Identifying Roles of Fishing Ports using Multi-source Data Aggregation

William Hsu[1], Hao-Hsun Wang[2], Min-Ruey You[3], Min-Hsuan Lu[4], Hsin-Yu Tsai[5]
National Taiwan Ocean University
Taiwan
wwyhsu@gmail.com
wangha81@gmail.com
yu830425@gmail.com
a0981801581@gmail.com
cd19941203@yahoo.com.tw

ABSTRACT: *Managing and maintaining fishing ports is crucial for fishing activities. With the lack of budget and labor force, some fishing ports face the brink of being shutdown. A first choice for demolishing would be fishing ports with low utilization, but this does not mean the port is totaly useless. When severe weather condition, i.e., typhoons, is in proximity of the costal line, offshore fishing vessels will return to dock in ports for safety. When major ports are fully docked, some of the minor ports or less active ports will be chosen as the haven. These ports are seldom used and lowly utilized, but should still be maintain for safety issues as they serves as a sanctuary for extreme weather conditions.*

This research uses big data aggregation approach to identify haven ports by processing fishing vessel states using collected voyage data recorder (VDR) records and correlates these information with the typhoon information data source.

Keywords: Data aggregation, Big data, Voyage data recorder, Typhoons, Fishing ports

Received:

1. Introduction

Taiwan is an island country surrounded by water, with offshore and costal fisheries industries producing an estimate of 20 billion TWD per year. Fishing port management is also an issues in fisheries management. How to utilize and maximize the cost and performance of ports and removing those unnecessary ones is a research issue. A solution to alleviate the financial burden is to shutdown and abandon ports with low utilization. However, by using this approach to decide the fate of fishing ports will impose another problem. Fishing ports are not only used for fishing operations but also as sanctuaries for severe weather conditions. Under normal weather and operations, some ports may be nearly empty. When severe weather conditions occur, i.e., typhoons, these ports will serve as a sanctuary for fishing vessels.

There are more than 200 fishing ports in Taiwan, excluding those anchorages used by fishing rafts and sampans. Maintaining each single fishing ports and posting human labors to record the utilization is very costly. To cut down expenses, ports must be closed and human observation is out of consideration. Our approach is to rely on the data available, and aggregate information from them to identify the purpose of each fishing port.

Currently, three systems are used in Taiwan to track vessels. First, the Automatic Identification System (AIS) uses either radar echo or satellites to detect vessels. However, International Convention for the Safety of Life at Sea (SOLAS) only requires vessels of gross tonnage (GT) over 30 or passenger vessels to install AIS. Second, the Vessel Monitoring Systems (VMS) is used to track overseas fishing. This information is realtime and relayed through satellite communication, but offshore and costal fishing vessels do not have them installed. Third, the Voyage Data Recorder (VDR) is used to track costal and offshore fishing vessels. In this research, we will use VDR information to track the position of fishing vessels. The difference between VMS and VDR is that VMS takes samples every 30 minute to 6 hours, and is relayed back to the base station through satellite communication immediately where as VDR takes samples every 3 minute and is downloaded when the fishing vessels returns for fueling. VMS are primarily used for overseas fishing, and the vessels tend to be larger and seldom returns to the original country. VDRs are used for offshore and costal fishing and provides much higher sample resolution, which is a suitable data source for our research. In contrast with overseas fishing operations, costal fishing operations tend to last shortly, usually returning to ports daily or for longer operations, ranging from 3 to 5 days.

Data source of typhoons are freely available from the internet [1], [2]. With this extra information combining our VDR samples, we expect to see the trends of vessels in fishing ports when typhoons are in proximity of Taiwan.

