A GIS-Centric Approach for Modeling Vessel Management Behavior System Data to Determine Oyster Vessel Behavior on Public Oyster Grounds in Louisiana

David X. Gallegos

University of New Orleans, dgallego@uno.edu

Follow this and additional works at: https://scholarworks.uno.edu/td

Part of the Databases and Information Systems Commons

Recommended Citation

https://scholarworks.uno.edu/td/1918

This Thesis is protected by copyright and/or related rights. It has been brought to you by ScholarWorks@UNO with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Thesis has been accepted for inclusion in University of New Orleans Theses and Dissertations by an authorized administrator of ScholarWorks@UNO. For more information, please contact scholarworks@uno.edu.
A GIS-Centric Approach for Modeling Vessel Management Behavior System Data to Determine Oyster Vessel Behavior on Public Oyster Grounds in Louisiana

A Thesis

Submitted to the Graduate Facility of the
University of New Orleans
In partial fulfillment of the
requirements for the degree of

Master of Science
in
Computer Science

by

David Xavier Gallegos

B.S. University of New Orleans, 2013

December 2014
Acknowledgements

Tremendous thanks to Dr. Thomas Soniat for everything he has contributed to this thesis, it would not have been possible without him. The chalkboard has been permanently marred by our efforts.

Thank you to my advisors Dr. Mahdi Abdelguerfi and Dr. Shengru Tu, whose guidance helped mature and steer this project to completion and was truly invaluable. Thank you to John Finigan for his assistance, technical guidance and support. Thank you to Susan Colley Theodosiou for her advice, feedback, encouragement and biological insight. Many thanks to Nathan Cooper, a great colleague and a sounding board for programming strategy.

Thank you to the Louisiana Department of Wildlife and Fisheries and for funding this endeavor, especially Patrick Banks and Mark Schexnayder who provided pivotal support and collaboration. Thank you to Marc Maniscalco who gave me a starting direction and valuable GIS guidance. Thank you to the many others in the LDWF who assisted along the way.

My gratitude to the University of New Orleans for its excellent foundation of knowledge and guidance. I could not imagine a better fit for my higher education.

Finally, my fondest thanks to my wife Juanita Garces, my mother Karen Gallegos, and my father Gary Gallegos. Thank you to Sam, Morgan, Winston, Laz, Nick and Ursula, and all of those I’m lucky enough to call friends and family. I would not have managed without their love and morale support.
# Table of Contents

List of Figures ...................................................................................................................... iv
List of Tables ........................................................................................................................ v
Abstract .................................................................................................................................. vi
Introduction ........................................................................................................................ 1
Methods ................................................................................................................................ 3
   Study Area .......................................................................................................................... 3
   Vessel Global Positioning Data ....................................................................................... 3
   Speed and Vessel Behavior ............................................................................................. 6
   Spatial Pattern and Vessel Behavior ............................................................................... 8
   Importing Vessel Behavior Data into a GIS Platform .................................................... 11
   Identifying Vessel Behavior ........................................................................................... 11
Model Implementation ........................................................................................................ 13
Results .................................................................................................................................. 16
Discussion ............................................................................................................................ 28
References ............................................................................................................................. 31
Appendix 1: How to Setup the PostgreSQL with PostGIS Environment ....................... 32
Appendix 2: Report SQL code ........................................................................................... 35
Vita ....................................................................................................................................... 37
List of Figures

Figure 1. Louisiana Public Seed Grounds. Sabine Lake (SL), Lake Calcasieu (LC), Vermilion Bay (VB), Atchafalaya Bay (AB), Sister Lake (SL), Terrebonne Bay (TB), Barataria Bay (BB), Lake Borgne (LB) and Breton Sound (BS). ........................................................................................................... 3

Figure 2. A plot of unanalyzed VMS data from Breton Sound, Louisiana. Dots indicate position at time (pings). Light blue areas are reefs as mapped by side scan sonar (SSS). ....................... 5

Figure 3. A plot of unanalyzed VMS data from Bay Long, Louisiana. Dots indicate position at time (pings). Light blue areas are reefs as mapped by side scan sonar (SSS). ................. 6

Figure 4. Linear pattern of vessel movement. ........................................................................ 8

Figure 5. Circular pattern of vessel movement. ....................................................................... 9

Figure 6. Small interval (1-minute) speed (m/s) vs. large interval (5-minute) speed (m/s). ................................................................................................................................. 10

Figure 7. All data points with red overlay for docked, orange overlay for gearing, green overlay for fishing and yellow overlay for travel. ................................................................. 12

Figure 8. Flowchart of how points are allocated into paths. ..................................................... 14

Figure 9. Daily behavior of vessels in Vermilion Bay, Louisiana (10/10/2012). ...................... 17

Figure 10. Single vessel behavior for the 2012-2013 fishing season, Black Bay, Louisiana. ................................................................................................................................. 18

