Crowdsourced Maritime Data:

Examining the feasibility of using under keel clearance data from AIS to identify hydrographic survey priorities

by

Christine Schultz

A Thesis Presented to the
FACULTY OF THE USC GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the
Requirements for the Degree
MASTER OF SCIENCE
(GEOGRAPHIC INFORMATION SCIENCE AND TECHNOLOGY)

August 2015

DEDICATION

Many thanks for the support of my colleagues, family, and friends throughout my graduate studies. In particular, thank you to NOAA Corps officers LTJG Anthony Klemm and LT Meghan McGovern for going above and beyond the call of duty to provide NOAA data and technical support at all hours of the day and night.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIS	Automatic Information System
ARGUS	Autonomous Remote Global Underwater Surveillance System
BAG	Bathymetric Attributed Grid
EEZ	Exclusive Economic Zone
GIS	Geographic Information Systems
GOS	Global Observing System
GPS	Global Positioning System
LiDAR	Light Detection and Ranging
MMSI	Maritime Mobile Service Identity
NHSP	National Hydrographic Survey Priorities
NMEA	National Marine Electronics Association
NOAA	National Oceanic and Atmospheric Administration
OCS	Office of Coast Survey
QC	Quality Control
UKC	Under Keel Clearance
UKC%	Under Keel Clearance as a percentage of draft
USCG	United States Coast Guard
VGI	Volunteered Geographic Information
VHF	Very High Frequency
VOS	Voluntary Observing Ship
VTS	Vessel Traffic System

WMO	World Meteorological Organization
Zcor	Tide corrected minimum water depth
Zmin	Minimum water depth

ABSTRACT

NOAA’s Office of Coast Survey annually reviews the NOAA Hydrographic Survey Priorities (NHSP) document to guide the prioritization, planning, and execution of its yearly hydrographic navigational surveys, allocating millions of dollars in assets to help ensure safe navigation in United States navigable waters. As the highest priority navigationally significant areas are completed with modern surveys, NOAA must re-examine how hydrographic surveys are prioritized. One potential source of information that NOAA can employ to analyze areas that might require surveying is ship-generated Automatic Identification System (AIS) data. Ship draft data from AIS can be compared with charted depths to reveal the under keel clearance vessels experience when transiting in and out of ports. The value of under keel clearance compared with a vessel’s draft, combined with the proportion of ships operating at or around under keel safety limits can provide information beyond traditional sources to assess navigational risk. This thesis project assessed the feasibility of using AIS ship draft data to calculate under keel clearance and explore its utility as a factor to determine hydrographic survey priorities. The results proved under keel clearances calculated from AIS vary by port and can be quantitatively used to assign relative risks to ports using draft information. However, the attribute data from AIS must undergo significant quality control measures to remove a large amount of erroneous draft information input by the ships’ crew. Because draft information in AIS messages is a static field, the reported draft carries a great deal of uncertainty; significant negative under keel clearance vessels were calculated during the study. With additional research into the nature of erroneous AIS draft entries and developing detailed, automated quality control measures, AIS data will

have the opportunity to become a variable in a quantitative tool for planning future surveys by NOAA hydrographers.

CHAPTER 1: INTRODUCTION

The United States' ports and waterways are critical to the nation's economy. More than 80 percent of the United States' international trade by volume is conducted by maritime shipping. This trade is responsible for 724 billion dollars of the nation's Gross Domestic Product and it maintains over thirteen million American jobs. Maritime trade in and out of America's approximate 400 commercial ports is made possible by reliable and accurate nautical charts that warn mariners of hazards and help them safely convey their goods (National Oceanic and Atmospheric Administration 2013a). The National Oceanic and Atmospheric Administration (NOAA) is the federal agency, under the Department of Commerce, that is responsible for charting the United States waterways. With the combined NOAA and private contract hydrographic survey assets available today, it will take approximately 300 years to map the entire United States areas of responsibility (NOAA Office of Coast Survey 2012).

NOAA dedicates four large ocean-going vessels, one small research vessel, and six small craft from its research fleet to survey the coast year-round (NOAA Office of Coast Survey 2013b). Hydrographic contracts worth millions of dollars are also awarded annually to contribute to this effort, providing depth data from ships and support in the form of airborne LiDAR (Light Detection and Ranging) shoreline detection, LiDAR depth and obstruction data, and remote tide gauge installations. Even with these dedicated resources, there is still an enormous amount of coastline that must be surveyed and constantly resurveyed to ensure accurate nautical charts. NOAA's Office of Coast Survey must prioritize the areas that need to be surveyed and carefully assign the surveys to the appropriate operational assets (NOAA Office of Coast Survey 2012).

Depending on the size of the survey, its complexity, and the capabilities of the vessels and crew assigned to complete the survey, hydrographic survey projects can take a period of weeks to months. For example, survey number H08878, completed in 1966 in Hampton Roads, took 48 days for the NOAA Ship *Whiting's* hydrographers to complete using single beam sonar (National Oceanic and Atmospheric Administration 1967). In comparison survey H12617 took 20 days for NOAA Ship *Fairweather* to complete using multibeam sonar (National Oceanic and Atmospheric Administration 2014). Although NOAA does not release typical cost estimates for surveys, normal expenses include crew salary (approximately 50 crew members), food for the crew, fuel for the vessels, and other expenses incurred for sailing a ship. Typical survey techniques and operations are discussed further in Chapter 2.

The current annual hydrographic survey prioritization review and methodology is not well documented and uses many qualitative components. The prioritization process takes into account many different criteria, starting with the depth of the water. Areas of the seafloor are deemed 'navigationally significant' based on the depth of the water and the typical characteristics of the seafloor. Waters less than 600 feet deep are deemed significant in some parts of Alaska because of their rocky and unpredictable nature, while the flat seafloor of the Gulf of Mexico is significant for waters less than 120 feet deep. The navigationally significant areas are then assigned a priority level from Critical and Emerging Critical areas, through priority levels 1 through 4 (4 being the lowest survey priority). The age of the last survey, shipping trends, and tonnage, types of cargo, and requests from local mariners and port authorities are among some of the criteria that are analyzed to determine the next year's survey priorities (NOAA Office of Coast Survey 2012).

Professional mariners and surveyors currently acquire hydrographic data and information necessary for the prioritization decision process, but mariners on commercial vessels also acquire useful data during daily operations that can help NOAA assign survey priorities.

Crowdsourcing, the phenomenon that allows users and interested parties to volunteer geographic data to create a collaborative dataset and product, is a data source that saves resources while providing information about a vast area of the ocean (Goetz and Zipf 2013). Large commercial and passenger vessels transmit geographically referenced data about their voyage and their ships' characteristics from their required Automatic Identification System (AIS) instrumentation. These data messages, used mainly for navigational collision avoidance, are also collected by the United States Coast Guard (USCG) for law enforcement and traffic pattern assessment purposes (United States Coast Guard 2014a). If these volunteered data could provide information about vessel traffic patterns, ship characteristics, and risk assumed during transit to help set survey priorities and initiate a survey where ships are vulnerable to new seafloor obstructions, it could provide quantitative information used to improve NOAA's hydrographic survey efficiency and ensure resources are used on the highest priorities.

1.1 Thesis Purpose

This study examined the feasibility of using crowdsourced maritime data as one variable in a new quantitative approach for the NOAA Office of Coast Survey (OCS) to prioritize and initiate surveys in busy ports. Automatic Identification Systems (AIS) required onboard commercial vessels are important tools aboard ships that contribute to collision avoidance at sea. Along with a vessel's name and location, AIS messages also report characteristics about the vessel such as the vessel type, length, draft, and destination. By analyzing the vessel drafts with charted depths in United States ports and harbors, the under keel clearances experienced by ships during transits were calculated.

Under keel clearance is the distance between the bottom of a vessel and the seafloor. This measure, shown in Figure 1, also describes many of the concepts that are used in the methodology and analysis throughout this thesis. Many ports have required minimum under clearances for ships to maintain when transiting into their waters. The percentage of vessels transiting through areas at or near their under keel clearance operational limits can be a valuable variable that can factor into how surveys are prioritized by the OCS each year. This thesis calculated the under keel clearances normalized by vessel draft and conducted a statistical analysis of the normalized under keel clearance values to prove that quantities derived from ship AIS draft data may be useful in future quantitative NOAA OCS survey priority analyses, but not before undergoing extensive quality control processes and further research into the uncertainties inherent to the ship draft inputs.

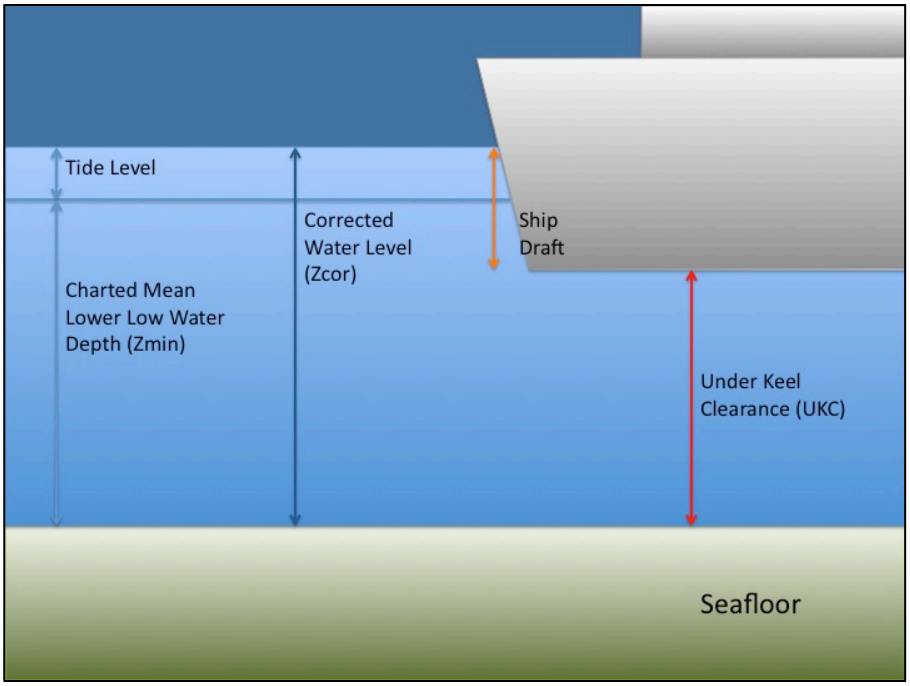

Figure 1 Graphic Representations of Draft, Under Keel Clearance, Charted and Corrected Water Depth, and Tide Level

1.2 Research Objectives and Methods

This research project was a first attempt to understand the nature of AIS draft data and how it can be used to quantify risks for United States ports by calculating and comparing under keel clearance values as a percentage of vessel draft for two ports. This was accomplished by applying quality control measures to vessel track lines derived from AIS data: horizontal position correction (GPS error and removing tracks intersecting the shore), and draft value filtering. The minimum charted depth was calculated and corrected for changing tidal conditions using verified observed tide levels. The corrected minimum depths were used in conjunction with the reported draft from each vessel to calculate the under keel clearance. Under keel clearance values were normalized by the reported draft of each vessel, resulting in the percentage of each ship's draft that was left as under keel clearance for the transit through the two case study areas. These draft percentages were then used to classify vessels into low, medium, and high risk categories. The percentages of vessels within these categories were compared and evaluated for their usability as a variable in a future quantitative survey prioritization model.

1.3 Thesis Organization

This thesis begins by providing background information on hydrographic surveys and charting in the United States, crowdsourcing as a means of data acquisition, and reviewing several crowdsourcing programs and studies that already exist in the maritime community. In the Methods chapter, the case studies, data sets and their sources, quality control, and calculations, are discussed. The Results chapter presents the data errors and calculated under keel clearance results. Finally, the Discussion and Conclusions chapter reviews the results and their implications to setting hydrographic survey priories, makes recommendations for using AIS draft data, and suggests future studies and research.

CHAPTER 2: BACKGROUND AND LITERATURE REVIEW

The nation's ability to conduct international trade hinges on the commercial shipping community. Despite the majority of international trade being conducted by maritime shipping, the industry remains surprisingly out of the spotlight (NOAA Office of Coast Survey 2012). This anonymity of the maritime shipping industry is the result of reliable and accurate nautical charts; these charts warn mariners of hazards and allow them to safely convey their goods though America's commercial ports.

2.1 Hydrography

Since Thomas Jefferson commissioned the first survey of the coast in 1807 to stimulate commerce in his newly formed country, the United States government has been responsible for a nautical charting program that now maintains over 1,000 traditional paper and electronic navigational charts. The extents of the navigational charts are shown in Figure 2. NOAA is responsible for charting the waters within the United States' Exclusive Economic Zone (EEZ). The EEZ extends from the shoreline out to 200 nautical miles offshore. In total, over 3.4 million square miles of seafloor fall into this area of responsibility (NOAA Office of Coast Survey 2012).

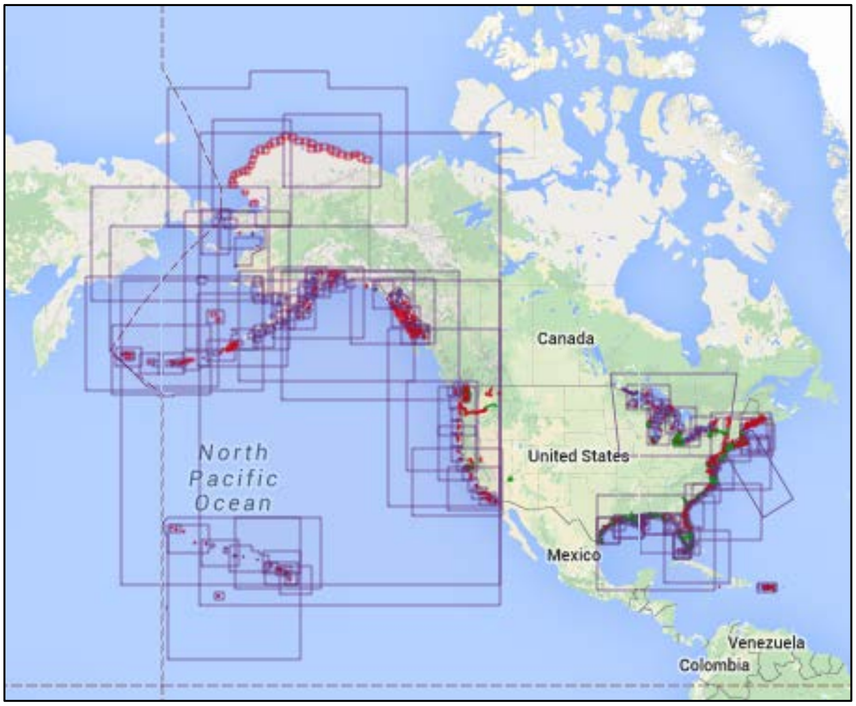

Figure 2 Extents of navigational charts maintained by NOAA

(Source: Office of Coast Survey 2015)

2.1.1 Hydrographic Survey Priorities

NOAA’s surveys are planned years in advance of data acquisition in order for survey platform logistics to be arranged, background data assembled, and reconnaissance to be performed. Before survey planning can begin, the survey areas must be prioritized (NOAA Office of Coast Survey 2012). Although most of the coastline has already been charted, much of the older survey data are considered to be inadequate; older hydrographic survey techniques such as lead line surveys or single beam sonar surveys inherently have data gaps that might not capture all hazards. Lead line surveys involved dipping a 10-pound weight into the water and measuring the amount of rope deployed when the weight reached the bottom. Similarly, single beam sonar does not provide full bottom coverage like the new multibeam sonar surveys (NOAA

Office of Coast Survey 2014). Also, navigation and location technology have advanced in recent decades, leading to increased positional accuracy not achieved before Global Positioning Systems (GPS) was available. Survey prioritization is completed in areas that are deemed "navigationally significant", or areas in the EEZ with depths less than the following criteria:

- 120 ft depth: Atlantic and Pacific coasts, East Gulf of Mexico, North Slope Alaska, Caribbean, Virgin Islands, Puerto Rico
- 300 ft depth: West Gulf of Mexico, West Alaska
- 600 ft depth: Pacific Islands, Alaska (except for West Alaska)

The navigationally significant areas are then prioritized based on:

- Shipping trends and tonnage
- Age of the last survey
- Technology used in the previous survey
- Under keel clearance needed for vessels
- Potential for previously unknown hazards
- Requests from the local community and government agencies

After compiling data for these requirements, the coastline is categorized into Critical and Emerging Critical Areas, and areas of Priority 1 through Priority 4. Critical and Emerging Critical are the highest priority, and Priority 4 is the lowest. The Critical and Emerging Critical areas are where high commercial traffic, hazardous cargo, and minimal under keel clearance (the bottom of the vessel is dangerously close to the seafloor) conditions exist. These areas contain the greatest risk factors for maritime incidents. As of 2012, approximately 28,000 square nautical miles were designated Critical areas. About 40% of the Critical and Emerging Critical Areas are located in Alaskan waters (NOAA Office of Coast Survey 2012).

2.1.2 Survey Data Acquisition Resources and Methods

NOAA's survey fleet is comprised of four hydrographic research ships, one small research survey vessel, and six small navigation response teams. Four of the vessels are large commissioned vessels that sail with fifteen to fifty crewmembers, complete twenty-four hour operations, and sail eight months out of the year. Two of these vessels are shown below in Figure 3 (NOAA Office of Coast Survey 2013b). The small research and navigation response vessels have limited range and can only complete daytime operations, but they can be quickly deployed to areas of need in emergency situations (National Oceanic and Atmospheric Administration 2013b).

Figure 3 NOAA Ship *Rainier* and NOAA Ship *Fairweather*
(Source: NOAA Office of Coast Survey 2013b)

Contemporary surveys are completed with three main types of sonar systems: single beam sonar, mulitbeam sonar, and side scan sonar. Single beam sonar creates a single, narrow beam of sound energy that detects the range from the sonar to the sea floor. This method of survey leaves gaps in the sea floor coverage, shown in Figure 4 below. Single beam sonar is used in areas close to shore where the risk of hitting a rock is greatest. Single beam sonar systems are relatively inexpensive in comparison with multibeam and sidescan sonar, so they are often deployed in areas where equipment could be damaged (NOAA Office of Coast Survey 2014).

Multibeam sonar can cover a wider swath of the sea floor than single beam sonar because multiple beams of sound energy are projected at once in a swath pattern underneath the survey vessel. Instead of receiving one sounding per sonar ‘ping’, some sonar can receive nearly 100 data points fanned out across the bottom, covering a much larger area of the sea floor in comparison with single beam techniques (NOAA Office of Coast Survey 2014).

