically assigning them labels based on their relative centroid values, it would be interesting to try and establish a multidimensional threshold value and distance that could be used to predict whether each cluster represented oil. When taking a closer look at the graphical representation of the results, you should observe that the dark blue cluster, which represents water, dominates the left side of the
graphic and surrounds the green oil clusters. It may be possible to distinguish these pixels to the left of the glint line from those surrounding the oil by dividing the data along the glint line or identifying additional features to use for clustering.

The results of clustering this first day in the dataset using the methodology outlined in the description of Experiment 2 can be viewed in image (B) of Figure 4. The corresponding accuracy values in Table 3 indicate that there was a slight decrease in performance, relative to the 5 cluster results from the previous experiment, with 55.39% true positives, 0.14 % false positives, and an overall F-Measure of 0.711 when clustering with 4 centroids. In this case, the methodology that separates the pixels into 2 different regions based on a sun glint line observed in a corresponding true color image does not improve the surface oil prediction accuracy of the clustering algorithm and does not warrant the exploration of a more formal technique to calculate the location of the sun glint line. Based on the very apparent change in pixel brightness as you cross over the obvious sun glint line in the true color image, I had expected a substantial increase in accuracy using this methodology.

The results generated while using the techniques outlined in Experiment 3 are shown in image (C) of Figure 4. It should be immediately obvious that the texture features enabled the clustering algorithm to identify a much higher percentage of the pixels contained in the NOAA shape-file, 70.35% true positives, 0.44% false positives, giving a total F-Measure of 0.821, and it only required clustering with 3 centroids and a single cluster to represent the oil pixels. Figure 12 shows a significant increase in accuracy for this day in the dataset when the texture features are included in the clustering activity.
The next day in the dataset, May 1\textsuperscript{st} 2010, had less accurate results when clustering was performed using the water-leaving radiances in the first experiment. Image (A) in Figure 5 shows us that the algorithm did a very good job of identifying true positive pixels in the lower portion and along the entire perimeter of the NOAA shape-file but a significant number of false positives were identified on the right side of the image. This resulted in 53.33\% true positives, 29.48\% false positives, and an F-Measure of 0.504. Dividing the data along the estimated glint line and performing the clustering procedure in 2 stages, the experiment 2 methodology, allowed for a slight decrease in the F-Measure value of 0.499 by decreasing the true positives to 42.98\% and also decreasing the false positives to 14.89\%. In contrast, executing the clustering algorithm using the bands and textures as features resulted in a sharp increase of the F-Measure to 0.748 with 73.00\% true positives and 11.41\% false negatives. In this case, we are looking at the results of clustering with 4 centroids and 2 clusters clearly represent the oil region. Another observation for this particular day is that a false positive cluster of oil in the lower left hand corner of the image was captured in all 3 experiments. The true color image generated by SeaDAS clearly shows a cloud in this area of the image that has not been masked using our l2\_flags. Improving the methodology to identify clouds in the satellite image would improve the accuracy of the clustering algorithm.

Image (A) in Figure 6 shows that the clustering results from the first experiment for May 27\textsuperscript{th} did not capture many of the oil pixels outlined by the NOAA shape-file. Clustering the data divided along the estimated glint line resulted in many fewer false positives and an improvement in overall accuracy, but image (B) in Figure 6 clearly shows that the algorithm still did not accurately capture the NOAA shape-file region.
However, as we can see in image (C) of Figure 6, introducing the texture features dramatically improved the clustering results. The F-Measure jumped from 0.117 to 0.373 with 67.59% true positives and 28.64% false positives. Even though the bottom right and left regions of the image still contain a significant number of false positives, the results still strongly suggest that the texture features significantly enhance the clustering algorithm’s ability to identify surface oil.