Figure 11. Behavior of all vessels for the 2012-2013 fishing season, Bay Long, Louisiana. ................................................................................................................................. 19

Figure 12. Behavior of all vessels for the 2012-2013 fishing season, Drum Bay, Louisiana. ................................................................................................................................. 21

Figure 13. Behavior of all vessels for 2012-2013 fishing season, Black Bay, Louisiana. ....... 22

Figure 14. Behavior of all vessels for 2012-2013 fishing season, Bay Long, Louisiana ......... 23

Figure 15. Travel markers only for Breton Sound, Louisiana................................................. 24

Figure 16. Fishing markers only for Breton Sound, Louisiana. ................................................ 25
List of Tables

TABLE 1. A sample of unanalyzed VMS data including ASSET_ID(Vessel ID) and SENT_DATE_UTC (Timestamp in Coordinated Universal Time). .................................................... 4
TABLE 2. Initial assignment of vessel behavior as related to speed in knots and meters per second (m/s). ............................................................................................................................ 7
TABLE 3. Fishing, travel, gearing, docked and unknown behavior in hours per day........ 26
TABLE 4. Required software names, versions and purposes........................................... 32
Abstract

The satellite communications system called the Vessel Management System was used to provide geospatial data on oyster fishing over the nearly 1.7 million acres of the public water bottoms in Louisiana. An algorithm to analyze the data was developed in order to model vessel behaviors including docked, gearing, fishing and traveling. Vessel speeds were calculated via the Haversine formula at small and large intervals and compared to derive a measure of linearity. The algorithm was implemented into software using Python and inserted into a PostgreSQL database supporting geospatial information. Queries were developed to obtain reports on vessel activities and daily effort expended per behavior. ArcGIS was used to display and interpret the patterns produced by the vessel activity, yielding information about fishing activity clusters and effort which implied the location and productiveness of oyster reefs.
Introduction

On December 7th, 2006, it became mandatory for all boats fishing federal waters to use the Vessel Management System (VMS) as mandated by the National Oceanic and Atmospheric Administration (NOAA) Fisheries Service Office for Law Enforcement (OLE). The system allows for monitoring of vessel activities to ensure compliance with fishing regulations by recording each vessel’s latitude, longitude and a timestamp at a rate of once per minute over public seed grounds (Gulf Council). This information can be leveraged to glean information about fishing activity and effort.

The company Pole Star Global is contracted for providing and administering the VMS systems (Salvi, personal communication) in Louisiana coastal waters. They use satellite-cellular terminals from Skywave Mobile Communications, a satellite communications company, aboard the fishing vessels (Skywave). IsatData Pro (IDP) 690 and 800 model terminals are installed innocuously to provide communications with Inmarsat 4th generation satellites, which are in equatorial orbit at around 22,000 miles (Inmarsat, 2014).

The use of VMS data to track fishing is not a novel idea. In the United Kingdom, researchers used it to model track fishing vessel behavior for fish, using speed as a single criteria ranging from around 1.5 knots to 4.5 knots, which is intended to envelope the speed spectrum of large vessels while fishing. They compared their speed measurements to logbooks to measure for accuracy (Witt and Godley, 2007). Another pair of researchers, also in the United Kingdom, allowed the frequency distribution of points with relation to speed to influence the fishing range and produced a similar result of 1.5 to 5.5 knots (Gerritsen and
Lordan, 2010). Both papers noted that due to the lack of metadata, such as the size of the vessel or the type of gear used, the ranges had to be broadened to account for equipment variances.

Prior to the usage of VMS, similar research could be conducted by relying on commercial vessel self-reporting. Using data from 1999-2001, researchers were able to track *Illex illecebrosus*, commonly known as the northern shortfin squid, fishery effort in the Atlantic Bight by aggregating the log data. The logs indicated when the trawls were lowered and raised, allowing exact points of fishing to be derived instead of calculated (Powell et al, 2005). This tracking method relies heavily on the error-prone human element, and is more labor-intensive than an automated method.

While speed as a sole indicator may be sufficient to identify fishing behavior of finfish species, oyster reefs in particular are often fished in a circuitous fashion to stay within the limited area of the reef. This thesis develops an algorithm to capture the circuitous nature of typical travel patterns to better distinguish fishing from other behaviors which can occur at similar speeds. The results provide a map of fishing effort on Louisiana public oyster grounds, as well as time expended per day fishing.

Using clusters of fishing activity, it is possible to infer oyster reef locations as well. This proves valuable because a large portion of Louisiana oyster reefs have not been formally mapped at the time of this paper. Concentration of fishing effort on certain regions can also highlight where LDWF may choose to plant hard substrate to maintain oyster reefs and stimulate the settlement of larvae.
Methods

Study Area

This study was conducted on the nearly 1.7 million acres of the public water bottoms ("seed grounds") in the state of Louisiana (Figure 1). The public areas include water bottoms in Sabine Lake, Lake Calcasieu, Vermilion Bay, Atchafalaya Bay, Sister Lake, Terrebonne Bay, Barataria Bay, Breton Sound, Chandeleur Sound, and Lake Borgne.