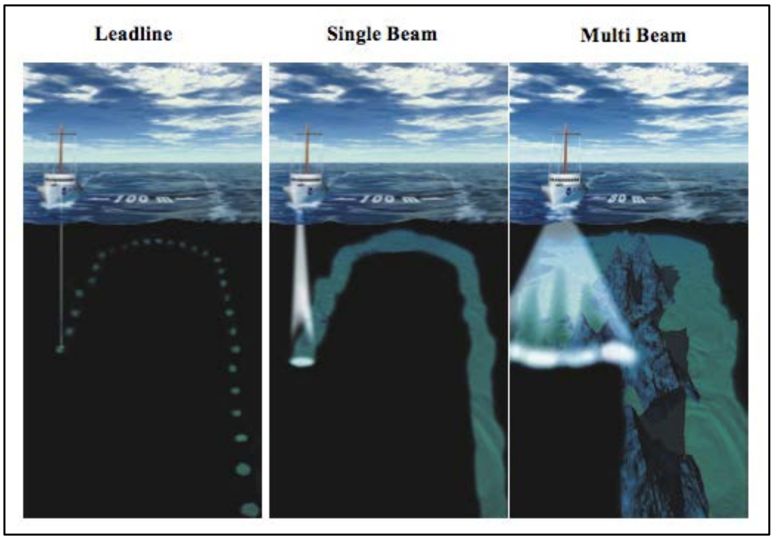

Figure 4 Sea Floor Coverage with Different Sonar Systems
(Source: NOAA Office of Coast Survey 2014)

Instead of using range detection from sonar pulses to determine water depths, side scan sonar uses sound to detect objects and obstructions on the sea floor. The sonar is usually towed behind a vessel instead of being attached to its hull. The beams of energy are directed to the side of the sonar in order to hit the sea floor and any objects at a steep angle. The returned energy is interpreted as an image. Shadows can be seen in the imagery, indicating the sonar has detected an object protruding above the sea floor. This principle is demonstrated in Figure 5. Because the sonar is towed and the sonar beams are not angled straight down, depth information cannot be derived from a side scan sonar’s data (NOAA Office of Coast Survey 2014).

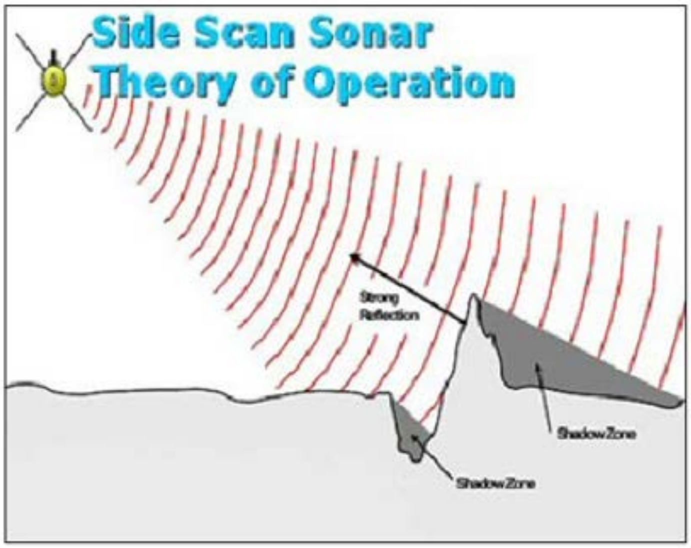

Figure 5 Side Scan Sonar Theory of Operations
(Source: NOAA Office of Coast Survey 2014)

The project instructions that accompany each NOAA survey assignment specify how and where each survey method will be conducted in the survey area. Often multibeam sonar and side scan sonar survey techniques are used together. In areas where the sea floor is known to be flat with little variation, such as the Gulf of Mexico and many areas of the East Coast, full multibeam sonar coverage is not required. By combining multibeam and side scan techniques, ships acquire a swath of depth data immediately below the vessel and can detect objects and obstructions to the side of the vessel. The sea floor is usually painted by 200% side scan coverage in these cases; multibeam seafloor coverage is limited to the track lines of the vessel. Survey areas where the seafloor is known to vary greatly in depth and have many obstructions is usually covered by 100% multibeam sonar soundings. Both depth and obstructions are detected with this method, but survey lines must be run closer together to cover the complete sea floor, making these surveys slower than side scan surveys (NOAA Office of Coast Survey, 2014).

2.2 Automatic Identification System

AIS is a system that is required for vessels 300 gross tons and larger that travel internationally, 500 gross tons and larger that travel domestically, and all commercial passenger vessels (International Maritime Organization, 2014). The system was developed in the 1990s as a secondary navigation and collision avoidance tool. AIS messages are broadcast over two VHF channels, reporting their position, ship characteristics, and voyage characteristics. The AIS is connected to the ship's GPS, so the position is very accurate. Details that are manually set by the ship bridge crews, including the vessel's draft, length, unique Maritime Mobile Service Identity number (MMSI), vessel type, and other parameters, are reported about the vessel. Its speed over ground, course over ground, destination, and other parameters are reported about the vessel's voyage. The shipboard AIS also receives messages from other vessels within VHF radio range and displays their characteristics to the bridge crew (Schwehr and McGillivery 2007).

The Nationwide Automatic Identification System (NAIS) is composed of over 200 land-based VHF receiver sites distributed throughout the United States. This system was designed to record AIS messages from United States ports and waterways and is used by the United States Coast Guard (USCG) and other government bodies. These AIS messages are collected mainly for search and rescue, emergency response, and maritime security, but the USCG and other agencies make certain datasets available to the public (United States Coast Guard 2014b).

2.3 Crowdsourced Data

Crowdsourcing, the growing trend that allows users and interested parties to volunteer data to create collaborative datasets and products, is becoming a popular way of acquiring geographic data and performing geographic analyses (Goetz and Zipf 2013). Crowdsourcing is a bottom-up approach to acquiring data on a subject instead of a top-down acquisition scheme (Aitamurto et al. 2011). Professionals and members of the general public create and contribute data based on their own measurements and experiences. OpenStreetMap is one example of a well-known geospatial crowdsourcing project. Informed users are able to add details to online maps, filling in data gaps that may exist in the information already available (Goodchild 2007). By using data provided by the general public, a vast network of observers is essentially created. Goodchild equated volunteers of geographic information to human sensors in his 2007 article in *GeoJournal*. He noted that networks of human sensors have a much greater potential to cover larger areas and capture new experiences than sensor networks constructed and deployed for specific purposes.

Quality control of crowdsourced data is an evolving subject that can be viewed in several different ways. Elwood, Goodchild, and Sui (2013) explain that there are three main methodologies to approaching crowdsourced data quality: consensus, moderated, and geographic. The consensus approach verifies crowdsourced data by involving as many reviewers and contributors as possible. The more data points there are, the more likely there will be a consensus around the correct answer. The moderated approach relies on trusted sources to review and verify data. The geographic approach assesses quality based on the spatial relationships of whatever topic is being studied. This framework is used to assess quality of AIS derived draft and under keel clearance data in this study.

In a way, crowdsourcing is already incorporated into NOAA’s nautical charting process, as the United States Coast Guard and the maritime community can report hazards and dangers to NOAA charting offices received from local sources (NOAA Office of Coast Survey 2013c).

There are other uses of crowdsourced data that are being explored by the hydrographic community. In a July 2013 report prepared by the Committee of Experts on Global Geospatial Information Management of the International Hydrographic Organization (IHO), the worldwide authority and governing body for hydrography, bathymetric crowdsourcing was specifically addressed: "Crowd-sourced bathymetry and satellite derived bathymetry cannot replace systematic, fully regulated hydrographic surveys, but these methods can provide rapid improvements to existing charts and help identify and prioritize those areas that require more comprehensive surveys. For many areas of the world, such techniques may be the only way to obtain at least a first coverage of indicative hydrographic information" (Ward and Bessero 2013, 8). The report points out that crowdsourced data from ships transiting near the Antarctic Peninsula have provided useful information where there previously was none available.

2.3.1 Crowdsourcing Technology

Crowdsourcing could not have become a popular and effective method of data acquisition without many modern innovative technologies. Goodchild presents five technological advances imperative to the successful development and implementation of geographic crowdsourcing: Web 2.0 advances allowing internet users to contribute to websites and databases, georeferencing and Global Positioning System (GPS), geotags, improved and diversified communication methods, and enhanced computer and mapping graphics (Goodchild 2007). GPS and improved communication are arguably the most important of the technology enhancements. The availability of precise positioning in common hand-held devices and the means to send and submit these positions and auxiliary data truly have transformed most citizens into potential human sensors.

2.3.2 AIS as a Crowdsourcing Tool

The carriage requirements for AIS and the trained users operating AIS make the system a nearly ideal tool for acquiring crowdsourced data in United States waters. Due to the tonnage and passenger requirements dictated by the International Maritime Organization, thousands of vessels carry AIS and their reports are automatically collected by the NAIS as they enter United States harbors (United States Coast Guard 2014b). This requires zero action by the ship's crew once pertinent information is entered into the system, and no reason for data to not be sent during regular operations and transits.

Ship bridge crews are also highly trained on their navigational and emergency communications electronics. The typical bridge watch stander is well versed in their equipment and likely to ensure the data they are reporting is accurate; many of the fields in AIS are input by hand and mistakes can be made when entering information. In a study completed by Harati-Mokhtari et al., errors in AIS messages were researched. Errors in ship length and beam were found to be the greatest, with nearly 47% of vessels reporting the incorrect lengths. Six percent of vessels were discovered that failed to report their vessel's name or call sign. Most importantly to this study, 17% of vessels were found to report a value of zero for their draft, and 14% reported drafts that were deeper than the length of the ship (Harati-Mokhtari et al. 2007). While some of these errors may be filtered and removed from analyses, the occurrence of errors demonstrate much of the vessel and voyage information input by hand are prone to error.

2.4 Government Crowdsourcing Programs

The maritime industry is already familiar with a crowdsourcing program called the Voluntary Observing Ship scheme (VOS). VOS is an international program that recruits and manages a fleet of commercial vessels that voluntarily transmit meteorological observations

(National Data Buoy Center 2009). These observations are distributed globally through the World Meteorological Organization's (WMO) Global Observing System (GOS). Ship observations are just one type of atmospheric data that GOS assimilates; upper air data, satellite soundings, and surface observations (to name a few categories of observations) are also available for use globally through GOS (World Meteorological Organization 2014)

The VOS program is organized internationally by a subsection of the International Oceanographic Commission and the WMO joint commission known as JCOMM, but run locally by individual member countries. Port meteorological officers stationed in ports across the globe maintain a fleet of ships outfitted with meteorological instruments by the program, assuring data quality. The officers also recruit new vessels into the program, adding to the number of observations available to forecasters, modelers, and researchers in traditionally data-sparse areas (National Data Buoy Center 2009).

2.5 Crowdsourcing Research and Programs Using AIS

Data reported and acquired by the AIS communications network is starting to prove its utility beyond collision avoidance and vessel traffic control. Several papers have been published over the past decade exploring and proving the data's usefulness is new and innovative ways.

A recent study using AIS data studied ship trajectories in New Zealand. Sampath and Parry's study revolved around using AIS point data to tease out meaningful information from the vast datasets they had available to draw conclusions about the types of vessels transiting through New Zealand waters. This study focused on ferries, passenger vessels, and high speed watercraft. These particular classifications were studied because their high rate of speed and ability to carry large numbers of passengers; these factors add an element of risk to any potential casualties at sea (Sampath and Parry 2013).

Much like this proposed AIS draft study, Sampath and Parry first dealt with the data management aspect of working with AIS data. Since the AIS messages in the study were collected every minute, millions of points accumulated over a short period of time in a small area. The team separated the AIS raw data sets into smaller temporal data sets to make the file sizes manageable. The raw AIS messages were then decoded and checked for errors and duplicate messages. This step was not necessary in this thesis because data were provided by NOAA and the USCG and they were already decoded and available in shapefile, geodatabase, and tabular formats. The data sets chosen for this study were limited to a 2GB file size so ArcGIS could handle the analyses.

Sampath and Perry's trajectory study used analytical tools in ArcGIS to calculate speed profiles for the different classes of vessels and study the characteristics of ships at anchor. The study concluded that movement patterns of vessels could be computed from spatial and temporal analyses, providing valuable insight into the patterns of vessel transits (Sampath and Parry 2013). This trajectory study encountered many of the same challenges that were experienced in this proposed AIS project, making it an excellent case to examine.

A Finnish research team developed a method to monitor marine traffic exhaust emissions using AIS data (Jalkanen et al. 2009). According to the authors, the results of their study can be used as a tool to make health and emissions policy decisions. For example, aerosol emissions, usually sulphate particles, pose health risks to residents of highly trafficked coastal areas such as along the English Channel and along the East Coast of the United States. The team created the Ship Traffic Emission Assessment Model (STEAM) to assess ship emissions pollution in the Baltic Sea. Using this model, AIS data are used to determine the position of vessels, their identification, and the speed at which they are sailing. The model then can match vessels with a volunteered engine and emissions profile database. The speed of the vessels also contributes to the engine operation information, leading to an estimate of the emissions at certain speeds. Wave data are also used in the STEAM model to help estimate the amount of ship fuel consumption in different sea conditions.

The limitations of the Jalkanen et al. study were mainly related to the ship and engine characteristics. Where possible, data regarding the engine and fuel were input into the database using parameters volunteered from the ship owner or from the Lloyd's database, which tracks information about commercial vessels. If these data were not available, assumptions were made about the engine based on the environmental conditions it faced at the time. The composition of the particulates at certain vessel speeds was an estimated value as well, although based on experiments on large commercial vessels.

Although the STEAM model proved to have limitations and inherent uncertainty built in, the study successfully used AIS data to prove that monitoring commercial ship traffic pollution is possible using crowdsourced data. Due to the high frequency of position reports via AIS, the authors concluded that the temporal and spatial resolution of their study was satisfactory for studying shipping and emissions trends.

Chinese researchers similarly studied real-time pollution from ship emissions and polluting discharge using AIS data (Qian et al. 2011). The group developed a real-time monitoring system that can track ships discharging hazardous waste such as oil, fuel, or other pollutants. The system can also display historical pollution data, perform statistical analyses, and run predictive models during events like oil spills to direct authorities to the most likely areas where cleanup and mitigation will be needed. The entire model is based on AIS data. AIS data are input and decoded, then fed into predictive models for pollution from vessels (atmospheric and hydrologic diffusion models) and how the pollution will move within individual harbors.

Given all of this demonstrated value in AIS data, the NOAA Office of Coast Survey has already started to use AIS positional data to identify areas of heavy traffic in support of updating nautical charts. NOAA's Arctic Nautical Charting Plan, released in February of 2013, includes one example of the use of AIS data. The positional data derived from AIS provided information about the shipping and transit trends in remote locations in Alaska and the Arctic. NOAA scientists confirmed that most arctic shipping and travel occurs during the summer, mainly from June through August using AIS point data and ship track lines. Analysts proposed using these temporal and spatial trends to prioritize and update arctic nautical charts (NOAA Office of Coast Survey 2013a).

2.6 Summary

Hydrographic survey prioritization takes many factors into account to plan the most efficient use of government resources, survey assets, and time to protect the life and property of sailors at sea and support international maritime commerce. Survey and navigation techniques and technologies have evolved over the years; this evolution has resulted in faster survey work, more precise and accurate bathymetry, and new sources of vessel traffic information for the scientific community. Survey prioritization is currently a qualitative process without extensive documentation. There are currently no concrete, documented standards for the quality of data are that used to make the decisions of where to conduct hydrographic studies in upcoming years.

AIS data and marine crowdsourcing have proved themselves useful in many traditional and non-traditional arenas such as ship navigation and pollution monitoring, respectively. The ability to acquire AIS by VHF signal and satellites makes it a powerful tool. The maritime community has shown that it is willing to participate in programs such as VOS and the experimental hydrographic programs. Together, there are many possibilities for gleaning information that is usually difficult to acquire from ships transiting through areas of interest. Many of these findings influenced the design of this study, as described in the next chapter.

CHAPTER 3: METHODOLOGY

To address the research objectives of this study, authoritative spatial data from multiple offices within NOAA National Ocean Service were used to calculate the under keel clearance values and additional non-spatial statistics necessary to evaluate operational risk at two ports used as case studies. This chapter presents the type of data, its sources, and how the data was used to calculate the under keel clearance and associated statistical values. The methodology and workflow in Figure 6 were used to calculate the under keel clearance and undertake risk analysis for this project.