2. Background

Voyage Data Recorder (VDR) has been used for recording and computing vessel operations since 2007 in Taiwan. This device has become an important basis for fishery policy making and research to conduct further studies in understanding fishing activities and marine resources. Similar to the VMS systems, the VDR systems use global positioning satellites (GPS) for trajectory tracking. The latitude, longitude, course bearing, course speed, and time stamp is recorded every 3 minutes. Upon returning to the harbor for refueling, the raw VDR records are uploaded to a central server at the Center of Systems and Naval Mechatronic Engineering (CSNME) for storing. VDR systems was first installed by the government to compute fishing oil stipends and now used for stock assessment and management. Over the years, more than 7000 fishing vessels have installed VDR system and approximately 5000 fishing vessels upload their trajectories to the servers. This vast and big data can be used to create value for further research.

Extensive use of GPS information in fisheries management and stock assessment is widespread. Few examples include, Alvard et al. collected fishing trajectories using small GPS devices to analyzed the foraging behavior of fishing vessels [3]. They used these information to locate Fish Aggregating Devices (FAD) in open seas based on foraging theory [4] by assuming that prey encounters should be spatiotemporally autocorrelated around FADs [5]. Coro et al. used VMS to compute fishing monthly effort [6]. First they classify vessel activities into 3 classes, hauling, fishing, and steaming. Next, they aggregate the fishing hours into 0.5 degrees square and estimate the fishing monthly effort (FME). Hintzen et al. has develop the VMSTools, a package of open-source software capable of parsing erroneous data, linking logbooks, mapping, and analyze fishing impacts [7]. For fishing groups of several vessels with the same exploitation pattern (referred as a "metier"), Russo et al. utilizes VMS data by using artificial neural networks with multilayer perceptrons to provide an assessment of fishing activity [8]. Baker Jr et al. developed a GPS-based binary logit model which can predict fishing effort in a vertical hook-and-line reef fish fishery [9]. They used a minimum of 1 minute data collection interval. They concluded that GPS-based models under perform GPS plus electronic video monitoring (EVM) systems, but is used widespread due to lower costs of collecting data.

3. Method

The objective of this research is to find fishing harbors which serves as a sanctury. These harbors are expected to be lowly utilized during normal status, but for severe weather conditions, they should be docked with many fishing vessels. Two steps are required to reach our objective: collecting and aggregating the data, and comparing plus summarizing the information. We will describe each step in this section.

3.1 Data Acquisition

In order to derive information without posting human sentinels, we can make use of currently available data. The first step is to gather information of typhoons, including its position, speed, radius of maximum wind (RMW), and the radius of outermost closed isobar (ROCI). The data flow of our system is shown in Figure 1. Typhoon information is gathered from both Japan [1] and from the Center of Weather Bureau of Taiwan [2]. These information are parsed and stored into our database. Raw VDR data is periodically uploaded to the Fisheries Agency of Taiwan from the VDR data center at the National Cheng-Kung University. We collect the raw data from the Fisheries Agency and then parse, construct voyages according to the method explained in Hsu et al. [10]. However, due to the vast amount of data to be process, we choose to use MongoDB as our engine, a document store model noSQL database [11]. For our case, we following the definition of a voyage as in [10] being:

• After a fishing vessel enters a port and before the fishing vessel exists a port. (Docking)

• When a fishing vessel leaves a port and then re-enters a port. (Operating)

• When a fishing vessel is in a port, and the next sample shows up in another port. (Possibly the device was carried on land to be installed somewhere else.)

The information required would be the docking status, where fishing vessels is in docking state.

Algorithm 1 Proximity Detection

Input: Set of fishing vessel voyages V, typhoons H, fishing ports P_f, docking threshold d, and parallelizing factor p.
Output: Candidate fishing ports that serves as a sanctuary P_h.

1: for each $v \in V$ using parallelizing factor p do

2: Select v in docking state from database.

3: Trace v and find docking port and docked time.

4: Compute the number of vessels station in each port P_f.

5: end for

6: Compute the monthly average docked vessels $\forall P_f$.

7: for each $h \in H$ using parallelizing factor p do

8: Mark the timeline t where the ROCI of h is in 6nm proximity of Taiwan.

9: If $p \in P_f$ has significantly increase in docked vessels during t, insert p to P_h.