Figure 1. Louisiana Public Seed Grounds. Sabine Lake (SL), Lake Calcasieu (LC), Vermilion Bay (VB), Atchafalaya Bay (AB), Sister Lake (SI), Terrebonne Bay (TB), Barataria Bay (BB), Lake Borgne (LB) and Breton Sound (BS).

Vessel Global Positioning Data

The data were received in an Excel format which was the result of a database export. It included a unique vessel identifier (ASSET_ID), a timestamp in Coordinated Universal Time (SENT_DATE_UTC), and latitude and longitude (Table 1). Raw VMS data overlaid on a map of
Table 1. A sample of unanalyzed VMS data including ASSET_ID (Vessel ID) and SENT_DATE_UTC (Timestamp in Coordinated Universal Time).

<table>
<thead>
<tr>
<th>ASSET_ID</th>
<th>SENT_DATE_UTC</th>
<th>LATITUDE</th>
<th>LONGITUDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>208</td>
<td>9/21/2012 15:51</td>
<td>29.248150</td>
<td>-91.129750</td>
</tr>
<tr>
<td>219</td>
<td>9/21/2012 15:53</td>
<td>29.254150</td>
<td>-90.795300</td>
</tr>
<tr>
<td>155</td>
<td>9/21/2012 15:56</td>
<td>29.739733</td>
<td>-89.794683</td>
</tr>
<tr>
<td>208</td>
<td>9/21/2012 15:56</td>
<td>29.248067</td>
<td>-91.129167</td>
</tr>
<tr>
<td>213</td>
<td>9/21/2012 15:58</td>
<td>29.253717</td>
<td>-90.794683</td>
</tr>
<tr>
<td>209</td>
<td>9/21/2012 15:59</td>
<td>29.277500</td>
<td>-90.937817</td>
</tr>
<tr>
<td>122</td>
<td>9/21/2012 16:01</td>
<td>29.817083</td>
<td>-89.650300</td>
</tr>
<tr>
<td>212</td>
<td>9/21/2012 16:05</td>
<td>29.265817</td>
<td>-90.940000</td>
</tr>
<tr>
<td>102</td>
<td>9/21/2012 16:06</td>
<td>29.667367</td>
<td>-89.567367</td>
</tr>
<tr>
<td>219</td>
<td>9/21/2012 16:23</td>
<td>29.254217</td>
<td>-90.795500</td>
</tr>
<tr>
<td>155</td>
<td>9/21/2012 16:26</td>
<td>29.739533</td>
<td>-89.794683</td>
</tr>
<tr>
<td>213</td>
<td>9/21/2012 16:28</td>
<td>29.253717</td>
<td>-90.795633</td>
</tr>
<tr>
<td>209</td>
<td>9/21/2012 16:29</td>
<td>29.277500</td>
<td>-90.938033</td>
</tr>
<tr>
<td>208</td>
<td>9/21/2012 16:30</td>
<td>29.248433</td>
<td>-91.127967</td>
</tr>
<tr>
<td>208</td>
<td>9/21/2012 16:33</td>
<td>29.248233</td>
<td>-91.130183</td>
</tr>
<tr>
<td>212</td>
<td>9/21/2012 16:35</td>
<td>29.265917</td>
<td>-90.944817</td>
</tr>
<tr>
<td>102</td>
<td>9/21/2012 16:36</td>
<td>29.667650</td>
<td>-89.567667</td>
</tr>
<tr>
<td>218</td>
<td>9/21/2012 16:37</td>
<td>29.349150</td>
<td>-90.832633</td>
</tr>
</tbody>
</table>
Louisiana produces a plethora of points in which vessel position is given (Figure 2 and Figure 3), but vessel behavior is unresolved. A total of 755,408 points were received spanning from September of 2012 to July of 2013.

Figure 2. A plot of unanalyzed VMS data from Breton Sound, Louisiana. Dots indicate position at time (pings). Light blue areas are reefs as mapped by side scan sonar (SSS).
Figure 3. A plot of unanalyzed VMS data from Bay Long, Louisiana. Dots indicate position at time (pings). Light blue areas are reefs as mapped by side scan sonar (SSS).

**Speed and Vessel Behavior**

Oyster fishermen were interviewed and onboard data were collected to categorize vessel behavior with respect to speed. Most commercial oyster fishing in Louisiana occurs via dredging, which involves operating the vessel slowly; in contrast, travel occurs at relatively high
speeds. Docked behavior, is that in which little or no translocation occurs. Docked is considered as an umbrella term to describe a vessel as tied to a dock, anchored, or otherwise stopped. (It may be tempting to assign a speed of zero to the docked behavior; however, the VMS uses IsatData Pro’s global L-band satellite communications modems, which have a positional error of up to 2.5 meters, which can result in slight spatial deviations over time.) “Gearing” is a single descriptor referring to activities including preparing, deploying and retrieval of the dredge and culling of oysters.