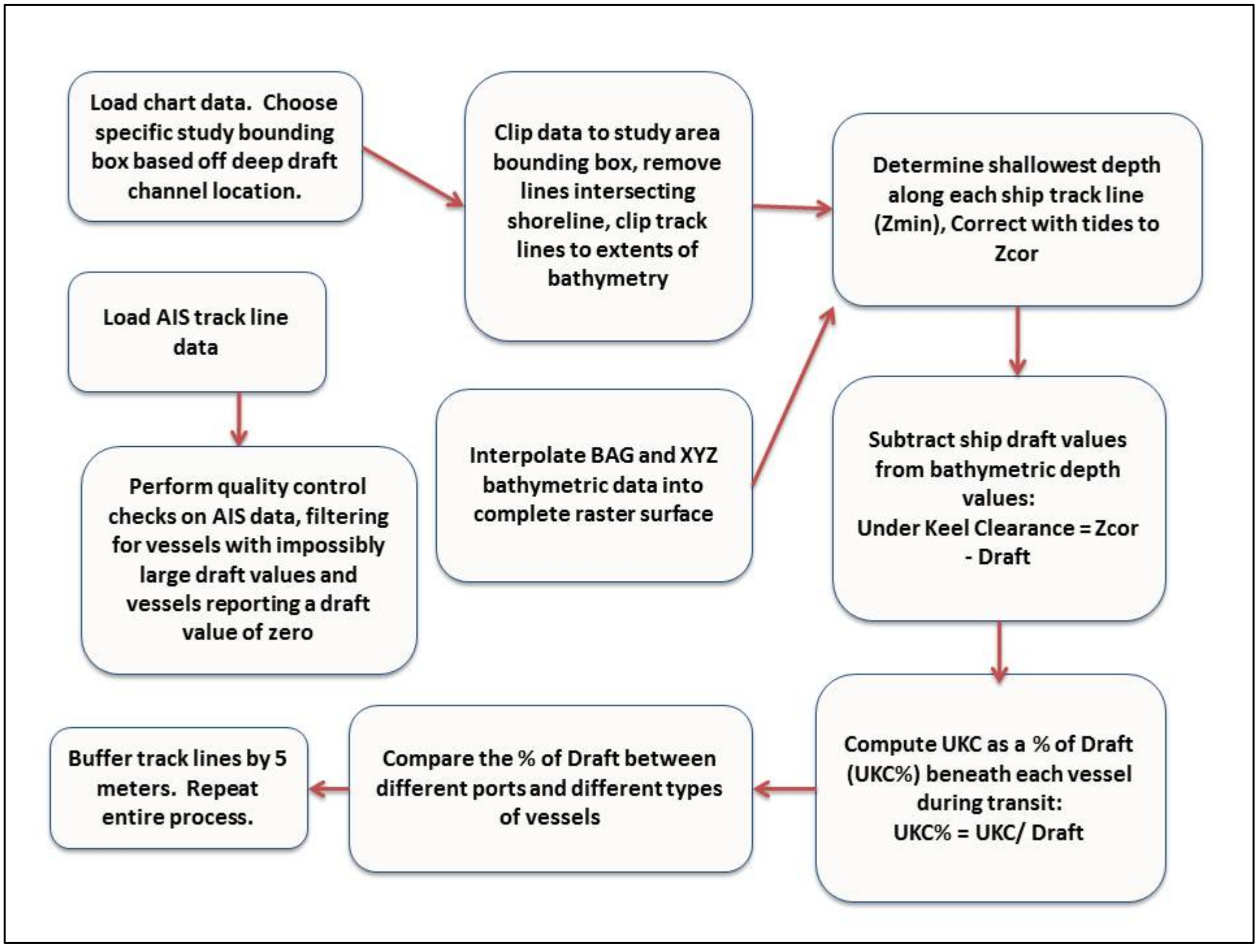

Figure 6 Methodology Diagram for Under Keel Clearance Computations and Analysis

3.1 Study Areas

Two study areas were chosen to be test cases for using AIS draft information and NOAA bathymetry to calculate and analyze ship under keel clearance. These study areas are the ports of Hampton Roads in Norfolk, Virginia and Los Angeles-Long Beach in California. They were chosen because these ports met the following criteria:

- United States ports: AIS data provided by the USCG available
- Major shipping hub: a variety of commercial and recreational ships transited through the area, guaranteeing a wide range of reported AIS draft values
- Deep draft channel: port terminal facilities accessed via deep draft channels that must be used by large commercial vessel traffic, such as container and tanker vessels that were constrained by where they could transit by their draft
- Located with 10 miles of a primary tide gauge: NOAA primary tide gauges are the most reliable type. Having a study area with a close proximity to a tide gauge minimized the uncertainty associated with under keel clearance values by using verified tide level corrections.

The Hampton Roads and Los Angeles case study locations provided comparable basins on the Atlantic and Pacific coasts to evaluate differences in under keel clearance values for normal vessel traffic. These values were used to assess the relative risk ships assume because of their draft when entering port. These case studies are representative only of one particular type of port and are not designed to be representative of the other ports of the United States that do have similar bathymetric, traffic, and structural qualities. Although these ports may not be representative of all ports, there is no indication that similar or identical methods would not be useful for other ports with different characteristics in the United States. However, if some of the datasets mentioned below are not available, data quality may be degraded and uncertainty will be introduced into the results.

3.1.1 Hampton Roads Case Study

Hampton Roads is the access point to the Port of Norfolk, Port of Portsmouth, and other points and terminals upstream in the James River, Elizabeth River, and Nansemond River. Hampton Roads comprises the entrance to these rivers from the Chesapeake Bay and the Atlantic Ocean and is a naturally occurring port area that is able to support ships with deep drafts. The Port of Virginia boasts six terminals, vessel berths dredged to support ships with 50-foot drafts, and convenient access to railways and highways (Port of Virginia 2015). Not only is this a commercial hub, but it also supports the United States Coast Guard Sector Hampton Roads (United States Coast Guard 2015) and Naval Station Norfolk. The naval station alone holds approximately 75 ships located on 14 piers, including large aircraft carriers (Military.com 2015).

The extents of the Hampton Roads case study area contained data within a box approximately bounded by the following latitude and longitude ranges: 37° 00’ 38” N and 36° 57’ 33” N latitude and 76° 17’ 24”W and 76° 21’ 00” W latitude. These extents are shown in Figure 7 by the pink boundary line. The extents of this study were chosen to include the narrow deep draft channel south of Old Point Comfort, which is the entrance into Hampton Roads from the Atlantic Ocean. The primary NOAA tide gauge associated with the Hampton Roads case study was Sewells Point. The tide gauge was established in 1927, shown in Figure 8, is routinely serviced by NOAA professionals. It is located on Pier 6 in Norfolk, within the extents of this case study (NOAA 2015c). Contemporary NOAA hydrographic surveys cover the study area as well (National Geophysical Data Center 2015a).

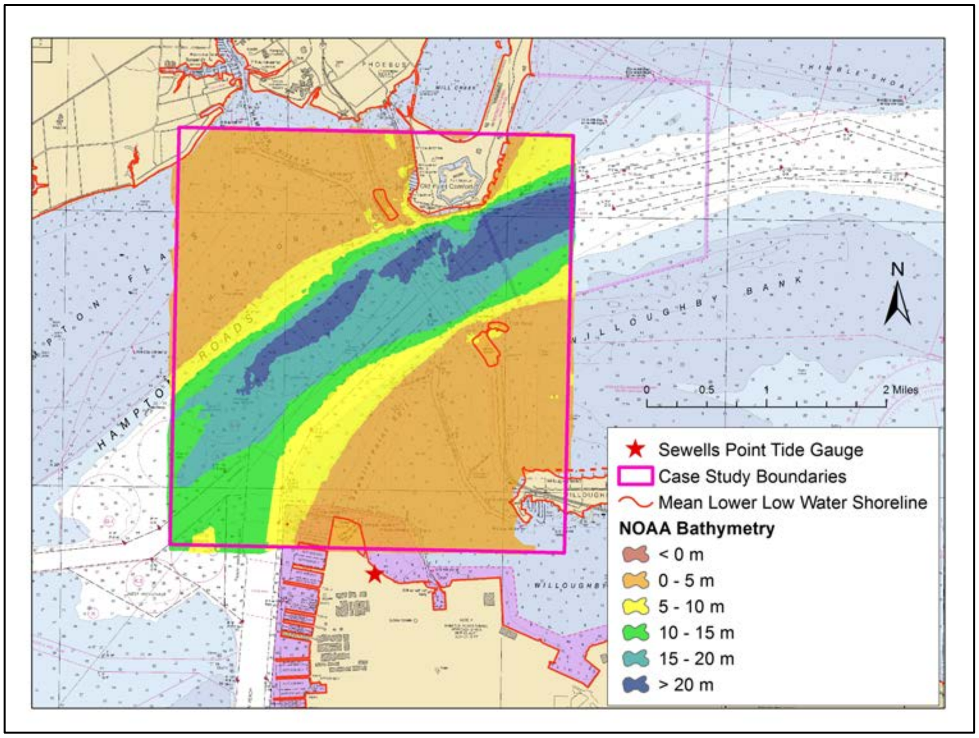

Figure 7 Hampton Roads Case Study Boundaries and Bathymetry

Figure 8 NOAA Primary Tide Gauge Installed at Sewells Point, VA (Source: National Oceanic and Atmospheric Administration 2015c)

Hampton Roads is an important waterway supporting one of the busiest ports on the East Coast. The availability of AIS data, supporting data necessary for quality control, and a blend of many types of vessels (commercial, military, recreational) makes it a prime location for this study.

3.1.2 Los Angeles/Long Beach Case Study

The Port of Los Angeles is spread out from the end of West Ocean Boulevard in Long Beach to Cabrillo Beach in San Pedro. The general geography of the port is illustrated in Figure 6. The majority of the actual port is located on Terminal Island, a man-made island that primarily houses commercial shipping and passenger cruise terminals. The port has 270 berths for commercial ships and enough marine space for 3,800 small boats. There are 8 terminals for container ships and 7 carrying liquid bulk, among the 23 ship terminals (Port of Los Angeles 2015a). The port also houses United States Coast Guard Station Los Angeles Long Beach (United States Coast Guard 2013).

The ship terminals and marinas in the Port of Los Angeles are located within a large breakwater, which was built in sections from 1871 to 1937 (Port of Los Angeles 2015b). This structure protects the deep water terminals and marinas from swell and waves from the Pacific Ocean. There are two main ship channels in and out of the breakwater and the port.

The extents of the Los Angeles case study area contained data within a box approximately bounded by the following latitude and longitude ranges: 33° 46’ 40” N and 33° 40’ 09” N latitude and 118° 09’ 43” W and 118° 17’ 53” W latitude. These extents are shown in Figure 9 below. The primary NOAA tide gauge associated with the Los Angeles case study was Los Angeles gauge. The tide gauge was established in 1923 and is routinely serviced by NOAA professionals. It is located in Berth 60 of Port of Los Angeles, within the extents of this case study. NOAA Ship *Fairweather* completed a hydrographic survey of San Pedro Bay in 2013. All data available from this survey were in the most complete digital format, the bathymetric attributed grid, explained in the next section (National Geophysical Data Center 2015b).

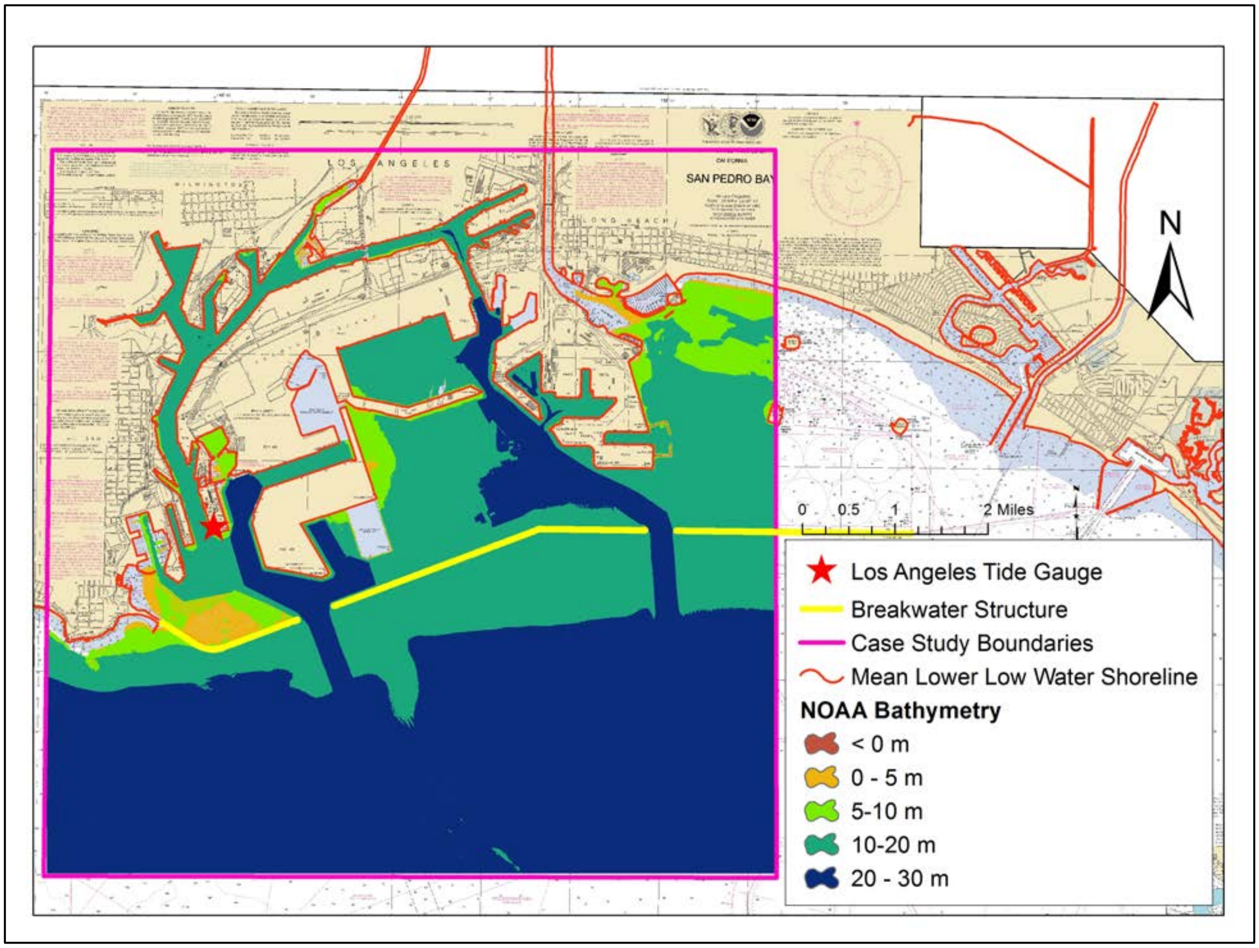

Figure 9 Los Angeles Case Study Boundaries and Bathymetry

The Port of Los Angeles provided an interesting contrasting case study to Hampton Roads because of the difference in the nature of the deep draft channel. Although Los Angeles' deep draft was not defined by natural features like Old Comfort Point and berthing in Norfolk, the extensive breakwater structure funneled all vessels, regardless of draft, into the narrow channels.

3.2 Data

Two primary sources of data were used to study the feasibility of using AIS data to assess under keel clearance to prioritize NOAA hydrographic surveys. The main data source was the AIS data from the USCG itself. AIS track lines made available by the USCG through a joint NOAA/Bureau of Ocean Energy Management web portal contain the positions and characteristics of all vessels transiting within United States waters. The second source was NOAA digital bathymetry data. Charted depths from NOAA provided the base data and depths that the ship drafts were compared against to calculate the under keel clearance of vessels within the case study boundaries. Additional data sets, presented below, were used to apply quality control and corrections for environmental conditions to the AIS data: mean lower low shoreline data and local tide gauge verified tide levels. Mean lower low shoreline data provided a boundary that delineated the interface of land and water. Any vessels intersecting this boundary could be assumed to have an incorrect horizontal position because they were technically on land. The local tide level data provided the dynamic correction to bathymetry since depths were charted on a specific datum that did not take into account changes in tide levels.

3.2.1 Automatic Identification System Track Lines

AIS point data are continuously collected by the USCG nationally to primarily assist with port security and search and rescue operations. Private companies also intercept and collect AIS point data. These data are sold in the form of pure AIS point and track line data, the same as is used in this study, or in the form of real-time online vessel tracking services. Prices for these data vary based on the company, support, and service provided.

Through an agreement with the USCG, NOAA is authorized to provide AIS point data to the general public via the MarineCadastre.gov website. MarineCadastre.gov is a joint NOAA – Bureau of Ocean Energy Management (BOEM) website that provides authoritative data for marine and offshore wind farm uses. Nearly 250 datasets are available to the public. To download data, users must register for a user name and password (MarineCadastre.gov 2015).

Public AIS data are available by month and by UTM zone for 2009 through 2012. The data is downloaded as a geodatabase with all necessary information included: points collected every one minute by broadcasting ships and data tables with vessel and voyage attributes. The attributes necessary to this study are the vessel’s draft, time stamp, vessel name, vessel type, and international MMSI number. Other information available is ship characteristics such as length and beam, port of departure, destination, estimated time of arrival, and other information not pertinent to this study. The MMSI numbers in the AIS data sets are scrambled to protect individual vessel privacy. While the numbers do not correspond to the numbers actually used by ships, all of the other attribute information is accurate (MarineCadastre.gov 2015).

For this study, AIS data were requested directly from the NOAA Office of Coast Survey. Three months of data from each case study were selected: February, June, and October from 2011. Because of the large file size of AIS data, only three months were requested. The limitations in file size were based on the analytical and storage capabilities of the computer systems used for analysis. Personnel at the NOAA Office of Coast Survey filtered the AIS point data to include only ships that were underway (excluding ships at anchor or alongside a pier) and converted the point data to individual track lines using the Track Builder script provided by MarineCadastre.gov. This script, designed to run with Esri ArcMap 10.1 software, used the ship identifier (MMSI number) and the date and time the AIS message was sent by the ship to create lines tracing ship voyages. This tool reduced millions of points into thousands of tracks, allowing data to be easily shared and manipulated on a personal computer with standard GIS software. The resulting track lines do not include any estimate of uncertainty that was introduced during the track line creation processing. NOAA personnel also provided all vessel and voyage attributes as a comma-separated value file (.csv) that was later added to the corresponding track line data via the vessel MMSI number (MarineCadastre.gov 2015).

The Hampton Roads case study had 3,438 AIS attributed track lines within the study area, and the Los Angeles case study had 18,554 track lines within the boundary of the study. The vast difference in vessel traffic was attributed to the size and infrastructure capabilities of both ports. Los Angeles had 23 commercial vessel terminals while the Port of Norfolk in Hampton Roads had 8 terminals. Since Los Angeles had nearly three times the number of terminals of Norfolk and the added passenger vessel traffic to Catalina Island, the near six-fold difference in vessel traffic was expected. In this study it is assumed that the difference in vessel traffic between the two ports does not influence the results because the time span was several months long and included three different months in 2011, creating data sets representative of typical traffic patterns for the two case studies. Table 1 shows the number of vessel tracks recorded in the case studies, broken down by the type of vessel as reported by the AIS signal.

Table 1 Case Study Track Lines Divided by Vessel Type

Vessel Type	Hampton Roads			Los Angeles		
	February	June	October	February	June	October
Anti-pollution	0	0	0	14	21	17
Cargo	379	415	439	923	989	938
Dredging	87	5	4	47	91	8
Fishing	0	3	4	108	165	163
High speed craft	0	0	0	304	621	467
Law enforcement	17	17	2	0	0	1
Military operations	17	34	38	7	28	20
Not available	77	155	201	275	306	429
Other type	15	14	17	124	228	242
Passenger	7	41	45	131	90	168
Pilot vessels	29	24	10	420	433	412
Pleasure craft	1	52	74	41	59	85
Port tender	0	0	0	58	19	98
Reserved for future use	0	0	0	99	46	8
Sailing	0	2	2	0	0	0
Search and rescue	9	9	4	0	0	0
Tanker – all vessels of this type	23	48	31	187	295	275
Towing	50	61	68	919	1170	1174
Towing – Length exceeds 250m	34	31	18	275	179	159
Tug	202	288	288	1248	1392	1355
Wing in Ground	1	5	3	0	2	0
All Vessels (Total)	**944**	**1224**	**1256**	**5183**	**6137**	**6026**

The total number of vessels from the different months for the two ports show a moderate seasonal signal. Vessel traffic in Hampton Roads was approximately 25% less in February than the summer and fall months, while it was reduced by approximately 15% in Los Angeles. This trend was mirrored in most of the classifications of vessels for both ports. The decrease in winter traffic was likely a manifestation of shipping patterns and recreational traffic adjusting to strong winter storms in the Atlantic and Pacific, which cause even the largest vessels to alter their course and speed to avoid dangerously high winds and waves.