The images in Figure 7 clearly indicate that the June 9th data benefited from the methodology in Experiment 2. As the only day in the data set where the best results were obtained using the sun glint line separation technique, the chart in Figure 12 shows that the Experiment 2 methodology resulted in an F-Measure value of 0.591 and the Experiment 3 methodology only produced an F-Measure 0.471. Image (B) of Figure 7 shows that this improvement in accuracy was the result of a sharp decrease in false positive pixels. In contrast, the visual results from June 10th and June 12th summarized in Figures 8 and 9, respectively, show that the clustering results using the Experiment 2 methodology were very poor. In both cases, clusters were created along the glint lines that were not at all representative of the NOAA shape-file regions. A large portion of the June 18th data was masked with clouds and there was no obvious sun glint in the true color image. These conditions resulted in both the 3-feature methodology from Experiment 1 and the 6-feature methodology from Experiment 3 enabling good results on the dataset. The results depicted in Figure 11, show that the data from the final day, June 25th, was best separated into true regions when clustering with the de-glinted water-leaving radiance features and 5 centroids. Image (C) in Figure 11, clearly shows that
including textures as features in the dataset caused an increase in the number of false positives.

In summary, including the texture products as features in the clustering operation improves the overall accuracy of the algorithm. The hypothesis that we could improve the accuracy by dividing the dataset along the glint line and clustering the data in 2 separate stages turned out to only improve the results for 1 day. In all other cases, a significant accuracy improvement did not occur between the first and second methodologies, and the third methodology enabled better results than the second. Given these results it does not make sense to pursue a more accurate technique to locate the actual sun glint line. Finally, the results also suggest that it may be necessary to establish a threshold and distance value to be used in determining whether a cluster represents surface oil. I briefly explored the feasibility of this direction and concluded that it will be necessary to normalize the features collectively across all days instead of as it is now, normalized by day. This will place all of the data on the same scale and hopefully make the discovery of a threshold possible.
CONCLUSIONS

As previously discussed, the results indicate that introducing textures as features improves the overall accuracy of the FCM clustering algorithm over using only deglinted water leaving radiances from the bands. It would be interesting to continue this exploration by evaluating the effect that other textures calculated from the GTSD matrices may have on the accuracy of the clustering algorithm. It may also be rewarding to explore the application of other image processing techniques on the MODIS satellite data.

Although the texture features clearly improve the accuracy of the algorithm, the chart in Figure 12 shows a general decrease in accuracy after the first day in the dataset. This could be due to the fact that the area of the surface oil continued to increase over this time frame, resulting in larger NOAA shape-files and therefore more pixels to cluster. This observation may suggest that as the size of the clustering region increases, a more diverse set of pixels will be encountered resulting in lower oil detection accuracies. Another explanation for this phenomenon suggested in the paper by Innman et al. [7] is that chemical dispersants sprayed by clean-up crews may have changed the properties and appearance of the surface oil over time making consistent regions difficult to locate. This observation also brings into question the accuracy of the shape-files generated by NOAA. Although shape-files were produced almost daily, only those created using the MODIS Aqua or Terra satellites could be used for evaluating the accuracy of the clusters.
because winds and gulf currents altered the shapes of the surface oil hourly. For example, a shape-file generated in the morning using an alternate satellite source would be very different than the shapes observed in a true color image created from MODIS data collected in the afternoon. This issue significantly reduced the size of the dataset making it difficult to gauge the accuracy of the NOAA shape-files. The question of how accurate the methodology used by NOAA to generate their shape-files from the MODIS satellite data still remains. Xiao et al. [10] performed FCM clustering on Level-1B MODIS data captured above the Deepwater Horizon incident on April 25th with entropy textures. Their paper concluded that the methodology used was successful because the total area of the clusters discovered was comparable to the official data, however they did not cite their ground truth data source. Unfortunately, I was not able to compare my clustering results with those published by Xiao et al. because they only worked with a single day that was outside of my dataset and they only reported the estimated area of the clusters identified.

Finally, the drawback with using FCM to identify regions of surface oil is that the researchers are required to specify the number of clusters to be generated and the number of centroids that represent regions of surface oil. In the discussion section, I proposed the establishment of a threshold in feature space and a maximum distance that could be used to identify clusters as regions of oil. Another alternative would be to assume that the NOAA shape-files represent absolute ground truth and train a model using a supervised learning technique such as Decision Trees, Neural Networks, or Support Vector Machines. Ultimately, the goal would be to construct a model that is capable of predicting the presence of oil across days in the dataset.
REFERENCES


[12] NASA MODIS Website, National Aeronautics and Space Administration. (http://modis.gsfc.nasa.gov/)


[18] SeaDAS Home Page, Ocean Biology Processing Group, (http://oceancolor.gsfc.nasa.gov/seadas/)