Four categories of vessel behavior were thus initially characterized based on speed alone: docked, fishing, travel and gearing. Other possible behaviors (e.g., exploratory polling of the bottom and excessively slow travel such as in no-wake zones) are unresolved. Fishing speed was initially set at approximately 1.0 to 3.0 knots, or 0.51 m/s to 1.54 m/s. Travel generally occurred at speeds in excess of 7 knots or 3.6 m/s. Docked behavior is distinguished as occurring at speeds under 0.1 knots or 0.05 m/s. Gearing occurs between 0.2 knots and 1 knots, or 0.10 m/s to 0.51 m/s (Table 2).

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Knots</th>
<th>m/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Docked</td>
<td>&lt; 0.1</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>Gearing</td>
<td>0.2 – 1</td>
<td>0.10 – 0.51</td>
</tr>
<tr>
<td>Fishing</td>
<td>1 – 3</td>
<td>0.51 – 1.54</td>
</tr>
<tr>
<td>Travel</td>
<td>&gt; 7</td>
<td>&gt; 3.60</td>
</tr>
</tbody>
</table>

To determine vessel speed, the distance between two geographical points must be calculated. This is done via the Haversine formula for great circles, which simplifies the Earth into a spherical shape with a radius $r$ of 6,371,000 meters. It determines distance $d$ in meters
with inputs of two latitudes, $\Theta_1$ and $\Theta_2$, and two longitudes, $\gamma_1$ and $\gamma_2$ (Equation 1). The timestamps between the two data points (vessel pings) are compared to obtain the duration in seconds, which yields speed in meters per second.

$$d = 2r \cdot \arcsin \left( \sqrt{\sin^2 \left( \frac{\Theta_1 - \Theta_2}{2} \right) + \cos(\Theta_1)\cos(\Theta_2)\sin^2 \left( \frac{\gamma_2 - \gamma_1}{2} \right)} \right)$$ \hspace{0.5cm} (1)

**Spatial Pattern and Vessel Behavior**

Speed as a solitary indicator is insufficient to determine vessel behavior except at high travel speeds. The spatial pattern of pings is an additional source of information for resolving vessel behavior.

Linear travel is demonstrated by a pattern as shown in Figure 4. The line represents a travel path, whereas the five stars represent the VMS ping points occurring at the typical one minute interval. In this case, the speed calculated between one-minute interval points is equal to the speed calculated between five-minute intervals. An example of linear motion is indicated by the point on the diagonal line as seen in Figure 6.

![Figure 4. Linear pattern of vessel movement.](image)

A perfectly circular of motion is shown in Figure 5. In this example, the speed of the vessel over a five minute interval is equal to zero and not equal to the speed calculated over a
one minute interval. This pattern of circular motion is illustrated in Figure 6 as the point on the x-axis.

Figure 5. Circular pattern of vessel movement.
Figure 6. Small interval (1-minute) speed (m/s) vs. large interval (5-minute) speed (m/s).

The diagonal of Figure 6 is useful in behavior analysis. When a point on the graph approached that line, the vessel is moving in a nearly linear pattern. Any points which are above the line indicate acceleration. Points below the line indicate deceleration or circuitous motion.

In practice, vessel speed was calculated as the average of a sliding scale using all five-minute intervals, and herein is termed “net speed”. The calculation of speed is thus weighted to all surrounding points.

While one minute intervals are the norm, the VMS sometimes fails to read pings. Therefore, points are grouped together based on consecutive timestamps requiring that they
occur no more than twenty minutes apart. This grouping is herein termed a “path”, and must include at least five points for the net speed calculation to be made.

**Importing Vessel Behavior Data into a GIS Platform**

The formulated algorithm yields a vessel speed, net speed, vessel behavior and duration of vessel behavior. The data was produced in a spreadsheet format which was imported into ArcMap 10.1. Vessel behavior was the topmost overlay over side-scan sonar reefs, public oyster ground boundaries and a map of Louisiana.

**Identifying Vessel Behavior**

All 337,016 filtered and analyzed data points were graphed with respect to speed and net speed (Figure 7). Based on the point distribution and the previously established boundaries of behaviors, the formula was slightly adjusted. Fishing speeds were broadened to include speeds of up to 1.75 m/s, as a large number of points with low linearity could be observed in that region. A 0.25 unit margin remains between fishing and travel where a discontinuity of data points occurs. Travel retains its original definition of all points with speeds of over 2m/s, but refined with the restriction that the point must be within 2 units of the diagonal to eliminate erroneous and unlikely data. Acceleration points with speeds of 0.5 m/s and above, which are those occurring above the diagonal with high linearity, are included in travel. As it was difficult to distinguish behavior occurring at low speeds with linear travel, a margin of 0.25 units was placed between the fishing and travel regions near the diagonal. Gearing is considered to be anything between 0.1 m/s and 0.5 m/s which lies below a 0.125 unit margin
above the diagonal. Docked retains its definition of points with both speeds and net speeds under 0.1 m/s.