3.2.2 NOAA Bathymetric Attributed Grids

Data acquired from the NOAA hydrographic surveys are available from the National Geophysical Data Center online archive. Contemporary surveys are stored and downloaded as bathymetric attributed grid (BAG) files, and older surveys are available as grid registered XYZ point data files. All NOAA surveys also have Descriptive Reports available for viewing and download. These reports are a detailed description of how the surveys were conducted, any problems that were encountered, and artifacts in the data. Since the data was already accepted after systematic and comprehensive quality control reviews by NOAA hydrographers and cartographers and applied to the navigational charts, any data acquisition problems and artifacts do not affect this study. Table 2 displays the surveys that were used in both case studies.

Table 2 Bathymetric Data Sources

Survey Number	Case Study Area	Survey Year	Survey Name
F00388	Hampton Roads	1994	Southern Chesapeake Bay Investigations
H06930	Hampton Roads	1944	Off Willoughby Spit, Virginia
H07171	Hampton Roads	1947	Hampton Roads, Virginia
H07824	Hampton Roads	1950	Old Comfort Pt, Virginia
H08878	Hampton Roads	1966	Hampton Flats, Virginia
H12617	Los Angeles/Long Beach	2013	San Pedro and Vicinity
H12618	Los Angeles/Long Beach	2013	Long Beach and Vicinity
H12619	Los Angeles/Long Beach	2013	Approaches to San Pedro

The two case studies have bathymetric data sources available from different decades. The Los Angeles data were acquired in 2013, while Hampton Roads had data from a wider range: from 1950 to 1994. With the expansive area NOAA is responsible for charting, surveys are not usually repeated within a decade to update navigational charts, unless the need is great. These data represent the most up to date data available, and are the soundings that are charted on the current navigational charts. When new surveys are completed the newest, most accurate data supersede the older surveys and are added to the charts (NOAA Office of Coast Survey 2012).

While the difference in survey age between the two cases is large, these data are still being used on the official NOAA navigational charts and are the depths mariners use to determine their track through the ports.

In order to create a complete bathymetric surface for the full extent of each case study that would combine charted depths recorded in the XYZ format into gridded data, it was necessary to use the same algorithms used by NOAA to create BAG surfaces. By using the same algorithms, bathymetry values directly matched the charted depth values and reduced uncertainty due to bathymetric data. The algorithms are proprietary and created by Caris, a marine GIS company. The NOAA Northeast Navigation Manager provided assistance with creating BAG surfaces directly from the XYZ and BAG survey information using Caris. This involved loading hydrographic data and shapefiles defining the case study limits into Caris BathyDatabase software. The depth and limiting files were used to create a triangulated irregular network (TIN), a 3-dimensional representation of the seafloor of the study areas. The TINs were then output as gridded depth files (BAG) with the same resolution as the original bathymetric files. A 4-meter grid was created for Los Angeles and 3-meter grid for Hampton Roads. These resolutions were chosen because they are the native resolution of the data provided by NOAA. The resolution provided is determined by water depth (NOAA Office of Coast Survey 2014). These BAG grids provided the water depths, which were the basis for calculating ship under keel clearance values in the study areas. The extents of the BAG files were also used to clip the track lines, cutting the tracks off where depth data was no longer available. Bathymetry for both case studies is displayed in Figure 10.

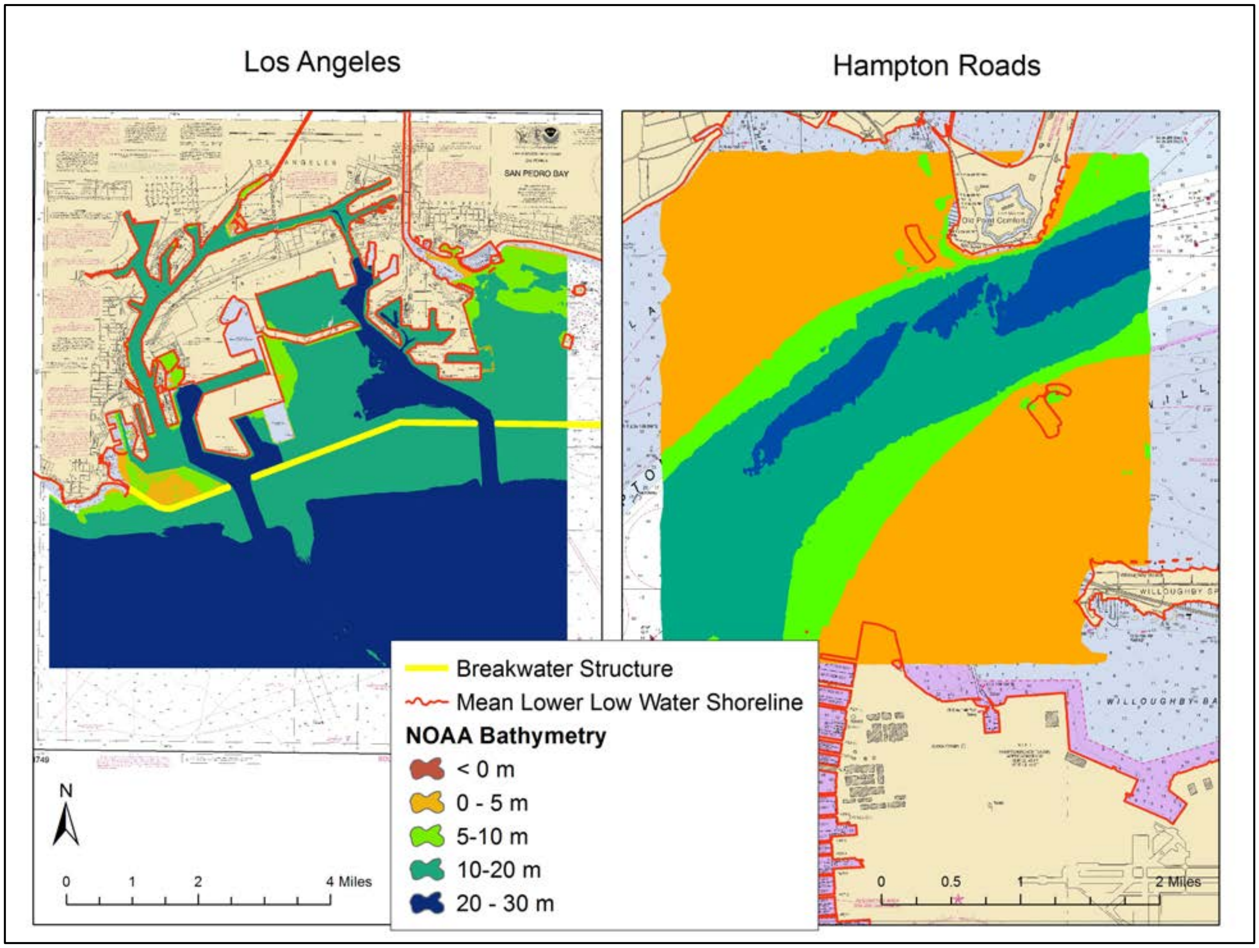

Figure 10 Bathymetry for Hampton Roads and Los Angeles Case Studies

The two images displaying charted bathymetric grids representing the seafloor in Los Angeles and Hampton Roads in Figure 10 show that Los Angeles is generally deeper than Hampton Roads and had multiple distinctive dredged channels leading into the ship's breakwater. Hampton Roads is a natural deep draft port following the flow of the James River. The differences in average depths factored into the under keel clearance results.

3.2.3 NOAA Shoreline

The shoreline on NOAA charts is defined as the mean high water line, which is the 19-year average of the highest daily tidal level. The mean high water datum is how NOAA defines the interface between charted water and land. Shoreline data were used to filter AIS track lines that reported positions across the NOAA shoreline, or over land. The shoreline dataset was acquired from the NOAA Continuously Updated Shoreline Project (CUSP). CUSP data may be downloaded in a variety of formats for United States and US territory shorelines. The CUSP shoreline data are compiled from imagery, shoreline vectors, and light detection and ranging (LiDAR) coastline surveys; these datasets are constantly updated. There is no version number associated with the CUSP dataset. The date the shoreline file was downloaded must be used to reference the version of the dataset in this study instead of a specific survey or version number (National Oceanic and Atmospheric Administration 2015b). These particular data were downloaded on February 20th, 2015 for Hampton Roads and February 28th, 2015 for Los Angeles. Shoreline data were downloaded as line shapefiles, vector files that were used to perform quality control on the AIS tracks by defining where lines intersect land. The CUSP shoreline data presented one measure of verifying or disproving the positional accuracy of the track line data and provided a means of removing lines that suffered large positional errors.

3.2.4 NOAA Verified Tides

Verified tide observations from primary NOAA tide gauges were used to correct clearance values. Charted depths on survey charts are referenced to the mean lower low water tidal datum, which is calculated from the average of 19 years of the daily lowest value in the tide cycle. Tide observations from the time each ship transited through the study areas were necessary to know the depth of the water, accounting for the tidal cycle at that time. All observed water levels from the Sewells Point (tide station 8638610) and Los Angeles (9410660) tide stations were downloaded as .csv files from the NOAA Tides and Currents products website (National Oceanic and Atmospheric Administration 2015d). Verified data have already been checked for quality and are the official tide levels. Tide levels reported by the hour in meters were downloaded for all days in February, June, and October 2011. These tide correctors were added to the depth data during the under keel clearance calculations and analysis in order to have the most accurate results possible. The maximum observed tide range was two meters above and below the mean lower low tidal datum. The adjustment to charted depths for tides was important for vessels that count on high tides to provide a safety buffer of clearance when transiting through ports.

3.3 Data Quality Control

Data errors and uncertainty are inherent to crowdsourced data. In order to reduce the uncertainty and the number of errors in the data used in this study, several quality control measures were taken. Data were reviewed with specific criteria to remove track lines that defy the normal operational limits of vessels. The criteria are discussed below. Corrections were applied to the track line and bathymetry data to create the most accurate dataset and reflect the environmental conditions the ships experienced during their transits into the ports of Hampton Roads and Los Angeles.

3.3.1 Reported Vessel Draft

The draft information contained in AIS track lines is entered in manually by the mariners operating the bridge electronics, and as a result, errors inherently exist. Several potential sources of error are: masters and mates forgetting to enter a draft information, resulting in a reported draft of zero meters; hitting the wrong button to enter an inaccurate number or forgetting the decimal point; and entering the draft value in feet instead of the internationally-mandated meter unit. By using the wrong unit, drafts are reported almost three times deeper than reality. The over-reporting of draft values creates artificially hazardous under keel clearance values.

Calculated under keel clearance values may be dangerously low for the vessel, and even negative, which implies that the vessel ran aground during their transit.

The quality control (QC) measures began by eliminating track lines with the most obvious error: draft data missing. All track lines were removed from the AIS datasets where the reported draft equaled zero. It could be assumed that no ship would ever have a draft of zero and that the ship's crew either neglected to enter a draft or entered zero in error. Track lines were also selected and removed that reported values over 20.0 meters (65.6 feet) for ships transiting through Hampton Roads. The maximum depth of the channel at the Port of Virginia is 15.2 – 16.8 meters (50.0 – 55.0 feet) (Wood 2012). Track lines were selected and removed that reported values over 24.0 meters (78.7 feet) for the Port of Los Angeles/Long Beach. The maximum depth in the main channel at the Port of Los Angeles/Long Beach is 19.8 meters (65.0 feet). These draft QC criteria exceeded the charted depths in both case study ports to allow for the heavily loaded ships that transit through the deep draft channels during high tides. It is unlikely that ships with these drafts would conduct business in these ports because they would risk grounding and would not be permitted to enter by harbormasters. Vessels with drafts deeper than the above criteria were assumed to have incorrect draft information entered and were removed from the track line datasets.

As part of the investigation of reported draft uncertainty, the distributions of draft values were also studied. The number of ships with each unique reported draft value in Hampton Roads was plotted against the draft, demonstrating the distribution of drafts among the study area and period. If a double bell curve resulted with the second peak draft value approximately 3 times the first peak draft value, this would indicate a high frequency of ship drafts entered using incorrect units (i.e. feet instead of meters). Such errors would result in vessels having negative minimum water depths and negative under keel clearances. When plotted, no clear trend appeared, indicating the draft errors were not consistent with units and a feet to meter correction could not be applied to improve draft data quality. This analysis is discussed in more detail in the Discussion and Conclusion chapter.

Draft values broken out into major types of commercial vessels also do not reveal a clear trend that would suggest the majority of crews enter feet instead of meters for their draft values. Figures11, 12, and 13 show the reported draft values for cargo, tanker, and tug/towing vessels for Hampton Roads. No double bell curve signature exists for cargo, tanker, or tugs. The tug and towing vessels had the greatest probability of displaying the double bell curve because most of these vessels originate from the United States, where the cargo and tanker vessels have a much greater percentage of internationally owned and operated ships. This possible correction cannot be applied to any of the vessel types in the case studies, leaving reported ship draft to be the greatest amount of uncertainty in the AIS dataset.

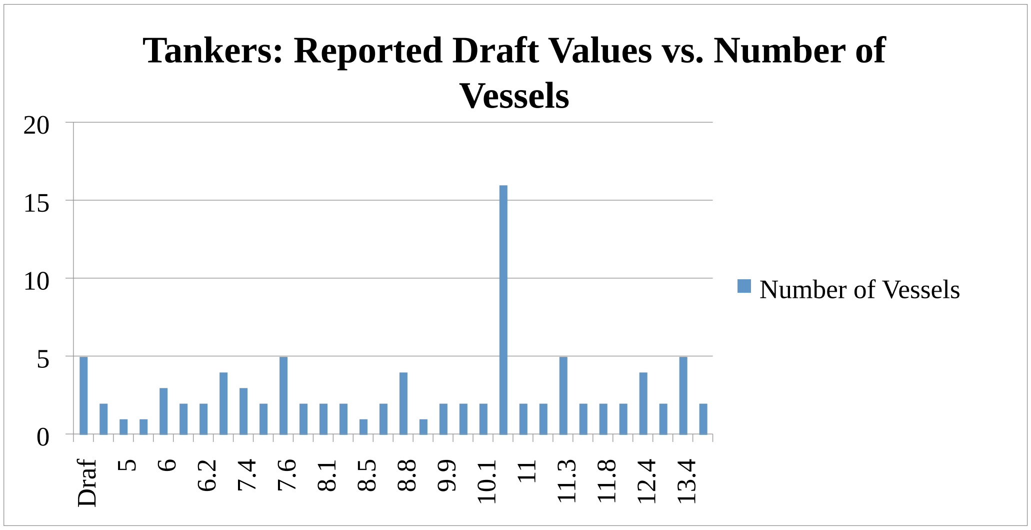

Figure 11 Reported Draft Values for Tanker Vessels Versus the Number of Vessels Reporting Each Draft Value, Hampton Roads

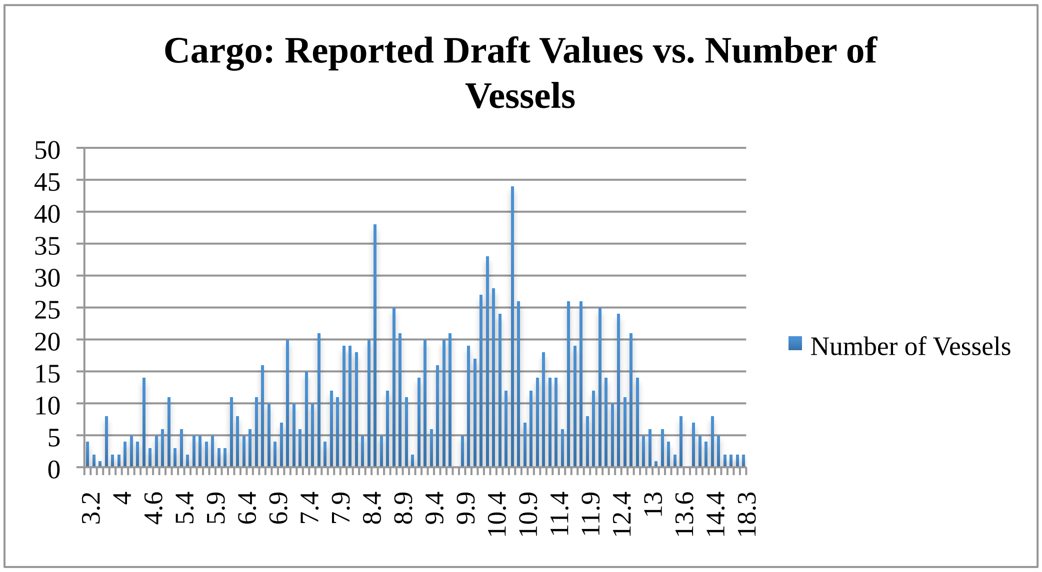

Figure 12 Reported Draft Values for Cargo Vessels Versus the Number of Vessels Reporting Each Draft Value, Hampton Roads

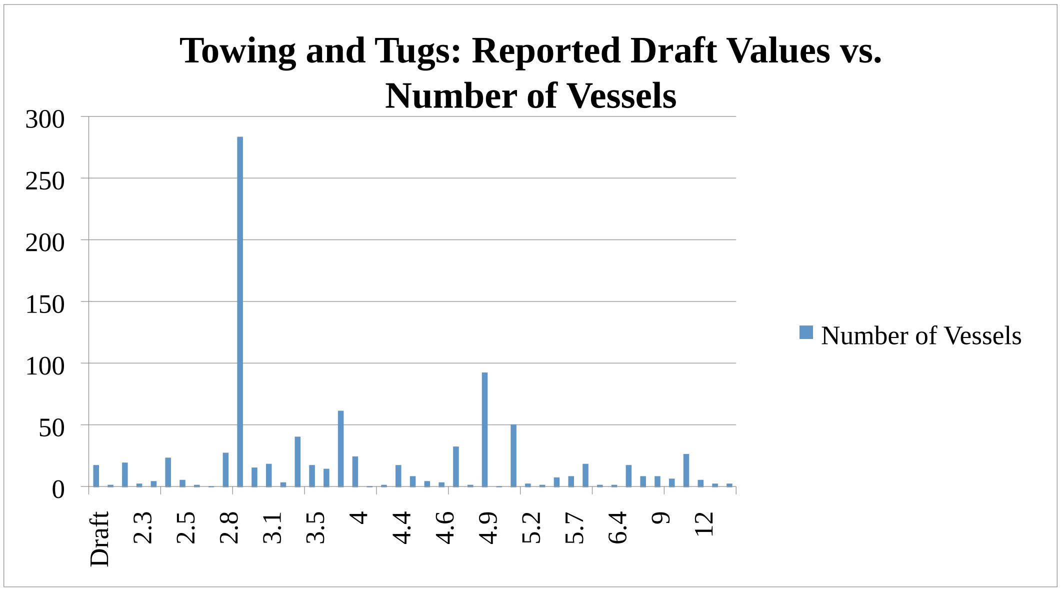

Figure 13 Reported Draft Values of Tug and Towing Vessels Versus the Number of Vessels Reporting Each Draft Value, Hampton Roads

3.3.2 Two-dimensional Horizontal Position

Marine GPS can experience the same positional errors as terrestrial GPS units. Multipath errors, degraded position during sub-optimal satellite geometry conditions, and signal loss can lead to minor or gross positional errors (latitude and longitude). Gross positional errors were eliminated from the AIS track line dataset by comparing ship positions with the NOAA charted shoreline data. Tracks that intersected the line indicating the shoreline were selected and removed from the dataset. GPS errors were the most straightforward to diagnose during this study because there were definitive data to compare them against (charted shoreline) and were reported from a sensor instead of a human, narrowing the reasons for large, obvious errors. Track lines remaining after erroneous lines intersecting shore and extending outside of the study boundaries were removed are shown in Figure 14 for Hampton Roads and Los Angeles.