10: end for

11: Output candidate ports P_h.

3.2 Data Aggregation

In the second step, we aggregate the Typhoon information and the voyage information to produce a summary of the utilization of fishing harbors. The number of fishing vessels docked can be computed by analyzing the voyages in docking state. We count the number of fishing vessels stationed at each port at a hourly interval. An assumption made here is that the VDR devices are kept on and never turned off1. If the devices were to be switched off after the captain docks their ship, further analysis will be required to identify the state of the vessel as it would be missing from our records. Algorithm 1 summarized the process that aggregates the data to produce our desired information. Fishing vessels voyages and typhoon trajectories are input into our algorithm. We compute the number of vessels stationed at each port using hourly resolution, and speed up can be obtained by utilizing multi-core parallelism by a factor of p-cores. Shown in Figure 2 is the status of Yan-Pu Dong-Gang fishing port in southern Taiwan from 2007/1/1 to 2015/12/31. The y-axis represents how many fishing vessels is docked according to the collected VDR data.

Next, we mark the period where the typhoon hits the baseline of Taiwan, which in other words, the typhoon is in 6nm proximity of the land. We compute the hourly average docking vessels of each port and compare to the period where the typhoon is in proximity. If this number increases significantly, i.e., 2-3 standard deviations, we can assume that this port is a candidate port which serves as a sanctuary for period of extreme weather conditions. We also created a visualization module to show data for references.

---

[1] Most newer VDR devices installed contains a rechargeable lithim batter that charges when the fishing vessel's engine is powered on.

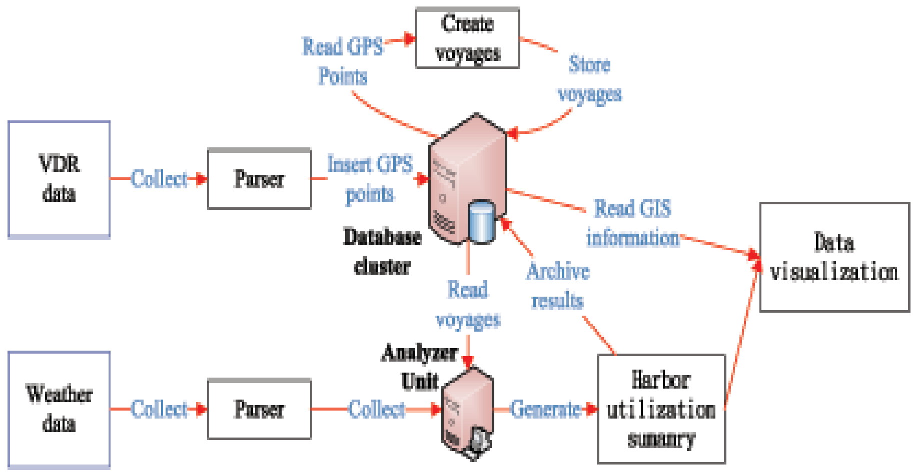

Figure 1. Aggregating multi-source data. This system collects data from weather sites and VDR readers.