Figure 7. All data points with red overlay for docked, orange overlay for gearing, green overlay for fishing and yellow overlay for travel.
Model Implementation

The software responsible for analyzing the VMS data for vessel behaviors based on the proposed algorithm was written in Python 2.7.5. Once analyzed, it inserts the data into a PostgreSQL database supplemented with PostGIS, which offers geospatial analysis. For information on setting up the database, see Appendix 1. The program then queries for two report types, an activity report and a fishing effort report, which are output as csv formatted files. The program is broken into four major components:

1. VMSC.py: Contains the main entry point of the program, console functionality and composes the queries.
2. DbFunctions.py: A layer on top of the database driver, Psycopg2, which provides query abilities and helper functions to build reports.
3. Path.py: A layer on top of Python’s set which houses trips. Each trip is a closely clustered set of data points.
4. Point.py: A structure for latitude, longitude, timestamp, speed, duration in seconds and net speed attributes of a point record.

The software is a straightforward representation of the algorithm. The logic of allocation of points (pings) into paths is shown in Figure 11.
Figure 8. Flowchart of how points are allocated into paths.

A full activity report which returns each point ordered by vessel identifier and then timestamp is created by a query. It converts the point from its geospatial representation in the database into a string format, per the ST_Y and ST_X PostGIS supported functions. A second query, the fishing effort report, yields a vessel ID and a sum of time per day for each vessel behavior defined (Appendix 2).
For the fishing effort query, the practice of isolating each behavior into its own subquery which is left joined with the rest of the results was necessary for efficiency. A more straightforward approach ran for over 30 minutes before the operation was cancelled, while this query was observed to run in under 15 seconds with a full season’s data.

Examining the first inner query, the main table aliased vpo provided each day and vessel ID and was paired to each behavior via a further subquery, aliased vp1 through vp5. The entire query was wrapped to become a subquery in order to provide the ability to filter by vessel ID and date, which would have otherwise been impractical.
Results

Queries produce spatially referenced visualizations of vessel behaviors. These include visuals of the behaviors of single vessels in a single day, single vessels for an entire fishing season, all vessels for a single day, all vessels for an entire season and single behaviors.

Example results are presented below.
A plot of vessel behavior for a single day (10/10/2012) shows distinct linear patterns of travel and a clustering of fishing, gearing and docked behavior (Figure 9). The behavior of a single vessel (vessel ID 103) is plotted for the entire season in its area of greatest activity (Figure 9).
10). Clusters of fishing activity are centered on a small portion of mapped reefs in the area. No docked (stopped) behavior is evident and gearing clusters with fishing, as expected.

Figure 10. Single vessel behavior for the 2012-2013 fishing season, Black Bay, Louisiana.
Figure 11. Behavior of all vessels for the 2012-2013 fishing season, Bay Long, Louisiana.

The behavior of all vessels for the entire season show a variety of patterns (Figures 11 - 14). Distinctly defined clusters of fishing and gearing activity are identified, much of which occurs near land and off of mapped reef areas (Figure 11). In confined channels and passages vessel behavior shows a less identifiable pattern of activity (Figure 12). Fishing and travel occur
together in channels and open bays. Fishing and docked behavior tends to cluster together, whereas much of the unknown behavior appears linear -- like travel behavior. The plot shown in Figure 13 is of vessel behavior on the most densely trafficked area (Figure 13). A tight clustering of fishing and associated gearing behavior is evident over the oyster reef. Most of the unknown behavior is linear and off of the reef, as is travel behavior. The patterns of behavior plotted in Figure 14 show boats travelling from a channel in the west and dispersing across open water in the east. Fishing is largely restricted to mapped reefs, with notably more gearing behavior on the reef in the southeast as compared to the reef located to its west. (Note that Figure 14 shows vessel behavior for the same region shown in Figure 3 in which positional data are displayed.)
Figure 12. Behavior of all vessels for the 2012-2013 fishing season, Drum Bay, Louisiana.
Figure 13. Behavior of all vessels for 2012-2013 fishing season, Black Bay, Louisiana.
Figure 14. Behavior of all vessels for 2012-2013 fishing season, Bay Long, Louisiana.
Figure 15. Travel markers only for Breton Sound, Louisiana.
Figure 16. Fishing markers only for Breton Sound, Louisiana.

As mentioned above, queries can be used to produce plots of individual behavior. Two examples are presented here -- travel behavior only (Figure 15) and fishing behavior only (Figure 16). Note that the travel tracks show a number of nodes of dispersal/confluence and a few distinct linear pathways (Figure 15). Fishing tends to cluster in the expected manner, largely but not exclusively over mapped reefs. Not all behavior identified as fishing however
shows a clustered pattern, as evidenced by the linear tracks apparent in the northeast sector of Figure 16.