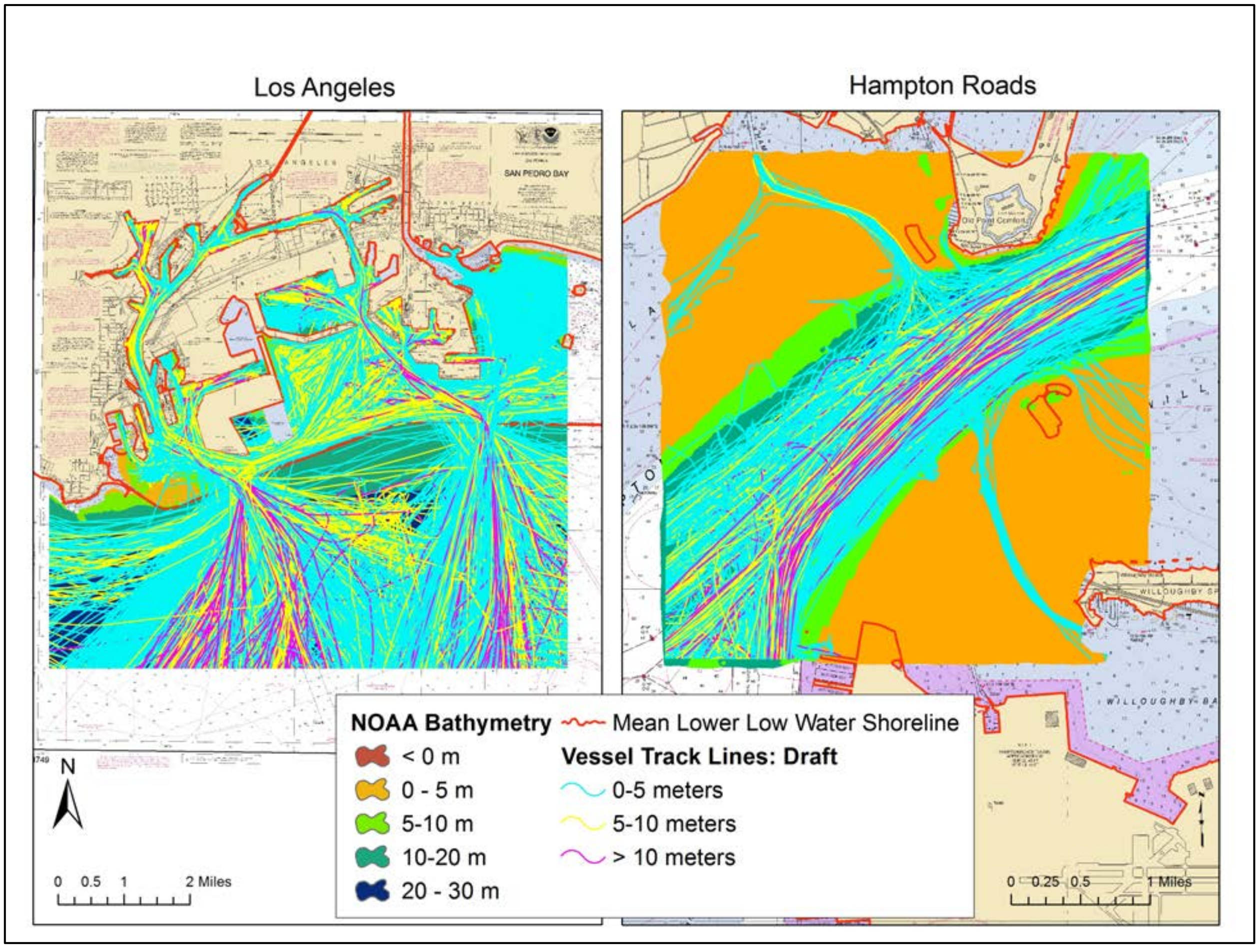

Figure 14 Hampton Roads and Los Angeles Vessel Track Lines by Reported Draft

Smaller positional errors were more difficult to detect and factor into the analysis. Because the ship's GPS accuracy is not reported along with the GPS position in this dataset, it is impossible to pinpoint track lines that have degraded positional data. A study completed by Januszewski surveyed the commonly installed marine GPS units and their accuracy. Out of the 21 units researched in the study, 17 reported accuracy better than 5.0 meters (16.4 feet) (Januszewski 2014). Thus to take the possibility of positional inaccuracies into account, a polygon was created that buffered each track line horizontally by 5.0 meters (16.4 feet). Five meters was chosen as the buffer distance to capture the average accuracy of marine GPS units, while maintaining a conservative estimate of accuracy. By creating polygons that represent a buffered line with uncertainty, positional errors difficult to detect such as poor satellite geometry or temporary loss of signal are captured in assuming that a ship may not exist at a certain point, but instead in a circle of uncertainty at any given time in the study. Portions of the buffered track line polygons are displayed in Figure 15.

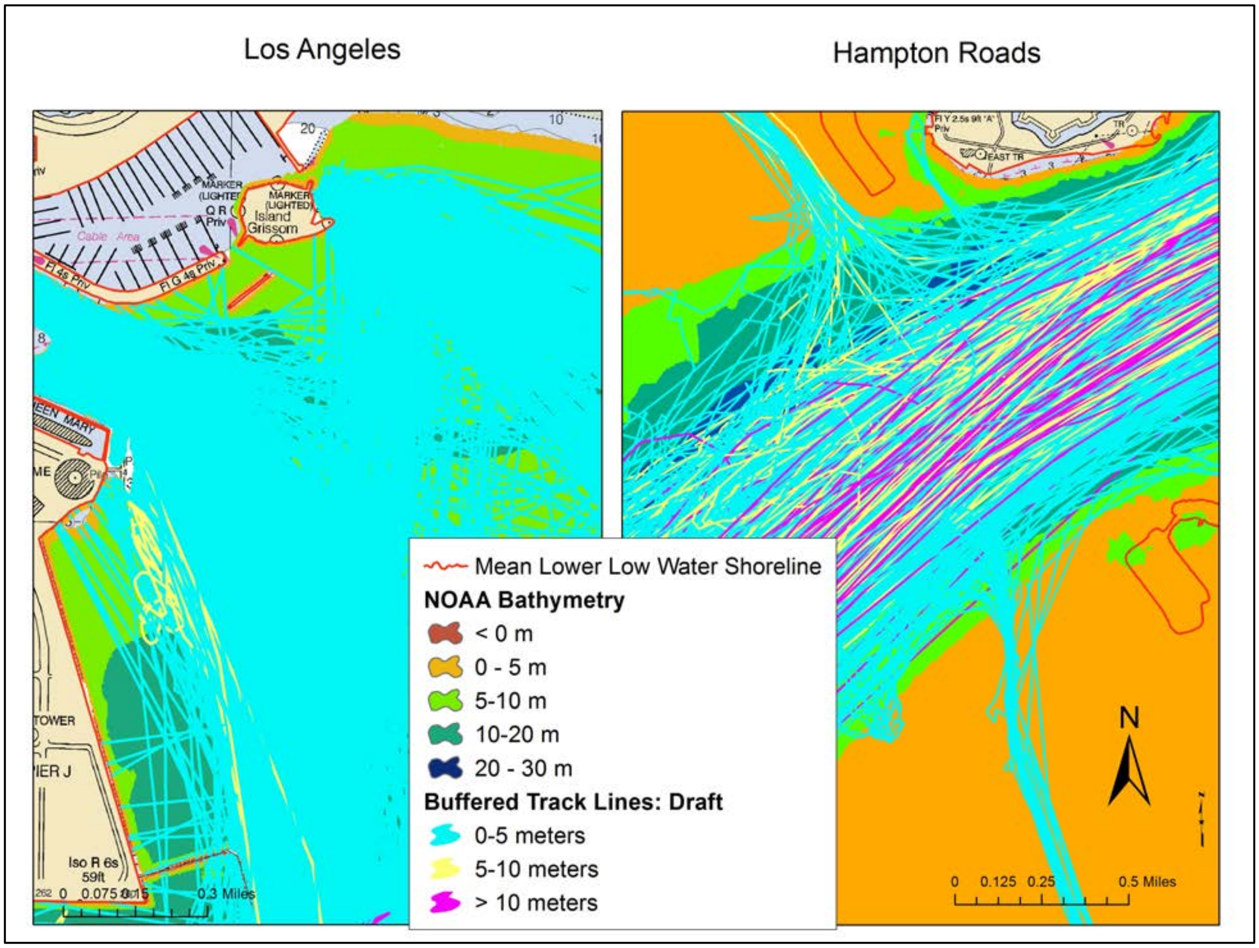

Figure 15 Hampton Roads and Los Angeles Buffered Vessel Track Line Polygons by Reported Draft

Finally, all retained track lines and buffered track line polygons, which represented lines with GPS uncertainty, were clipped to the extents of the available bathymetry so calculations of minimum water depth could be calculated. Segments of the lines and buffer polygons outside the bathymetric surface would not have a water depth for comparison and would report an inaccurate minimum water depth. These procedures addressed the horizontal accuracy of GPS receivers and prepared the track line data to be corrected for changes in the environment: tidal data creating positional errors in the vertical direction.

3.3.3 Water Level Correction

Additional data was needed to correct the position of the track lines in the vertical dimension because it could not be addressed by GPS uncertainty corrections made by the two-dimensional surface created by buffered track line polygons. The height of the tides in both case studies constantly changes throughout the day. While the tidal level does not significantly impact the operations of vessels with shallow drafts, the tides can play a large role in the decision-making process of large vessels such as tankers and cargo vessels that operate near the depth limits of the port. The forecast tide levels determine how much cargo is loaded onto these vessels and when they can plan to enter and exit port. During the Hampton Roads study period, the minimum observed tide was 0.213 meters below the mean lower low water datum and the maximum was 1.402 meters above mean lower low water. In the Los Angeles/Long Beach study, the minimum tide level was 0.473 meters below mean lower low water and the maximum was 2.15 meters above the mean lower low tidal datum. The tidal corrections to the bathymetric data minimized the uncertainty associated with the natural cyclical daily change of the water levels, but could not address other environmental conditions such as ocean swell, which impacts each ship differently.

Verified NOAA tides were used to correct the minimum water depth values that ships experienced during their transits through the case study ports. Because both the verified tide values and the bathymetric surfaces were referenced to the mean lower low tidal datum, the tide values could be added to the charted water depth to obtain the actual water depth values during ship transit. For any ship position within the study area, the tide correction values, determined by using the time of day the ship track line entered the study area, were added to the charted mean lower low water depths.

The values in Table 3 demonstrate the majority of track lines in Hampton Roads that were removed were due to incorrect draft values, and not due to inaccurate positions. In Los Angeles the track lines removed for positional errors nearly equaled those removed for draft errors. Although AIS data are internationally mandated for commercial vessels, there is little oversight or regulation, resulting in a dataset that has noise and errors such as the ones removed in this quality control process.

Table 3 Summary of Lines Removed in Quality Control Process

Quality Control Procedure	Hampton Roads Track Lines	Los Angeles/Long Beach Track Lines
Positional errors: tracks intersecting shoreline data	104	3562
AIS draft error: Draft = 0	440	3715
AIS draft error: Draft exceeds maximum values Hampton Roads > 20m Los Angeles > 24m	2	0
Tide correction not available	6	0
Percent of original tracks removed during QC process	16%	39%

3.4 Under Keel Clearance Calculation Procedures

Spatial and non-spatial procedures and analyses were necessary to assess the feasibility and utility of using draft data to help assign relative risk of different ports in order to set survey priorities. This chapter outlines the steps taken to take the raw AIS ship data and compare it with the charted NOAA bathymetry to produce a dataset that can be analyzed by NOAA hydrographers to begin to quantify navigational survey priorities based on ship under keel clearance.

The values necessary to understand and calculate under keel clearance are displayed in Figure 1. The ship draft is displayed as an orange line, showing the draft is how deep, measured from the surface, the hull of the vessel extends into the water. The charted minimum water level for a ship track or polygon, or the Zmin, is added to the tide level to result in a corrected minimum water level (Zcor). The under keel clearance (UKC) is the corrected water level (Zcor) minus the ship draft.

3.4.1 Minimum Water Depth (Zmin)

In order to calculate the minimum under keel clearance of any vessel it is essential to know the minimum water depth the vessel passes over as it transits through a port. The minimum water depth was found by comparing the location of each point comprising a track line with the raster value of the NOAA bathymetric surface at the same location. The minimum value found during the comparison was saved as an attribute of the line, becoming the minimum water depth (Zmin) value of the line.

This same comparison operation was also completed for the buffered track line polygons. All raster values corresponding with the buffered line polygons were queried for the bathymetric depth and the shallowest depth was assigned as the Zmin value attribute to the polygon.

3.4.2 Corrected Water Depth (Zcor)

Once the relationship was established between the track lines, Zmin values, and verified tide levels, the tide levels were added to the Zmin values, creating a corrected minimum water depth value for each track (Zcor).

3.4.3 Minimum Under Keel Clearance (UKC)

The minimum UKC was calculated for each line and polygon in both case studies by subtracting the draft from the Zcor, as shown in Figure 1. By subtracting the ship's draft from the corrected minimum depth it experienced during its transit, the amount of water left as clearance under the vessel, the UKC, was determined. Table 4 summarizes the calculations necessary to derive the UKC for each QC'd ship track line in the case study

Table 4 Calculations Used to Derive Under Keel Clearance (UKC)

Variable	Description	Equation
Zmin	Minimum uncorrected depth for each track line	Extracted from track position data and NOAA bathymetry
Zcor	Minimum water depth corrected using verified NOAA tide levels	Zcor = Zmin + Tide correction
UKC	Under keel clearance: the amount of water between the bottom of a ship's keel and the seafloor	UKC = Zcor - Draft
UKC%	Under keel clearance as a percentage of ship draft	UKC% = (UKC/Draft) * 100

3.5 Analysis

Once the Zmin, Zcor, and UKC values were calculated for each line and polygon the track lines and their attributes were exported as a table so additional calculations could be completed in a spreadsheet versus in GIS software. There are greater options for data comparison, calculations, and tabular/graphic output in spreadsheets than other software that were used for this study. These geographically generated and referenced values were the basis for the non-geographic analysis to follow.

3.5.1 Under Keel Clearance as a Percentage of Draft (UKC%)

The UKC was then used to calculate the percentage of each ship's draft left as UKC for the vessels (UKC%). This was a metric suggested by representatives from the NOS Marine Charting Division. By understanding and comparing the magnitude of the margins of safety being left by ships during their transits, relative risk between ports may be derived from the AIS track data. Table 4 shows the percentage calculation:

The resulting UKC% values were analyzed by port and commercial vessel type to assess trends, patterns, and dependencies within the dataset. In addition to this analysis, the results of the UKC% calculations that returned negative values were analyzed by port and commercial vessel type. Similarly, positive values were also analyzed by port and commercial vessel type in order to draw conclusions about sources of error and uncertainty within the dataset. The numbers and percentages of each type of vessel were broken down into categories of risk: negative UKC%, 0-10%, 10-15%, 15-20%, 20-50%, 50-100%, and greater than 100% UKC%. By breaking out the types of vessel in each port into these categories, hydrographers could quantify how many vessels were approaching, meeting, or exceeding their operational limits imposed by reported draft values while entering port.

3.5.3 Reported Draft Comparison

The reported AIS drafts of a small sample of vessels were compared with the ship characteristic information publicly available on a popular maritime website used by maritime industry enthusiasts and professionals: Vesselfinder.com. The privately owned and operated website hosts a database of characteristics of ships sailing internationally and also shows a real-time plot of ships via AIS broadcast data. The ships may be searched by name, which is a reported attribute of the AIS track lines. The commercial vessels with negative UKC% values reporting the deepest drafts from each major category (cargo, tanker, towing, tug) for the two case studies were compared with the information available online. Approximately half of the vessels with the deepest AIS drafts in each category reported draft information within one meter of the vessel characteristics reported by Vesselfinder.com. The other half of vessels did not have ship characteristics available in the Vesselfinder.com database. For the vessel drafts that could be verified, it can be concluded that the reported draft value was not the primary reason why the vessel resulted with a negative UKC% value. Ships reporting high draft numbers could not be automatically or reliably removed from the dataset without losing valuable information about the port and its deep vessel traffic. The results from the comparison of the drafts from vessels with negative UKC% values are shown in Tables 5 and 6 below.