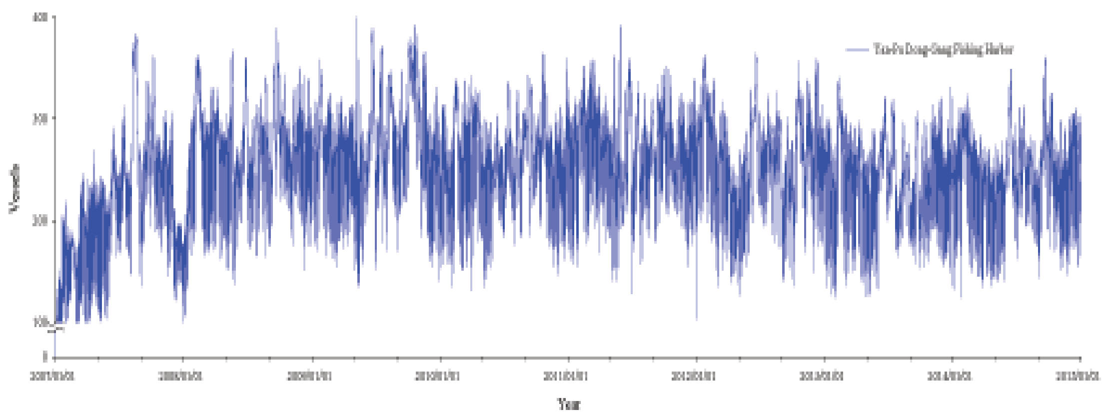

Figure 2. Yan-Pu Dong-Gang fishing port utilization. This plot shows the number of fishing vessels docked in Yan-Pu Dong-Gong port

4. Experiment Result

Our experiments were conducted on our private cloud cluster composing of 6 Intel Xeon E5-2620 v2 2.10GHz, with a total of 36 cores, 320GB of memory and 24TB of distributed cloud storage. VDR data used for the experiments contains more than 52 billion points collected from 2006/01 to 2015/12. Typhoon information is acquired from [1] and [2], with its trajectory, speed, and size (RMW, ROCI) constructed into our database.We have archived 1650 typhoons within the period of 1951 to 2015. The primary operating system for computation is Ubuntu linux version 4.8.2-19. The database software is MongoDB version 3.2 [11]. For the frontend visualization and user interfaces, we use Microsoft Windows Server 2012 R2 along with IIS 8.0 and Cesium 3D globe visualization library version 1.18 [12]. The underlying GIS is obtained from OpenStreetMap [13]. Shown in Figure 3 is Typhoon TEMBIN (Birth: 2012-08-19 06:00:00 UTC, Death: 2012- 08-30 12:00:00 UTC, and lifetime of 270 hours). The inner circle represents the RWM and the outer circle represents the ROCI of TEMBIN. When Typhoon TEMBIN is close to southwestern Taiwan, fishing vessels return to port for safety. When Typhoon TEMBIN leaves, the fishing vessels leave the port for operations (see Figure 4).

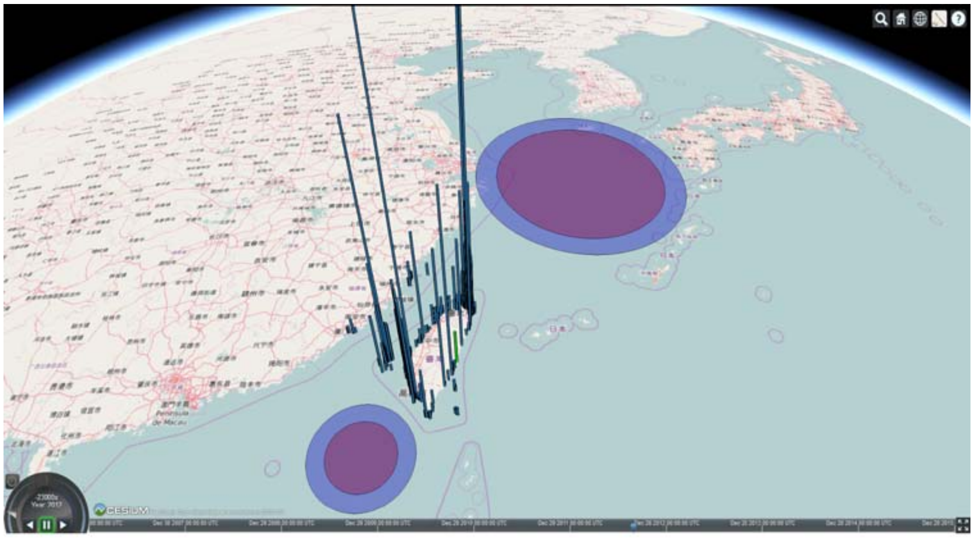

Figure 3. Typhoon TEMBIN (lower left) entering Taiwan. Ports in western Taiwan and southern Taiwan have many vessels docked