Table 3. Fishing, travel, gearing, docked and unknown behavior in hours per day.

<table>
<thead>
<tr>
<th>Vessel ID</th>
<th>Date</th>
<th>Fishing</th>
<th>Travel</th>
<th>Gearing</th>
<th>Docked</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>103</td>
<td>2/4/2013</td>
<td>7.4</td>
<td>3.0</td>
<td>0.2</td>
<td>0.0</td>
<td>0.7</td>
</tr>
<tr>
<td>114</td>
<td>10/23/2012</td>
<td>4.8</td>
<td>0.9</td>
<td>0.0</td>
<td>0.0</td>
<td>6.6</td>
</tr>
<tr>
<td>126</td>
<td>4/13/2013</td>
<td>8.2</td>
<td>0.6</td>
<td>0.1</td>
<td>0.0</td>
<td>0.4</td>
</tr>
<tr>
<td>135</td>
<td>4/12/2013</td>
<td>0.1</td>
<td>0.0</td>
<td>1.8</td>
<td>1.4</td>
<td>0.2</td>
</tr>
<tr>
<td>152</td>
<td>4/20/2013</td>
<td>3.5</td>
<td>2.1</td>
<td>0.5</td>
<td>1.3</td>
<td>1.0</td>
</tr>
<tr>
<td>158</td>
<td>3/18/2013</td>
<td>0.0</td>
<td>7.6</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>173</td>
<td>4/9/2013</td>
<td>6.6</td>
<td>0.5</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>179</td>
<td>12/19/2012</td>
<td>1.3</td>
<td>3.0</td>
<td>0.0</td>
<td>0.0</td>
<td>3.1</td>
</tr>
<tr>
<td>197</td>
<td>4/9/2013</td>
<td>4.8</td>
<td>2.4</td>
<td>0.4</td>
<td>0.0</td>
<td>0.4</td>
</tr>
<tr>
<td>255</td>
<td>12/18/2012</td>
<td>4.0</td>
<td>2.3</td>
<td>0.7</td>
<td>0.1</td>
<td>0.3</td>
</tr>
</tbody>
</table>

A budget of the time involved in various behaviors was constructed, selected results of which are shown in Table 3. Some vessels had expected distributions of activity (vessels 103, 126, 173), in which a large majority of time was spent fishing with some travel and a small amount of unknown activity. Others (vessels 152, 197, 255) showed more fishing than travel activity, but to a slighter degree than the previous group. The budgets of vessels 114 and 179 showed that more time was unaccounted for than spent in fishing, whereas the time budget for vessel 135 is inexplicable at this time. Vessel 158 apparently simply travelled across the public grounds. From summing up the time allocated for all vessel behaviors for the 2012-2013 fishing
season, it was determined that 12.6% of the time was included in the unknown behavior category. Thus, most vessel behavior was resolved.
Discussion

Compared to the unfiltered analysis of the data, the usage of the algorithm to assign behaviors to each vessel’s points was successful. This can be visually confirmed by examining the maps, where dense clusters of fishing activity compose the filling of granular green polygons. While the existence of the reef can only be inferred from these shapes, it would make little economic sense for the fishermen to frequent an area where no reefs are present.

Common travel routes can be observed on the map, which are consistent in behavior. Vessels tend to travel along the same bands of the Gulf to the reefs, creating thick ropes of activity. Gearing often occurs within reef areas, as would be expected following travel into a reef’s boundaries. Docking occurs rarely over ocean, and is likely an indicator of an anchored vessel or particularly slow gearing.

Some outliers can be found in the maps, such as in Figure 10, which manifests as solitary points or small clusters. This could be a point of refinement on the formula. However, any behavior analysis will have some shortcomings; it would be virtually impossible to detect all fishing activity accurately without real-time management or more sophisticated VMS metadata. There could also be other legitimate explanations for the noise: The fishermen may be testing out new areas for reefs, which will cause small loci of fishing activity. They may be polling the Gulf bottom in a straight line to locate new reefs. Any number of behaviors beyond the four basic determined types that this thesis attempts to distinguish is possible.

The weakest point of this thesis is demonstrated most prominently in Figure 12, where the distinction between behaviors is least clear. It appears as though the proximity of land may
have a negative correlation on a vessel’s fishing speeds, which is not accounted for through the algorithm, and would cause the point to register as gearing. Fishing behavior can still be distinctly observed; however, any attempt to calculate fishing effort for the day would be complicated by the multiple behaviors.

The fishing effort report provides a summary of each vessel’s day, and could be used for interesting statistical analysis. This thesis does not attempt to explore all of the output possibilities, however many interesting data characteristics could be derived such as which vessels were outliers for each behavior, which vessels spent the highest ratio of time fishing and which months had the highest fishing activity.