Table 5 Hampton Roads Vessel Draft Comparison: Negative UKC Sample Vessels

Vessel Name	Reported Vessel Type	Reported AIS Draft (m)	Vesselfinder.com Draft (m)
Navios Pollux	Cargo	17.8	18.2
Golden Zhejia	Cargo	18.3	18.12
CGC Elm	Military	13	Not available
Kanawha	Military	13	Not available
Asir	Not available	12.6	11.5
Girrasol	Other type	20	Not available
Big Horn	Other type	12	12
Carnival Glory	Passenger	8.5	8.2
JMJ	Pleasure craft	6	Not available
JUSMAN	Pleasure craft	5.5	Not available
Shannon Dann	Reserved for future use	16	3.9
M/V Sunchaser	Reserved for future use	12.7	Not available
CG25403	SAR	3.3	Not available
SKS Tyne	Tanker	15.1	15.7
APL Cyprine	Tanker	13.5	13.4
Taft Beach	Towing	10.6	Not available
Chesapeake	Tug	16	Not available
Sea Raven	Wing in ground	8.4	Not available

Source: Vesselfinder.com (Vesselfinder.com 2015)

Table 6 Los Angeles Vessel Draft Comparison: Negative UKC Sample Vessels

Vessel Name	Reported Vessel Type	Reported AIS Draft (m)	Vesselfinder.com Draft (m)
Zoe	Anti-pollution	Not available	Not available
Hyundai Tokyo	Cargo	19.9	14.02
Chang Hang Ji Hai	Cargo	14.7	11.1
Catalina Jet	High speed craft	3.7	Not available
SPT Vigilance	Other Type	4	3.7
Catalina King	Passenger	3.3	Not available
Pilot Boat Polaris	Pilot	1.5	1.2
Crystal II	Pleasure craft	3	Not available
Patriot II	Port tender	3.5	Not available
A	Reserved for future use	3.5	Not available
Habari	Tanker	20	22.5
Genmar Victory	Tanker	18.6	19.02
Larcona	Towing	9	Not available
Patcona 2	Towing	10	Not available
Lynn Marie	Tug	17	13.4
Campbell Foss	Tug	16	Not available

Source: Vesselfinder.com (Vesselfinder.com 2015)

The negative track lines in Hampton Roads and Los Angeles did not show a distinct spatial pattern. The tracks with negative UKC% are displayed in Figure 16. Instead of indicating one or more areas where ships are exceeding their minimum depth limits, the track lines appear to be a subset of the total tracks without any additional trends. The lack of a pattern could be because the negative values are being caused by random draft data input errors, and the minimum UKC values are determined at a single point, yet assigned to an entire track line. In order to further understand the spatial pattern of negative track lines, individual points that compose a track line or smaller line segments should be used to find the minimum UKC for a vessel’s transit.

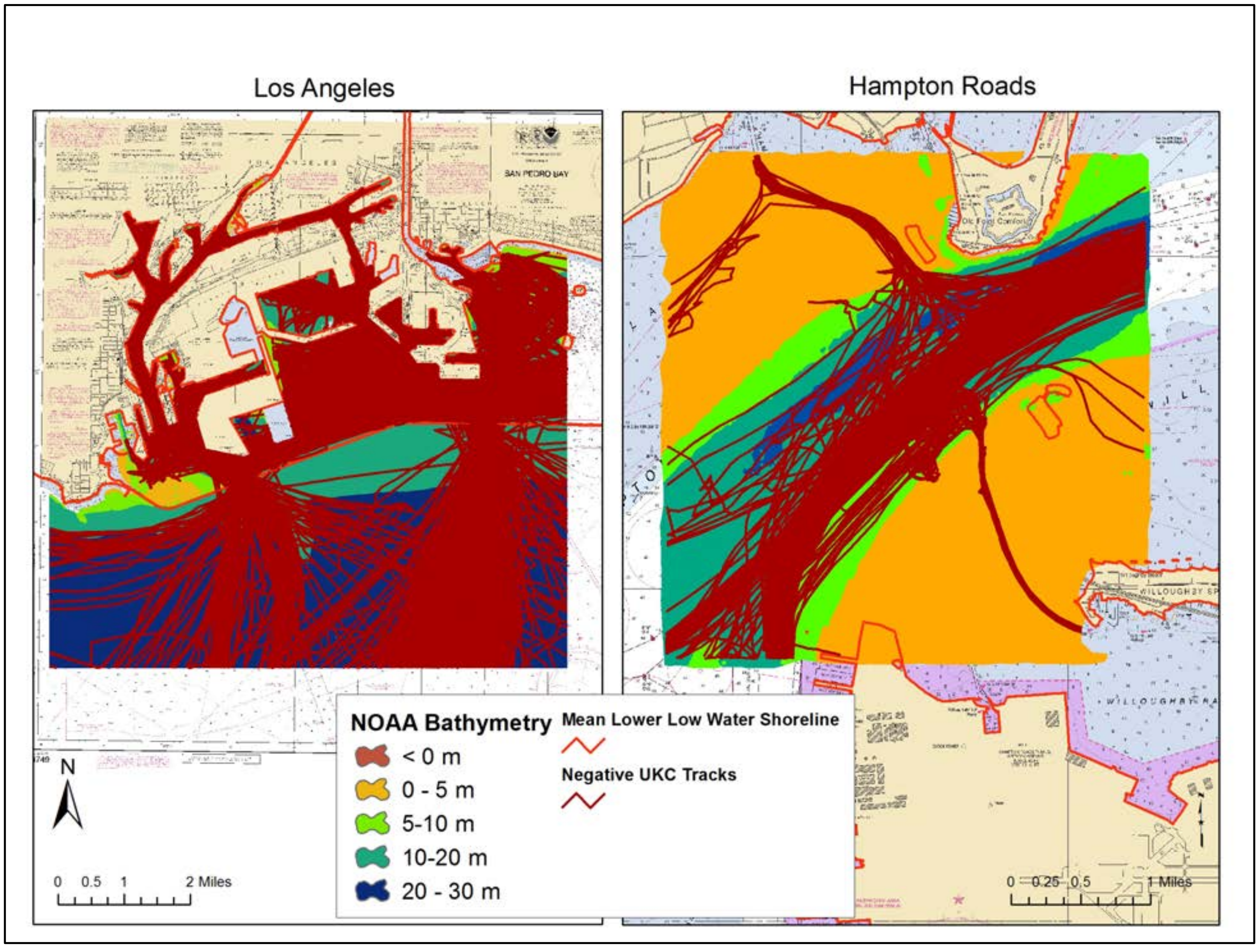

Figure 16 Track Lines with Negative UKC and UKC% Results

In summary, the data processing in this study involved quality control, calculation of key variables, and data manipulation of the resulting tabular data. Spatial data, originating from sources within the NOAA Office of Coast Survey and BOEM were subjected to quality control procedures, environmental corrections, and calculations using crowdsourced AIS data to output a non-spatial quantitative dataset of UKC values that will help hydrographers assess risk and assign hydrographic surveys to areas of need in the future. Track lines from vessels that reported positions on land and draft values equaling zero were removed from the study. After the minimum bathymetric depths were calculated for each line, these values were corrected using verified tide data to provide an accurate depth under each ship. Finally, the UKC and UKC% values were calculated and analyzed to provide hydrographers with quantitative comparisons of ship UKC in two busy commercial ports. The next chapter reviews the results from the calculations and demonstrates the results have errors apparent in the data despite quality control measures and environmental corrections, but can still be useful to hydrographers.

CHAPTER 4: RESULTS

The main objective for this study was to calculate and evaluate the UKC values and UKC as percentages of ship draft (UKC%) remaining under the keel for each vessel transiting through Hampton Roads and Los Angeles/Long Beach. These results were used to assess if the UKC% can be used by hydrographers to begin to quantify risks posed to mariners and help set annual hydrographic survey priorities. In this chapter, a method for using the results from these calculations to produce quantitative measures that useful in the comparison of risks at different ports is outlined. The following sections explain the results achieved from this study.

4.1 Negative UKC and UKC% Results

Despite the quality control measures taken to reduce the positional error and reported draft inaccuracies, the remaining errors that could not be easily filtered out or corrected were partially manifested in negative calculated values for Zcor, UKC, and UKC% for the ports of Hampton Roads and Los Angeles. The percentages of ships with negative values calculated were significant and could not be ignored in this study because they comprised approximately one fifth of the vessels included in each port. Because these percentages were so high, the number of vessels with negative values must be taken into account when the calculations from NOAA bathymetry and reported AIS position and draft are potentially used as a quantitative input into charting priorities. The types of vessels with negative calculated UKC and UKC% values provided additional information about vessel traffic in the two ports.

4.1.1 Negative Track Line Values

In Los Angeles, 1,688 of the 11,335 vessel track lines studied reported negative UKC and UKC% values: according to these values, 15% of vessels should have run aground during their transits. This is a signification number, indicating additional errors associated with the AIS data provided by the USCG and NOAA remain after quality control measures outlined in the Methods chapter were not sufficient to remove all erroneous tracks. To further study the reasoning behind the negative values, the negative tracks were separated by the vessel type, shown in Table 7 below.

For the negative UKC and UKC% tracks in Los Angeles, the majority of vessels registering negative UKC and UKC% values were tug vessels: 56% of the negative tracks belonged to tug boats. The next largest category belonged to tow boats, making up 22% of negative UKC vessel traffic lines. Tankers and cargo vessels combined only made up 5% of the negative tracks. High speed craft comprised 10% of the negative values for Los Angeles. Unlike tankers and cargo vessels, tugs and towing vessels may maneuver outside of deep draft channels and near man-made constructions and shallow waters to conduct their normal operations. Tugs and towing vessels, many of them possibly using the wrong units or inputting incorrect values for AIS drafts, are more likely to transit those shallow areas because they have relatively shallow drafts and may operate inshore of their ship or barge, resulting in negative UKC and UKC% values.

Of the 2983 vessel track lines that passed quality control measures in Hampton Roads during the study period, 703 of them, approximately 22%, had draft and Zcor values that resulted in negative UKC and UKC% values. This is number indicates the Hampton Roads data also suffers the same types of undetected errors as the Los Angeles data.

In the case of Hampton Roads, the majority of vessels registering negative UKC and UKC% values were cargo vessels: 69% of the negative tracks belonged to cargo ships. Surprisingly, the next largest category of negative track lines did not belong to tanker vessels, the commercial vessel type with similar draft characteristics. Tankers only made up 6% of the negative tracks, while passenger vessels comprised 14% of the negative values for Hampton Roads. Tankers have much deeper drafts than passenger vessels, but it is more likely that passenger vessels would transit through shallower water depths.

Table 7 Percentages of Negative UKC and UKC% Percentages by Vessel Type for Hampton Roads and Los Angeles

Vessel Type	Hampton Roads		Los Angeles	
	Track Lines: Negative UKC%	**Buffered Track Line Polygons: Negative UKC%**	**Track Lines: Negative UKC%**	**Buffered Track Line Polygons: Negative UKC%**
Anti-pollution	0%	0%	0%	0%
Cargo	69%	68%	2%	3%
High speed craft	0%	0%	10%	11%
Military	1%	2%	0%	0%
Not available	0%	0%	0%	0%
Other type	2%	2%	0%	0%
Passenger	0%	0%	4%	5%
Pilot	0%	0%	0%	0%
Pleasure craft	5%	6%	0%	0%
Port tender	0%	0%	0%	0%
Reserved for future use	2%	2%	1%	2%
Search and Rescue	0%	0%	0%	0%
Tanker	6%	6%	3%	3%
Towing	1%	1%	22%	22%
Tug	14%	14%	56%	53%
Wing in ground	0%	0%	0%	0%

While Hampton Roads and Los Angeles show similar overall percentages of negative UKC and UKC% values, the ships that comprise the negative track lines are quite different. In Hampton Roads where the bathymetry is more dynamic and the deep draft channel is natural, the large cargo vessels report the greatest amount of negative values. In contrast, the tug vessels, which primarily escort large vessels in and out of harbor, make up the majority of negative UKC and UKC% values. The deep draft channels are dredged, and outside the channels the bathymetry is also flat and deep, with the exception of the terminal, channel marker, and breakwater structures. Additional studies conducted in other ports may provide insight into if the percentages of negative UKC and UKC% remained relatively similar percentage-wise for ports overall.

4.1.2 Negative Buffered Track Line Polygon Values

The UKC and UKC% results from the buffered track line polygons also resulted in a significant number of negative values, with only slight differences from the track line results. Both ports had an increase in negative values: 24% of Hampton Roads buffered line polygons and 16% of Los Angeles buffered line polygons displayed negative UKC and UKC% values. This difference can be accounted for by the extra 5 meter buffer area on either side of the track line. When maneuvering in narrow channels, areas of rapidly changing depths, or near man-made structures, more vessels are likely to show shallower Zcor depths than from the original track lines alone. Table 8 shows the categories of negative polygon results for Hampton Road and Los Angeles.

When the polygons with negative values are categories by vessel type, the results also mirror the categories of ships reporting negative values for the original track lines. For both Hampton Roads and Los Angeles, the individual values generally only vary by a few percentage points. The overall percentage of these combined vessel types remained nearly the same, but relative to each other, the ratio shifted. All of the other types of vessels had percentages change by 1% at most from lines to buffered line polygons.

The small changes in vessel numbers help indicate positional errors are unlikely to contribute significantly to AIS track analysis. Table 8 below shows the number of ships in each vessel type category that displayed negative UKC% for the track lines and for the buffered polygon tracks, for Hampton Roads and for Los Angeles.

Table 8 Comparison of Vessels with Negative UKC% Values in Hampton Roads and Los Angeles

Type of Vessel	Hampton Roads		Los Angeles	
	Number of Negative UKC% Recorded for Track Lines	Number of Negative UKC% Recorded for Buffered Polygon Tracks	Number of Negative UKC% Recorded for Track Lines	Number of Negative UKC% Recorded for Buffered Polygon Tracks
Anti-pollution	0	0	1	2
Cargo	483	477	38	49
Military	10	11	0	0
High speed craft	0	0	173	206
Not available	3	3	0	0
Other	12	12	6	5
Passenger	1	2	76	91
Pilot	0	0	4	4
Pleasure craft	37	43	3	5
Port tender	0	0	4	5
Reserved for future use	13	12	24	30
Search and rescue	2	2	0	0
Tanker	39	40	51	52
Towing	7	7	368	412
Tug	95	96	940	985
Wing in Ground	1	1	0	0

The number of vessels with negative UKC and UKC% values increased when the track lines were expanded into larger areas by 5 meters of either side, simulating the uncertainty in the reported GPS positions. Although the numbers increased, the ports of Hampton Roads and Los Angeles showed similar trends when comparing the negative UKC and UKC% values of track lines and polygons. The greatest shift was seen in the relative percentages of tug and towing vessels in Los Angeles, but the prevailing trend remained unchanged: tug vessels remained the greatest percentage of negative lines and polygons, and towing vessels had the next highest percentages. By reviewing the results and recognizing how the percentages of negative tracks and polygons remained nearly the same for the case studies, horizontal GPS positional errors can be ruled as a small source of error for using AIS data for the purpose of hydrographic surveys. However, the number of track lines that reported with negative UKC and UKC% values is significant and must be considered when using AIS data for quantitative analysis of survey priorities. These values are still important to include because although many of the negative values may be errors, these track line values may also indicate a shift in bathymetry and local knowledge being used for route planning, or high risks being taken by ship operators in certain areas.

4.2 Track Lines: Positive UKC and UKC% Results

The following section is an analysis of the track lines and buffered track line polygons that had positive Zcor and UKC values in the Hampton Roads and Los Angeles case studies.

4.2.1 Hampton Roads

The Hampton Roads case study resulted in the tracks of 2310 positive vessel draft comparisons with the charted depths, or 77% of vessels. Of the vessels and track lines in this dataset, 3% of vessels allowed 0 to 5% of their draft values for UKC, and another 3% left 5 to 10% of their draft as UKC. These are considered categories of high risk to vessels and are of interest to hydrographers. The vast majority of vessels (41%) had water depths of over 100% of their draft under their keel during the transits. Table 9 shows the findings for the entirety of quality-controlled track lines in Hampton Roads in February, June, and October 2011.

Table 9 Calculated UKC as a Percentage of Vessel Draft (UKC%), Hampton Roads

Range of UKC% Values	UKC% All Vessels	UKC% Cargo Vessels	UKC% Tanker Vessels	UKC% Tug and Towing Vessels
0-5%	3%	6%	3%	1%
5-10%	3%	6%	12%	1%
10-15%	3%	5%	4%	0%
15-20%	3%	5%	1%	1%
20-50%	13%	21%	6%	6%
50-100%	11%	13%	18%	8%
> 100%	39%	7%	14%	73%

These data indicate that most of the ships transiting through Hampton Roads are maintaining safe under keel clearances at all times. A small percentage of the total vessel traffic operate in a high risk environment, leaving 10% of less of their draft value under their vessel during transits. These results are displayed in Table 9.

4.2.1.1 Hampton Roads Track Lines by Vessel Type

To further explore the vessel traffic and UKC trends of Hampton Roads, the UKC% was separated into three major vessel types: cargo, tanker, and tug/towing. These categories are a compilation of several vessel types as reported by AIS. The cargo category includes all cargo vessels, including those reported as hazardous, cargo that is reserved for future use, and cargo vessels that do not include any other information. The tanker category is similar to cargo, including hazardous classifications, future use, and tankers without other information. The tug/towing category includes all tug vessels, towing vessels, and tows that exceed 250 meters.

The UKC% for cargo vessels differs drastically from the trends for all vessels transiting through Hampton Roads. Table 9 shows that a significant number of vessels are in the high risk categories of 0-5% and 5-10% of draft remaining as UKC during the transits. Of the cargo vessels in Hampton Roads, 6% report 0-5% UKC, and 6% report 5-10% UKC. Combining these values, 12% of cargo vessels in Hampton Roads were considered to assume high risk for running aground during their transits. While this is a large number, due to the average dimensions of cargo vessels, this percentage was expected and may be used to compare cargo vessel risk relative to other ports.

Similar to the Hampton Roads cargo vessels, the tankers also displayed greater high-risk UKC% values than the entirety of vessels in the case study. This was expected due to the nature of tanker vessels: they have deep drafts that change due to changing loads, and are known to push the limits of their vessels and the port to maximize shipping efficiency and profits. Table 9 shows and 15% of the vessels fell within the 0-5% and 5-10% UKC% categories during the analysis. Tanker vessels displayed only slightly greater risk trends than cargo vessels. Tankers and cargo vessels have similar dimensions need to transit through the same ship channels and into similar terminals in the Port of Norfolk.

The tug and towing vessels in Hampton Roads display a very different array of UKC percentages than the cargo and tanker vessels. In the high risk categories of 0-5% and 5-10% UKC, only 2% of tug and towing vessels transit through Hampton Roads. The majority (71%) of tugs and towing vessels have at least 100% of their draft depth under their keel during transit. The results in Table 9 are expected, as the drafts of tug and towing vessels are significantly shallower than the drafts of the cargo or tanker vessels. However, these vessels are much more maneuverable than the other commercial vessels studied and may be themselves in situations of high risk UKC% situations where getting close to shore to dock or help maneuver a larger vessel into it’s berth.