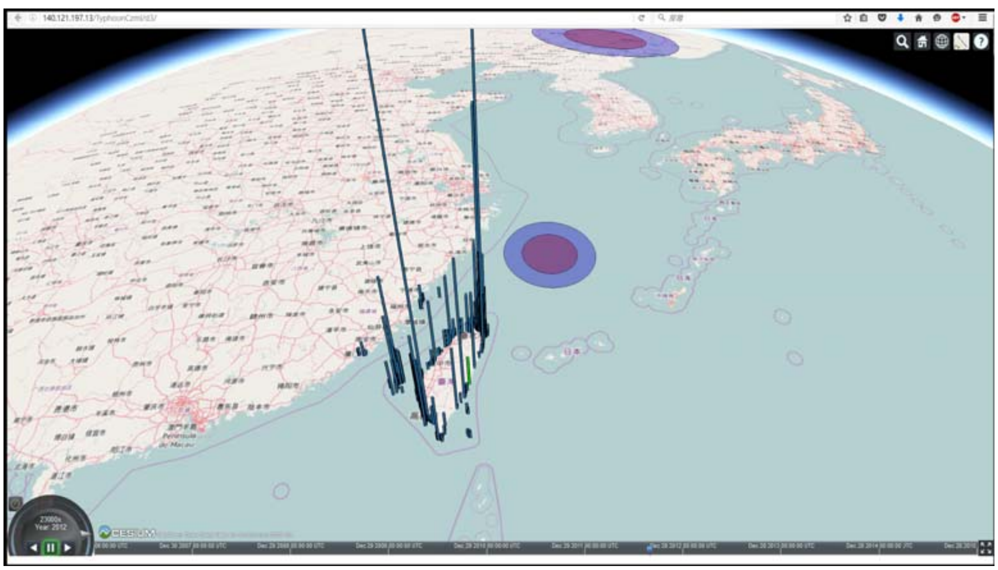

Figure 4. Typhoon TEMBIN leaving Taiwan. Vessels in western Taiwan, especially in Penghu islands, leave for fishing operations

Selected experimental results are shown in Figure 5 and Figure 6, which are records of Chima Fishing port's and Badouzi Fishing port's utilization over the period of 2006/1/1 to 2015/12/31 respectively. We find that Chima Fishing port have docking spikes during typhoon periods, which is marked blue in the figures. The average fishing vessels docked in this port is significantly lower than the maximum during typhoon periods, which indicates the possibility of this port being a sanctuary during extreme weathers. Badouzi Fishing port is known as a recreational and touristic fish market. The average number of vessels docked in this busy port is higher than Chima Fishing port, but our analysis shows that Badouzi has extra space to serve as a sanctuary (see Figure 6. For already busy fishing port, i.e., Yan-Pu Dong-Gang Fishing port in Figure 7, we see some trends that fishing vessels return for safety during typhoons, but the average number of fishing vessels docked does not differ too much in a monthly window.