Confirmation of the results would serve greatly to solidifying the algorithm used in this analysis. Groundtruthing, which would involve an individual recording and comparing pertinent information such as vessel speed, linearity and behavior, would provide invaluable feedback. Polygons could be created from the denser areas of the maps and compared to reef areas by means of additional side scan sonar or direct testing. Those areas with frequent markers of unknown behavior would also be excellent candidates for reef testing.

A density analysis could be produced on the map. With this, fishing efforts could be shown in a gradient fashion that would demonstrate which reefs are getting the most fishing activity. This could be especially useful in determining which regions were subjected to overexploitation.

One of the future directions of this project, and the motivation for the fishing effort report, is to compare the time spent fishing for each vessel to its trip tickets which contain self-
reported catch numbers. If these two pieces of information could be linked together, then a catch per unit effort (CPUE) calculation could be derived. An accurate CPUE could be valuable for limiting a season’s catch in order to aid in the sustainable fishing effort.

This project used a PostgreSQL database with PostGIS extension that allows for geospatial analysis, but did not heavily exploit its features. PostGIS supports linestrings, which allow points to be drawn together as a long line instead of a series of points. This could be used to demonstrate an exact travel path.

Finally, the algorithm could be further refined. Additional fine tuning of the sliding window size for net speed and the duration for which points are considered consecutive for trips could be performed. An adjustment for proximity to land or possibly water depth could prove valuable. A continuity factor could help to prevent consecutive points from frequently crossing between two different behavior definitions. More sophisticated analysis could also shrink or eliminate the unknown behaviors.
References

Article: Integrating vessel monitoring systems (VMS) data with daily catch data from logbooks to explore the spatial distribution of catch and effort at high resolution. Gerritsen, H., and Lordan, C. (2010).


Salvi, Nicholas. Vice President of Pole Star Global North America. "Questions regarding VMS Data." Message to the author. E-mail.

Appendix 1: How to Setup the PostgreSQL with PostGIS Environment

This project relies on numerous software components and packages. Table 6 outlines the product version of various components used in this thesis. Although any version later than those outlined would likely work, the safest reproduction of this thesis’ environment should use the same version numbers.

<table>
<thead>
<tr>
<th>Name</th>
<th>Version</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>PostgreSQL</td>
<td>9.3</td>
<td>Database</td>
</tr>
<tr>
<td>PostGIS</td>
<td>2.1.1</td>
<td>Geospatial analysis</td>
</tr>
<tr>
<td>Python</td>
<td>2.7.5</td>
<td>Interpreter for .py files</td>
</tr>
<tr>
<td>Psycopg2</td>
<td>2.5.2</td>
<td>PostgreSQL driver for Python</td>
</tr>
<tr>
<td>XLRD</td>
<td>0.9.2</td>
<td>Excel package for Python</td>
</tr>
<tr>
<td>ArcMap (optional)</td>
<td>10.1</td>
<td>Importing and viewing report spreadsheet exports</td>
</tr>
</tbody>
</table>

In accordance with Table 4, the following software should be acquired:

- PostgreSQL
- Python
- ArcGIS with ArcMap

The following packages should be installed which work with the above software packages.

- PostGIS
With PostgreSQL installed, create a new database called vms. Although it is possible to hook in the PostGIS functionality through the graphic user interface (GUI), it’s easier to run the following scripts from the command line. Note that these scripts assume a database name of “vms” and a username of “postgres”. If those are not used, change the parameters below accordingly. It also makes assumptions on the installation paths of PostgreSQL and PostGIS, which must be updated if inaccurate.

```
psql -U postgres -d vms -f "C:\Program Files\PostgreSQL\9.3\share\contrib\postgis-2.1\postgis.sql"
psql -U postgres -d vms -f "C:\Program Files\PostgreSQL\9.3\share\contrib\postgis-2.1\spatial_ref_sys.sql"
```