4.2.2 Los Angeles

The Los Angeles/Long Beach case study resulted in the tracks of 9647 positive vessel draft comparisons with the charted depths, or 84% of vessels included in the case study In the Port of Los Angeles/Long Beach complete dataset, 1% of vessels allowed 0 to 5% of their draft values for UKC, and another 1% left 5 to 10% of their draft as UKC. This was a smaller percentage of vessels in categories of high risk to vessels compared with Hampton Roads. The vast majority of vessels (43%) had water depths of over 100% of their draft under their keel during the transits. Table 10 shows the findings for the entirety of track lines surveyed in Los Angeles for February, June, and October 2011.

Table 10 Calculated UKC as a Percentage of Vessel Draft (UKC%), Los Angeles

Range of UKC% Values	UKC% All Vessels	UKC% Cargo Vessels	UKC% Tanker Vessels	UKC% Tug and Towing Vessels
0-5%	1%	< 1%	2%	1%
5-10%	1%	< 1%	2%	< 1%
10-15%	1%	1%	4%	< 1%
15-20%	2%	2%	3%	1%
20-50%	13%	27%	24%	6%
50-100%	23%	38%	32%	16%
> 100%	43%	29%	27%	51%

Nearly three quarters of all vessels transiting through Los Angeles left at least one half of their vessel draft as clearance beneath their vessels. This is considered to be a safe and conservative clearance by NOAA hydrographers in the OCS. The risk comes with the relatively small percentage of vessels beneath that number – the types of vessels that make up those percentages are explored in the following sections.

4.2.2.1 Los Angeles Track Lines by Vessel Type

As with the Hampton Roads case study, the UKC% was separated into three major vessel types to explore vessel traffic trends: cargo, tanker, and tug/towing. These categories are a compilation of several vessel types as reported by AIS, the same as used in the Hampton Roads classifications.

The UKC% for cargo vessels differs drastically from the trends for all vessels transiting through Hampton Roads. The number of vessels in the high risk categories of 0-5% and 5-10% of draft remaining as UKC during the transits is on par with the trends for all vessel traffic in the port, falling between 0 and 2%. Of the cargo vessels in Hampton Roads, 0% report 0-5% UKC, 1% report 5-10% UKC, and 1% report 10-15% UKC. The vast majority of cargo vessels transiting through Los Angeles operate within typical safe UKC risk tolerances. Only 3% of cargo vessels were calculated to have high-risk UKC% values. Except for a few vessels, cargo ships in Los Angeles kept a safe and conservative amount of water under the keel of their vessels during transits.

Los Angeles tanker vessels had greater percentages of negative and high-risk UKC% values than the Los Angeles cargo vessels, but had lower negative percentages and nearly double the high-risk values of the port as a whole. Although tankers have deep drafts that change due to changing loads, it appears few ships need to push the limits of their vessels and the port to maximize shipping efficiency and profits in Los Angeles. Table 10 shows that 7% of UKC% values were negative, and 4% of the vessels fell within the 0-5% and 5-10% UKC% categories during the analysis. Overall, 11% of vessels were considered to have high risk UKC values, and 4% moderate risk (10-20% UKC%); few tankers reportedly enter Los Angeles without conservative UKC% values, according to AIS draft information.

The tug and towing vessels in Los Angeles form a very different pattern of UKC percentages than the cargo and tanker vessels. Similar to the cargo vessels, the high-risk categories of 0-5% and 5-10% UKC have only 2% of tug and towing vessels. Half (50%) of tugs and towing vessels have at least 100% of their draft depth under their keel during transit. The positive values in Table 10 are expected; tug and towing vessels have shallower drafts when compared with the drafts of the cargo or tanker vessels.

4.3 Buffered Track Line Polygons: Positive UKC and UKC% Results

This section provides an analysis of the buffered track line polygons, created to account for horizontal uncertainty of marine GPS. It also examines the ranges of UKC% values calculated during this study.

4.3.1 Hampton Roads

The distribution of track lines that fell into different percentages of clearance when divided by draft was nearly identical for buffered lines when compared with the original track lines. The majority of lines buffered by 5 meters on each side of the line had clearance values over 100% of the draft values A total of 6% of vessels allowed up to 10% of their draft for UKC, and an additional 6% allowed 10-20% of the draft of UKC. As with the original track lines, a significant number of vessels indicated they took high and medium risks with their UKC when transiting through Hampton Roads. These results are shown in Table 11.

Table 11 Caclulated UKC as a Percentage of Draft (UKC%) for Hampton Roads Buffered Track Line Polygons

Range of UKC% Values	UKC% All Vessels	UKC% Cargo Vessels	UKC% Tanker Vessels	UKC% Tug and Towing Vessels
0-5%	3%	5%	4%	< 1%
5-10%	3%	6%	10%	< 1%
10-15%	3%	5%	4%	< 1%
15-20%	3%	4%	1%	< 1%
20-50%	13%	20%	9%	6%
50-100%	11%	13%	15%	8%
> 100%	40%	7%	15%	73%

4.3.1.1 Hampton Roads Buffered Track Lines by Vessel Type

For the Hampton Roads buffered lines the UKC% was separated into three major vessel types to explore vessel traffic trends: cargo, tanker, and tug/towing. These categories, shown in Table 11, are a compilation of several vessel types as reported by AIS, the same as used in the Hampton Roads and Los Angeles original track line classifications.

The buffered lines for cargo vessels only also reflected the results of the non-buffered lines, but minor differences were seen in this data set. This result was expected because a positional error of 5 meters can place a ship inside or outside of a navigational channel or into water depths too shallow for a vessel’s draft. Cargo vessels generally operated in areas that did not have great seafloor slopes, so a 5 meter positional error did not have an significant effect on the Zcor values or UKC%.

Tanker vessels reporting clearances from 50 – 100% of their drafts decreased by 3%. The middle values of UKC% that changed the most, slightly shifting toward shallower values. The 20-50% UKC% range increased by 2% as the 50-100% UKC% range decreased by 3%. A similar shift is seen in the 0-5% and 5-10% UKC% ranges as more vessels reported shallower Zcor values over the buffered area as opposed to the more narrowly-defined original track lines. These shifts were expected. Although ships operated in areas of slowly changing seafloor slope, the bathymetry did slightly change within 5 meters of the vessel’s position. The extra 5 meter horizontal position difference could have changed the depth of the Zcor, caused the UKC% to decrease, and increased the risk the vessel took during its transit.

Adding the 5-meter buffer to the tug and towing vessel track lines did not impact the distribution of UKC% values. The majority of vessels still reported leaving over 100% of the vessel draft as UKC (73% of tugs and towing vessels. There were very few vessels that fell into the 0-10% or 10-20% UKC% ranges.

4.3.2 Los Angeles

The distribution of UKC% for Los Angeles buffered track line polygons follows the same pattern as the original track lines. The majority of lines buffered by 5 meters on each side of the line had clearance values over 100% of the draft values (39% of vessels). The next-largest category, 24% of vessel tracks, had clearance values from 50-100% of vessel draft values. A total of 2% of vessels allowed up to 10% of their draft for UKC, and an additional 3% allowed 10-20% of the draft of UKC. These results are shown below in Table 12.

Table 12 Calculated UKC as a Percentage of Draft (UKC%) for Los Angeles Buffered Track Line Polygons

Range of UKC% Values	UKC% All Vessels	UKC% Cargo Vessels	UKC% Tanker Vessels	UKC% Tug and Towing Vessels
Negative	17%	2%	7%	26%
0-5%	1%	< 1%	2%	1%
5-10%	1%	1%	2%	1%
10-15%	1%	< 1%	4%	1%
15-20%	2%	3%	4%	1%
20-50%	14%	29%	23%	7%
50-100%	24%	37%	32%	19%
> 100%	39%	28%	26%	44%

4.3.2.1 Los Angeles Buffered Track Lines by Vessel Type

The Los Angeles buffered lines were separated into three major vessel types to explore vessel traffic trends: cargo, tanker, and tug/towing. These categories are a compilation of several vessel types as reported by AIS, the same as used in the Hampton Roads and Los Angeles original track line classifications.

The buffered track line representing cargo vessel movements in Los Angeles harbor had almost an identical distribution of UKC percentages compared with the original track line values. There was a slight shift of UKC% values toward the shallower values with high risk – the shift was only 1% for several categories. Approximately 97% of cargo vessels had safe UKC% values, leaving over 20% of their draft values for UKC during transit. The difference in UKC% values between the track lines and buffered polygon tracks can be accounted for by slight slopes in the charted bathymetry; the large percentage of low risk UKC% values were due to the regular use of dredged deep draft channels by vessels

The buffered tanker lines exhibited slight changes similar to the buffered cargo lines in Los Angeles. The difference in categorical values between the two data sets was very small: only 1% change for several of the UKC% categories. The majority of vessels were within safety margins when lines were buffered as well: 81% of vessels registered over 20% draft values for UKC in the Port of Los Angeles. Tanker vessels had a larger percentage of vessels operating with higher risk UKC and UKC% values.

The tug and towing vessels showed the greatest difference in UKC% values from UKC and UKC% values calculated from lines and then from the buffered lines. The high-risk UKC% categories remained the same: 4% of both regular and buffered lines were in the 0-20% UKC% value range. The safest category, over 100% UKC%, decreased 6%, which the 20-50% and 50-100% UKC% categories slightly increased. This shift toward shallower values shows the trend followed by the other categories of lines.

The small changes in the percentages of positive track line and buffered polygon values indicate positional errors are unlikely to contribute significantly to AIS track analysis for these two case studies. The position of vessels would make a larger impact where the seafloor depths change rapidly, such as in Alaskan ports. The techniques used in this analysis may prove valuable to NOAA hydrographers, who need to categorize and prioritize very diverse ports throughout the entirety of the United States.

4.4 Minimum Depths from Track Lines and Bathymetry

Given the errors in the AIS track line attribute data indicated by negative UKC and UKC% values in both ports, another data set may be valuable for creating hydrographic survey priorities: minimum bathymetry values (Zcor) ships experience during normal operations. Combined with the trends of UKC% between ports, Zcor values may help provide additional information about relative operational risk by removing the human input component of the draft data from crowdsourced AIS data set.

4.4.1 Hampton Roads Zcor Values

Due to the quality control measures taken for the two ports, Hampton Roads and Los Angeles had no negative Zcor values. A small percentage of vessels carrying AIS transited over Zcor depths less than 5 meters – 5% of vessels in this test case. Table 13 shows the distribution of Zcor with all Hampton Roads vessel track lines. The following section breaks the Zcor values down by type of ship.

Table 13 Distribution of Zcor Values for Hampton Roads

Zcor Range	Track Lines		Buffered Track Line Polygons	
	Number of Vessels	Percentage of Vessels	Number of Vessels	Percentage of Vessels
0-5 meters	157	5.4%	157	5.4%
5-10 meters	1162	40.2%	1162	41.5%
10-15 meters	1490	51.5%	1490	50.3%
15-20 meters	77	2.7%	77	2.5%
20-25 meters	7	0.24%	7	0.24%

The buffered track line Zcor values were nearly identical to the original track line values. The percentages of vessels in each Zcor category were within one percentage point, showing that a 5-meter horizontal positional error also does not cause a large change or error in the results of the AIS analysis. No negative Zcor values were recorded, helping to confirm that the quality control measures filtered out vessels that might have crossed onto charted land. The bathymetric model did not indicate a body of land where water was charted and ships were transiting.

4.4.2.1 Hampton Roads Zcor Values by Vessel Type

When the Zcor values are separated by ship type, the distribution of numbers of vessels reflects the size and draft of the vessel. Table 14 shows no cargo vessels reported entering waters less than 5 meters, and only a small percentage of tankers enter waters that shallow.

Tanker and cargo vessels had similar distributions in Zcor depths, tending toward deeper water ranges, where tugs had a greater distribution of Zcor ranges.

Table 14 Distribution of Zcor Track Line Values for Main Vessel Types in Hampton Roads

Zcor Range	Percentage of Cargo Vessels	Percentage of Tanker Vessels	Percentage and Tug and Towing Vessels
0-5 meters	0	1.1	8.7
5-10 meters	40.5	41.5	34.4
10-15 meters	56.9	50.0	54.3
15-20 meters	2.6	7.4	2.5
20-25 meters	0	0	0.2

Similar to the results of the overall track lines, the buffered track Zcor values in Table 15 for each vessel type were also nearly identical to the original track line values. While the difference between the tug/towing track lines and buffered polygons were within one percentage point, the cargo and tanker percentages varied by almost 3 percentage points in some categories, which is still a very minimal difference. Most vessels, regardless of type, showed that vessels tend to stay in waters 10-15 meters deep. A large amount of vessels also transit in waters from 5 to 10 meters in depth.

Table 15 Distribution of Zcor Values for Buffered Track Line Polygons for Main Vessel Types in Hampton Roads

Zmin Range	Percentage of Cargo Vessels	Percentage of Tanker Vessels	Percentage and Tug and Towing Vessels
0-5 meters	0	1.0	8.7
5-10 meters	42.6	42.11	34.9
10-15 meters	54.8	47.37	54.2
15-20 meters	2.6	9.5	2.1
20-25 meters	0	0	0.2
25-30 meters	0	0	0

In both cases the tugs and towing vessels have the most occurrences of transiting in waters less than 5 meters deep. Due to the draft and size characteristics of these vessels, as well as their mission of safety conveying larger vessels through narrow channels and into port berthing, this nearly 9% of vessels operating in less than 5 meters is expected. Few vessels remained in water of 15 meters or deeper, indicating most vessels were headed toward piers and few vessels transited through the region via deep draft channel only.

4.4.3 Los Angeles Zcor Values

Vessels in the Port of Los Angeles showed a great distribution of vessels operating at different depths when transiting through the port. Table 16 shows the wide distribution of all vessels in Los Angeles during the case study period.

Table 16 Distribution of Zcor Values for Los Angeles Vessels

Zcor Range	Track Lines: Number of Vessels	Track Line: Percentage of Vessels	Buffered Polygons: Number of Vessels	Buffered Polygons: Percentage of Vessels
0-5 meters	2755	24.1%	3052	26.7%
5-10 meters	2319	20.3%	2380	20.8%
10-15 meters	3094	27.1%	2927	25.6%
15-20 meters	2741	24.0%	2526	22.1%
20-25 meters	373	3.3%	369	3.2%
25-30 meters	51	0.4%	44	0.4%
30-35 meters	1	< 0.1%	0	0%
>35 meters	1	< 0.1%	0	0%

The buffered track line values for Los Angeles, also shown in Table 16, were within three percentage points of the original track lines values for each Zcor range

4.4.3.1 Los Angeles Zcor Values by Vessel Type

The Zcor values associated with cargo, tanker, and tug/towing vessels show a similar distribution to the Hampton Roads track lines. Few tanker and cargo vessels reported transiting in waters less than 10 meters. The majority of cargo vessels (72.8%) reported Zcor values between 15 and 20 meters, and 18.5% had Zcor values from 10-15 meters deep. Tanker vessels had a wider distribution of Zcor values, but these vessels stayed between 10 and 25 meters during most transit. Tug/towing vessel traffic was also similar to Hampton Roads distributions –

track lines tended to be in shallower waters than the tankers and cargo vessels. Table 17 shows the distribution of the three vessel types in Los Angeles for track line Zcor values.

Table 17 Distribution of Zcor Values for Main Vessel Types in Los Angeles

Zcor Range	Percentage of Cargo Vessels	Percentage of Tanker Vessels	Percentage and Tug and Towing Vessels
0-5 meters	2.1	0.5	23.8
5-10 meters	3.5	0.5	30.6
10-15 meters	18.5	27.1	36.3
15-20 meters	72.8	43.7	7.1
20-25 meters	2.7	24.6	1.1
25-30 meters	0.3	3.6	< 0.1
30-35 meters	0	0	0

The buffered track line Zcor distribution values in Table 18 show similar values as Hampton Roads Zcor values: Zcor range distribution was within several percentage points to the original track line Zcor values. There is a slightly larger difference in the LA data: tracks and buffered track polygon values are within 5 percent vessels in the Port of Los Angeles.

Table 18 Distribution of Zcor Buffered Track Line Polygon Values for the Main Vessel Types in Los Angeles

Zcor Range	Percentage of Cargo Vessels	Percentage of Tanker Vessels	Percentage and Tug and Towing Vessels
0-5 meters	2.7	0.6	27.2
5-10 meters	3.8	0.8	32.4
10-15 meters	21.9	30.7	32.0
15-20 meters	68.8	39.7	5.7
20-25 meters	2.6	25.0	1.1
25-30 meters	0.2	3.2	< 0.1
30-35 meters	0	0	0

The distribution of Zcor values for vessel types in both Los Angeles and Hampton Roads is a function of the structure and bathymetry of the port as well as the vessel's dimensions and the nature of the commerce it conducts. Similarities in results are found between the two ports, but because the Zcor values are not percentages or scaled to vessel characteristics in any way the use will be limited, as is discussed in later sections.

4.5 Estimated UKC% and Zcor Comparison

Given the uncertainty for ship drafts as recorded in the AIS data, another approach to analyzing these data is to make assumptions regarding the drafts for all ships of a certain type. To demonstrate this assumption, tracks for the tanker and cargo vessel types were used. Tanker and cargo vessels are classified according to size, as described by the Average Freight Rate Assessment (AFRA) scale. Because of the depths of the deep draft channels of both ports, the Panamax ship classification, what was named for the Panama Canal, was chosen as the representative ship dimensions for tanker and cargo vessels. Panamax vessels are the largest size vessels that could transit through the Panama Canal prior to its most recent upgrade. These ships boast maximum dimensions of 965 ft (294.13 meters) length, 106 ft (32.31 meters) width, and 39.5 ft (12.04 meters) draft (Maritime Connector 2015). These ship dimensions are commonly used to describe commercial vessel traffic supported by ports and are therefore appropriate as an assumption for further interpretation of Zcor results.

Using track line data only and the assumption that cargo vessels operate at Panamax limits, the results from the Zcor analysis can be re-interpreted to assign relative risk. In Hampton Roads, nearly 98% of cargo traffic records Zcor values of less than 15 meters. In contrast, less than 25% of cargo vessels in Los Angeles transit through waters less than 15 meters deep. For tanker vessels, 91% of vessels in Hampton Roads and 28% of vessels in Los Angeles record Zcor values less than 15 meters. Of course, making this draft generalization is a stretch because it is unknown whether most or any of these vessels without AIS draft values or other ship dimensions meet the upper limits of the Panamax classification. It is more likely that the cargo and tanker vessels have a wide variety of lengths and drafts that compose the typical vessel traffic. In Los Angeles in particular, it is common for tanker vessels with dimensions larger than

Panamax limits to offload a small percentage of their cargo offshore to minimize their draft and enter the port safely during periods of heavy seas and large amplitude, long period swell. With this practice, known as lightering, additional vessels are required to carry the small percentage of cargo from near the coastline into the terminal (Port of Los Angeles 2008). Thus, it is concluded that this approach to estimating handling the unknown draft is unsuccessful.