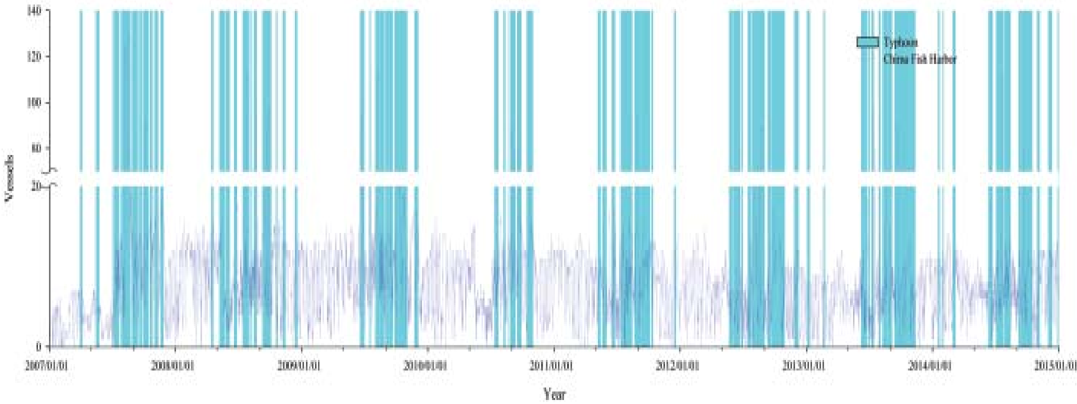

Figure 5. Chima Fishing port utilization during typhoon. This plot shows the number of fishing vessels docked in Chima Fishing port. The duration of a typhoon is marked in light blue

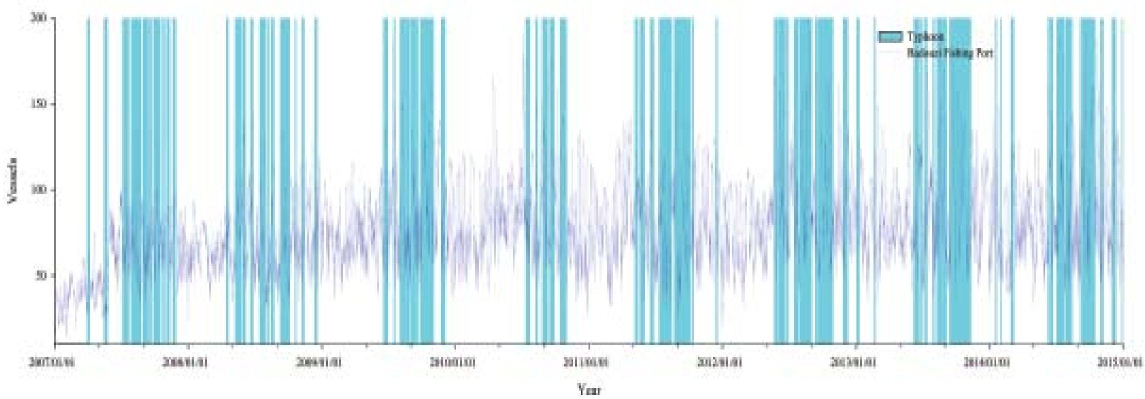

Figure 6. Badouzi Fishing port utilization during typhoon. This plot shows the number of fishing vessels docked in Badouzi Fishing port. The duration of a typhoon is marked in light blue

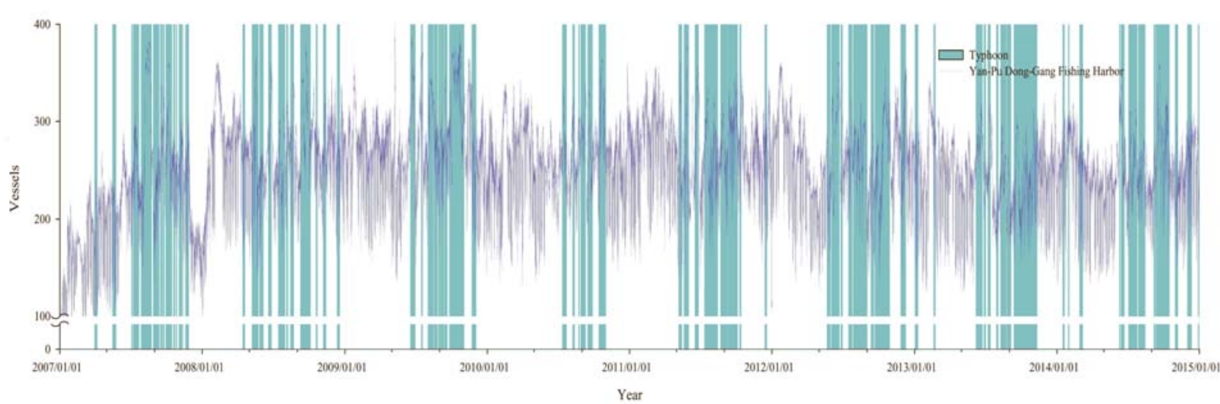

Figure 7. Yan-Pu Dong-Gang Fishing port utilization during typhoon. This plot shows the number of fishing vessels docked in Yan-Pu Dong-Gang Fishing port. The duration of a typhoon is marked in light blue