Run the following SQL on the PostgreSQL database to recreate the tables and populate the behaviors. If the script fails on the last line involving AddGeometryColumn, then either PostGIS isn’t installed correctly or the above scripts did not run against the correct database.
CREATE TABLE vessel_behavior (  
vessel_behavior_id SMALLINT PRIMARY KEY,  
vessel_behavior_name TEXT UNIQUE NOT NULL
);
INSERT INTO vessel_behavior (vessel_behavior_id, vessel_behavior_name) VALUES (1, 'Docked');
INSERT INTO vessel_behavior (vessel_behavior_id, vessel_behavior_name) VALUES (2, 'Fishing');
INSERT INTO vessel_behavior (vessel_behavior_id, vessel_behavior_name) VALUES (3, 'Travel');
INSERT INTO vessel_behavior (vessel_behavior_id, vessel_behavior_name) VALUES (4, 'Gearing');
INSERT INTO vessel_behavior (vessel_behavior_id, vessel_behavior_name) VALUES (5, 'Unknown');
CREATE TABLE vessel_path (  
vessel_path_id serial NOT NULL,  
vessel_id integer NOT NULL,  
vessel_path_start TIMESTAMP NOT NULL,  
CONSTRAINT vessel_path_pkey PRIMARY KEY (vessel_path_id),  
CONSTRAINT vessel_path_vessel_id_vessel_path_start_key UNIQUE (vessel_id, vessel_path_start)
);
CREATE TABLE vessel_point (  
vessel_point_id SERIAL,  
vessel_path_id integer REFERENCES vessel_path (vessel_path_id),  
point_timestamp TIMESTAMP WITHOUT TIME ZONE NOT NULL,  
duration_in_seconds smallint NOT NULL,  
speed numeric(8,6) NOT NULL,  
net_speed numeric(8,6) NOT NULL,  
vessel_behavior_id integer REFERENCES vessel_behavior (vessel_behavior_id),  
CONSTRAINT vessel_point_duration_in_seconds_check CHECK (duration_in_seconds BETWEEN 1 AND 1200)
);  
SELECT AddGeometryColumn('vessel_point', 'coords', 4269, 'POINT', 2);

The Python scripts should now be ready to run. If any packages were not installed correctly, the script will fail on the import lines. Also note that the conn_string constant should be updated in accordance with the established PostgreSQL environment.
Appendix 2: Report SQL code

The following queries are used for the full activity and fishing effort reports, respectively.

```sql
SELECT
  vessel_id,
  point_timestamp,
  vessel_behavior_name,
  ST_Y(coords) latitude,
  ST_X(coords) longitude,
  duration_in_seconds,
  speed,
  net_speed
FROM vessel_path
INNER JOIN vessel_point on vessel_path.vessel_path_id = vessel_point.vessel_path_id
INNER JOIN vessel_behavior ON vessel_point.vessel_behavior_id = vessel_behavior.vessel_behavior_id
ORDER BY vessel_id, point_timestamp
```
SELECT * FROM
(
    SELECT
        vpo.vessel_id,
        (vpo.vessel_path_start - (5 * '1 hour'::INTERVAL))::date trip_day,
        COALESCE(sum(path_fishing_time_hrs), 0.0) day_fishing_time_hrs,
        COALESCE(sum(path_travel_time_hrs), 0.0) day_travel_time_hrs,
        COALESCE(sum(path_docked_time_hrs), 0.0) day_docked_time_hrs,
        COALESCE(sum(path_gearing_time_hrs), 0.0) day_gearing_time_hrs,
        COALESCE(sum(path_unknown_time_hrs), 0.0) day_unknown_time_hrs
    FROM vessel_path vpo
    LEFT JOIN (
        SELECT round(SUM(duration_in_seconds) / 3600.0,1)
        path_fishing_time_hrs, vessel_path_id
        FROM vessel_point
        WHERE vessel_behavior_id = 2
        GROUP BY vessel_path_id
    ) vp1 ON vpo.vessel_path_id = vp1.vessel_path_id
    LEFT JOIN (
        SELECT round(SUM(duration_in_seconds) / 3600.0,1)
        path_travel_time_hrs, vessel_path_id
        FROM vessel_point
        WHERE vessel_behavior_id = 3
        GROUP BY vessel_path_id
    ) vp2 ON vpo.vessel_path_id = vp2.vessel_path_id
    LEFT JOIN (
        SELECT round(SUM(duration_in_seconds) / 3600.0,1)
        path_docked_time_hrs, vessel_path_id
        FROM vessel_point
        WHERE vessel_behavior_id = 1
        GROUP BY vessel_path_id
    ) vp3 ON vpo.vessel_path_id = vp3.vessel_path_id
    LEFT JOIN (
        SELECT round(SUM(duration_in_seconds) / 3600.0,1)
        path_gearing_time_hrs, vessel_path_id
        FROM vessel_point
        WHERE vessel_behavior_id = 4
        GROUP BY vessel_path_id
    ) vp4 ON vpo.vessel_path_id = vp4.vessel_path_id
    LEFT JOIN (
        SELECT round(SUM(duration_in_seconds) / 3600.0,1)
        path_unknown_time_hrs, vessel_path_id
        FROM vessel_point
        WHERE vessel_behavior_id = 5
        GROUP BY vessel_path_id
    ) vp5 ON vpo.vessel_path_id = vp5.vessel_path_id
    GROUP BY vessel_id, (vpo.vessel_path_start - (5 * '1 hour'::INTERVAL))::date
    ORDER BY vessel_id, trip_day
) t
Vita

The author was born in Ocala, Florida. He obtained his Bachelor’s degree in Computer Science from the University of New Orleans in 2013. He joined the University of New Orleans graduate program to pursue a Master’s Degree in Computer Science while working as a software engineer.