5. Conclusion

This preliminary work shows the possibility of aggregating data to produce results for future fishing port management. We have constructed a system capable of handling vast amount of VDR GPS point data and analyze this data against a foreign source, i.e., typhoon database. Our system is capable of producing candidate fishing ports that have low utilization in normal days, but serves as a sanctuary during extreme weather conditions. However, much details have not yet been considered into our model. For example, operation time of fishing vessels differ for each season, day-of-time, moon phase, and fishing method, i.e., torchlight fishery, long-line fishing, seine fishery, and pole and lines boote.

We do not know the maximum capacity of each fishing port, so estimates of a full capacity fishing port is not possible at this point. This will be included in our future research direction to increase the accuracy of our model. Moreover, we can also attempt to identify which of the fishing port has transition into a recreational or touristic port. This can help management authorities understand and help promote economic activities among those ports.

References

[1] (2016) Digital typhoon: Typhoon images and information. [Online]. Available: http://agora.ex.nii.ac.jp/digital-typhoon/

[2] (2016) Typhoon database of Taiwan. [Online]. Available: http: //rdc28.cwb.gov.tw/TDB/

[3] Alvard, M., Carlson, D., and McGaffey, E., "Using a partial sum method and GPS tracking data to identify area restricted search by artisanal fishers at moored fish aggregating devices in the commonwealth of Dominica," *PloS one*, vol. 10, no. 2, p. e0115552, 2015.

[4] Stephens D. W.,and Krebs, J. R., *Foraging theory*. Princeton University Press, 1986.

[5] Benhamou, S., "Efficiency of area-concentrated searching behaviour in a continuous patchy environment," *Journal of Theoretical Biology*, vol. 159, no. 1, pp. 67–81, 1992.

[6] Coro, G., Fortunati, L., and Pagano, P., "Deriving fishing monthly effort and caught species from vessel trajectories," in *OCEANS-Bergen, 2013 MTS/IEEE*. IEEE, 2013, pp. 1–5.

[7] Hintzen, N. T., Bastardie, F., Beare, D., Piet, G. J., Ulrich, C., Deporte, N., Egekvist, J., and Degel, H., "VMStools: Open-source software for the processing, analysis and visualisation of fisheries logbook and VMS data," *Fisheries Research*, vol. 115, pp. 31–43, 2012.

[8] Russo, T., Parisi, A., Prorgi, M., Boccoli, F., Cignini, I., Tordoni, M., and Cataudella, S., "When behaviour reveals activity: Assigning fishing effort to m´etiers based on vms data using artificial neural networks," *Fisheries Research*, vol. 111, no. 1, pp. 53–64, 2011.

[9] Baker Jr, M. S., Sciance, M. B., and Halls, J. N., "Potential for a simple gps-based binary logit model to predict fishing effort in a vertical hookand- line reef fish fishery," *Marine and Coastal Fisheries*, vol. 8, no. 1, pp. 118–131, 2016.

[10] Hsu, W. W.-Y., Wu, Y.-W., You, M.-R., Liao, C.-H., Lu, C.-Y., and Wang, H.- H., "Constructing an efficient state space query system for the voyage data recorder," *New Trends on System Science and Engineering: Proceedings of ICSSE 2015*, vol. 276, pp. 294–305W, 2015.

[11] (2016) MongoDB. [Online]. Available: https://www.mongodb.com/

[12] (2016) Cesium: An open-source JavaScript library for world-class 3D globes and maps. [Online]. Available: http://cesiumjs.org/

[13] (2016) OpenStreetMap. [Online]. Available: https://www.openstreetmap. org/