4.6 Hampton Roads and Los Angeles Comparison

Comparing the positive UKC% values and percentages of vessel traffic, patterns of risk can be determined from the data. For the final analysis, low risk UKC% values were assigned to UKC values 20% of a vessel's draft and greater. Medium risk UKC% values were between 10 and 20% of a vessel's draft, and high risk were from 0 to 10% UKC

When comparing the two ports in Table 14 by strictly including all AIS vessel traffic, Hampton Roads percent of traffic that has high risk UKC% values is approximately 4 percentage points higher than Los Angeles, making it a higher survey priority. When the traffic is broken down into vessel types the results and differences are more pronounced. Cargo and tanker vessels assume much greater risk in Hampton Roads where ship channels and surrounding waters are not as deep as Los Angeles. The medium risk categories for all vessels have the same pattern, but differ on the vessel type level: Los Angeles has a higher percentage of medium risk tankers than Hampton Roads. The percentages of risk for tug and tow vessels are nearly the same in both ports, suggesting these vessels do not contribute significant risk to the ports and also may not as important as the larger vessels when studying hydrographic survey priorities. More ports will need to be added to the study to draw definite conclusions about the significance of this finding.

This comparison, shown in Table 19, demonstrates that ports display different levels of operational risk when looking solely at UKC and UKC% values. Based on the bathymetry and vessel traffic trends, differences can be determined in the number of vessels that are at high, medium, and low risk of running aground or having a draft-related incident in the areas they transit.

Table 19: Comparison of High, Medium, and Low Risk UKC% for Hampton Roads and Los Angeles

Vessel Type	Risk Level – UKC%	Hampton Roads		Los Angeles	
		Number of Vessels	Percentage of Vessels	Number of Vessels	Percentage of Vessels
All Vessels	Low (>20%)	1867	62.3%	9064	79.3%
	Medium (10-20%)	159	5.3%	322	2.8%
	High (0-10%)	**194**	**6.5%**	**261**	**2.3%**
Cargo	Low (>20%)	497	41.1%	2480	94.6%
	Medium (10-20%)	116	9.6%	76	2.9%
	High (0-10%)	**142**	**11.8%**	**25**	**1.0%**
Tanker	Low (>20%)	36	38.3%	625	15.3%
	Medium (10-20%)	5	5.3%	53	7.0%
	High (0-10%)	**14**	**14.9%**	**27**	**3.6%**
Tug/Tow	Low (>20%)	811	86.7%	4081	72.4%
	Medium (10-20%)	11	1.2%	99	1.8%
	High (0-10%)	**12**	**1.3%**	**95**	**1.7%**

4.7 Results Chapter Summary

UKC, UKC%, and Zcor values vary from port to port depending on the type of vessel traffic, the bathymetry, and the character of the port. After QC measures were applied, minimizing the amount of uncertainty associated with the bathymetry, track line, and buffered track line polygon data, a significant percentage of tracks and polygons with negative UKC and UKC% values resulted. This was due to drafts reported by AIS being larger than the Zcor values calculated for each line and polygon. The main sources of these negative results varied in each port. In Hampton Roads most vessels with negative UKC% values were cargo vessels, while Los Angeles had significant number of tug and tow vessels with negative results. Additional research is necessary to understand the nature of the draft errors in order to correct them during the QC process and decrease the number of erroneous negative UKC% results. Studies could include determining how many vessels entered draft values using incorrect units and lead to a proposal for a quality control method that might be applied to solve the problem. Further research and understanding into the uncertainty of AIS draft data may help hydrographers determine how UKC% values contribute to a future quantitative survey prioritization model.

In the Hampton Roads case study, 6% of vessels operated with high risk UKC% values. When the UKC% results were categorized by vessel type, 12% of cargo vessels, 15% of tankers, and 11% of tug and towing vessels had high risk UKC% values during their transits. These values are expected due to the nature of their operations and the gradual slope of the bathymetry in Hampton Roads. Most of the vessels entering this port operate within safe UKC% limits for their vessel drafts. These results were mirrored in the UKC% results for the Hampton Roads buffered track line polygons, which expanding the radius for Zcor data around the individual track lines to account for typical GPS uncertainty. There was a slight shift of UKC% percentages towards the higher risk categories because when transiting near sloping bathymetry or near the edge of a channel, shallower depths will be chosen to the Zcor values compared with the Zcor values directly beneath the vessel.

The Los Angeles case study had different results in comparison with Hampton Roads. Los Angeles had significantly fewer vessels transiting with high risk UKC% values; only 1% of all vessels left 0-10% UKC% during the transit. Less than 2% of cargo vessels and only 4% of tanker vessels operated with high risk UKC% results. Tug and towing vessels also had very low values: less than 2% had UKC% percentages between 0 and 10%. The distribution of UKC% for the Los Angeles buffered track line polygons also mirrored the values of the original tracks, indicating that there were no significant advantages to performing the analysis for polygons in areas with gently sloping bathymetry.

The additional track line analysis using Zcor values demonstrated that unless more information about the individual vessels, Zcor is a less valuable factor for determining survey priorities than UKC%. For it to be useful, Zcor data must be used in conjunction with large assumptions of ship dimensions, while AIS data does not need assumptions – only quality control and measures of uncertainty.

The analytical techniques used in this chapter demonstrate that using AIS data in conjunction with bathymetry can provide additional information beyond traffic density analyses. The final chapter will discuss the challenges with using AIS data, suggestions for using UKC% derived from AIS and bathymetry, and possible future research.

CHAPTER 5: DISCUSSION AND CONCLUSIONS

AIS data are gaining traction in the scientific community as a means of studying vessel traffic patterns, aiding in environmental studies, and assessing operational risks of vessels. Despite the growing us of AIS for research, studies that use vessel data manually entered and volunteered by ship personnel must be used carefully. The AIS data reporting parameters derived from GPS have been studied and found to have high levels of accuracy (Januszewski 2014), but the human element incorporates a new level of uncertainty that is difficult to quantify and predict based on the nature of errors and the dynamic nature of ships. The uncertainty in ship draft data makes it necessary for extensive data quality control and corrections to be completed in order for the data to be useful and as accurate as possible for assessing hydrographic survey priorities. Chapter 5 discusses the inherent errors and uncertainty of the UKC% results for the case studies, suggests how AIS draft data should be used for setting hydrographic survey priorities, and makes recommendations for future research into crowdsourced AIS data for hydrographic purposes.

5.1 Data Errors and Verification

The most significant problem with using AIS data for hydrographic studies is the human factor. In most of the cases the GPS data can be trusted, as confirmed by the small number of track lines that were eliminated during the quality control process when vessel tracks intersected the charted shoreline. The vessel attribute information is a large source of errors. The reported values such as ship’s draft and type of vessel are manually input by the crew of the vessel. In conversations held with professional mariners during the US Hydro 2015 conference, deck officers stated that these parameters were buried within menus in AIS units and were hardly ever changed during the standard operating procedures for their vessels when getting underway or entering port.

In addition to the infrequency of updating vessel information in the AIS unit, it is easy for a mariner to enter incorrect information into the fields necessary for this study: draft and vessel/cargo type. Metric units are the international standard for draft, length, and beam information in AIS, but it is evident from the data that these standards and guidelines are not always followed. In the dataset studied in these two cases, multiple vessels were removed from the dataset from having unreasonably high drafts. Incorrect vessel characteristics can also be entered in error by pushing incorrect buttons. NOAA’s LTJG Anthony Klemm explained that in many cases mariners enter their draft information using the incorrect units or attempt to enter a string of numbers that should include values after a decimal point. Since the AIS unit only allows whole number entries, these types of draft entries are much larger than reality (for example 1.2 meters would display as 12 meters). These errors are difficult to detect and correct because they are inconsistent and vary based on the individual ship.

Draft data and vessel characteristics may be communicated by the ships’ masters and mates directly to the harbormaster and pilots, the authorities responsible for ship routing and safe navigation into and out of ports. With the direct line of communication, there is less emphasis on the AIS draft information and ship crews may not see the importance in keeping these values in their AIS unites accurate or up to date as they change.

Although the vessel type should not need to be changed during a cruise or while getting underway, the vessel type is also information that is directly input by ship crew, which allows the possibility for additional errors in AIS datasets. If the ship crew is not diligent or detail-oriented enough to input the correct category of vessel, they might not input the correct length or draft of the vessel either.

The inconsistent nature of errors associated with the human component of crowdsourced AIS data makes a great deal of human intervention necessary to create a quality dataset; even a complex set of automated controls may not catch all possible errors. Verification is needed for the results calculated from the AIS and bathymetry datasets, but at present additional data and resources are not available to perform any quantitative validation. Not having any quality measures built into the AIS, it is impossible to verify the actual quality of the AIS attributes.

Therefore, instead of employing a quantitative verification process for this project, it was necessary to rely on expert review to confirm the quality of the results, as suggested by Fonte (2015). The data ingested and produced, and the results generated were reviewed by two subject matter experts within the NOAA OCS: Lucy Hick and LTJG Anthony Klemm. In personal conversations, Ms. Hick and LTJG Klemm provided guidance during the planning, execution and review stages of the project. They confirmed that quantitative verification is not possible at this time and further study is needed (Hick 2015). LTJG Klemm indicated that the negative UKC and UKC% values are an inherent problem that needs to be addressed in the AIS draft data when NOAA moves forward to use AIS attribute data in operations and in studies such as those which might incorporate quantitative methods into the hydrographic survey prioritization process in the future (Klemm 2015). In Elwood, Sui, and Goodchild’s (2013) crowdsourcing quality assurance framework, the geographic approach the one used here, as lines were removed and decisions were made based on geographic locations and spatial relationships.

5.2 Data Error Solutions

One possible solution for the need for strict, in-depth quality control is to determine if the harbormasters and USCG keep a detailed database of ship characteristics and add the correct draft values to the dataset using the ship MMSI numbers. This solution may be possible for larger ports where VTS exists and where pilots are mandatory for larger vessels. Data on smaller vessels may not exist. The original MMSI numbers would also be necessary for this solution: to protect the individual identities of ships, the USCG scrambled MMSI numbers used in this study.

In addition to acquiring supplementary ship characteristics from sources other than AIS, additional education for mariners may help improve the quality of vessel attribute information in AIS. In the past AIS data has not been regularly used outside of the maritime industry. As it becomes more popular and useful for the scientific community and other industries, more entities will examine the data and create a greater need for accurate information. If international bodies such as the International Maritime Organization set stricter educational requirements for mariners, data quality may begin to improve over time.

Removing the human element to crowdsourced AIS data is also a possible long-term solution. Adding sensors into the data stream to report the instantaneous draft would remove errors from the system. Such an initiative that would need discussion between electronics companies and the USCG, which helps set standards for AIS transmissions. By adding an additional field to the AIS message, the automatically collected draft could be added into the automatic transmission by reading in a simple NMEA string, a standard electronics message format from the National Marine Electronics Association (NMEA) from a ship’s calibrated fathometer and report the actual draft or even the under keel clearance at the time of the message

(National Marine Electronics Association 2015). This would eliminate a significant contributor to the calculated UKC% uncertainty.

5.3 Using AIS-derived UKC for Survey Prioritization

Despite the inherent errors and uncertainty associated with the vessel characteristics of AIS data, the dataset can still provide valuable information for hydrographers and can help quantify risk for setting hydrographic survey priorities. Using the number and percentages of vessels in negative, high, and moderate risk UKC% categories can help assign relative risk levels. As seen in this comparison between Hampton Roads and Los Angeles, two prominent deep draft vessel ports with VTS control, measurable differences in the numbers and percentages of risk level assumed by vessels can identified using the combination of traffic patterns and bathymetry. This comparison can be difficult to quantify using the previous prioritization approach that depends upon factors such the age of bathymetry and ship density. By including the ship draft information, even including errors the analysis can provide valuable information that can be included in the hydrographic survey planning decision process. As models are developed by NOAA to set survey priorities, the percentages of risk categories may be included to set the total operational risk of the port. Although UKC% is an important factor, because of the errors and uncertainty in data, it should be weighted accordingly, taking the uncertainty and importance of other factors into consideration.

Similar to UKC%, Zcor data can be useful if it is properly broken down into specific categories of draft and vessel class. Understanding that the vessels conducting commerce and conveying passengers are at the most risk, these categories should be the main focus. The relative comparison of Zcor values of vessel categories may be included in the final hydrographic survey decisional model, but will be more difficult to use because draft must also be factored in to complete more extensive quality control. UKC% is a better variable to add into a model because the value and analysis already factors in the draft of vessels and is scaled to assess risk with the same criteria. Zcor values help establish trends, but when using this approach the track lines still require several extra steps of quality control and vessel/draft classification as compared with the UKC% analysis.

Further study is needed to bring to light the human errors found in the AIS vessel attribute data and to determine how the UKC% and Zcor data can be incorporated into a quantitative model for NOAA OCS. Although studies have been conducted examining the error rates in crowdsourced AIS data variables, additional information is needed about the specific nature of the errors in AIS draft data to gain insight into how these errors may be identified and either filtered or corrected. One possible study could leverage the aid of harbormasters and USCG VTS in specific ports to query ships entering their ports for their MMSI number and actual draft at the time of port entry. Over a period of several months the reported draft data could be compared with the AIS draft and additional information about draft error may be determined. It is likely that draft errors vary from port-to-port and season-to-season based on the types of vessels that transit in and out of port and where the vessels are traveling. If the errors are found to be consistent, a correction factor may be discovered and use to correct the AIS draft data when it is used in a new survey prioritization model.

As crowdsourced AIS data are used more regularly as a source of information for shore-based operations and studies, as opposed to collision avoidance at sea, the international maritime community is likely to respond to the uncertainties inherent in the volunteered data and help increase the accuracy and credibility through awareness and training. Draft data from AIS messages is valuable to the hydrographic survey community despite the errors that exist.

Relative trends in UKC% can be used to assess relative operational risk from under keel clearance and help quantify the need for surveys to be completed along the United States coast by NOAA. Further study and research into the specific errors and the incorporation of enhanced quality control measures to the data will create an increasingly robust solution and confidence in the data acquired from AIS messages.

REFERENCES

Aitamurto, Tanja, Aija Leiponen, and Richard Tee. 2011. "The Promise of Idea Crowdsourcing - Benefits, Contexts, Limitations." Crowdsouricng.org. Last updated July 9 2011. Accessed August 10, 2013. http://www.crowdsourcing.org/document/the-promise-of-idea-crowdsourcing--benefits-contexts-limitations/5218.

Cooper, Paul and John Hersey. 2013. “Crowdsourcing the Ocean Floor: How Mariners Can Gather Valuable Information for Better Decision-Making.” gCaptain.com. Last modified March 1, 2013. Accessed August 10, 2014. http://gcaptain.com/crowdsourcing-ocean-floor-mariners/.

Elwood, Sarah, Michael Goodchild, and Daniel Sui. 2013. “Prospects for VGI Research and the Fourth Paradigm.” In *Crowdsourcing Geographic Information: Volunteered Geographic Information (VGI) in Theory and Practice*, edited by Sara Elwood, Michael Goodchild, and Daniel Sui, 361-375. Springer Netherlands.

Fonte, Cidália C., Lucy Bastin, Linda See, Giles Foody, and Favio Lupia. 2015. “Usability of VGI for validation of land cover maps.” *International Journal of Geographical Information Science*: 1-23.

Goetz, Marcus and Alexander Zipf. 2013. “The Evolution of Geo-Crowdsourcing: Bringing Volunteered Geographic Information to the Third Dimension.” In *Volunteered Geographic Information, Public Participation, and Crowdsourced Production of Geographic Knowledge.* Edited by Daniel Sui, Sarah Elwood, and Michael Goodchild, 139-159. Berlin: Springer.

Goodchild, Michael F. 2007. "Citizens as Sensors: The World of Volunteered Geography." *GeoJournal* 69 (4): 211-221. doi:10.1007/s10708-007-9111-y.

Harati-Mokhtari, Abbas, Alan Wall, Philip Brooks, and Jin Wang. 2007. "Automatic Identification System (AIS): Data Reliability and Human Error Implications." *The Journal of Navigation* 60 (3): 373-389. doi:10.1017/S0373463307004298.

Hick, Lucy. 2015. Interviews by author. Silver Spring, Maryland. March 2015.

International Maritime Organization. 2014. "AIS Transponders." International Maritime Organization. Accessed November 2, 2014. http://www.imo.org/OurWork/Safety/Navigation/Pages/AIS.aspx.

Jalkanen, Jukka-Pekka, Joseph A. Brink, Lewis Kalli, H. Pettersson, Jussi Kukkonen, and Tapani Stipa. 2009. "A Modeling and System for the Exhaust Emissions of Marine Traffic and its Application in the Baltic Sea Area." Atmospheric Chemistry and Physics 9 (23): 9209-9223. doi:10.5194/acp-9-9209-2009.

Januszewski, Jacek. 2014. “Nominal and Real Accuracy of the GPS Position Indicated by Different Maritime Receivers in Different Modes.” TransNav – The International Journal on Marine Navigation and Safety of Sea Transportation 8 (1): 11-19. doi: 10.12716/1001.08.01.01

Klemm, Anthony. 2015. Interviews by author. Silver Spring, Maryland. February and March 2015.

MarineCadastre.gov, 2015. “Data Registry.” Marine Cadastre. Accessed 14 March 2015, http://marinecadastre.gov/data/.

Marine Connector, 2015. “Ship Sizes.” Marine Connector. Accessed 11 April 2015, http://maritime-connector.com/wiki/ship-sizes/.

Marine Exchange of Southern California, 2014. “Harbor Safety Plan.” Marine Exchange of Southern California. Accessed 21 March 2015, http://www.mxsocal.org/HARBOR-SAFETY-AND-SECURITY/HARBOR-SAFETY/Harbor-Safety-Plan.aspx.

Military.com, 2015. “Naval Station Norfolk.” Military.com. Accessed 15 March 2015, http://www.military.com/base-guide/naval-station-norfolk.

National Data Buoy Center. 2009. "The WMO Voluntary Observing Ships (VOS) Scheme." National Oceanic and Atmospheric Administration. Last updated January 28, 2009. Accessed September 30, 2014. http://www.vos.noaa.gov/vos_scheme.shtml.

National Geophysical Data Center, 2015a. “Bathymetry & Digital Elevation Models.” National Oceanic and Atmospheric Administration. Accessed 8 March 2015, http://maps.ngdc.noaa.gov/viewers/bathymetry/?layers=nos_hydro.

National Geophysical Data Center, 2015b. “NOS Hydrographic Survey Data.” National Oceanic and Atmospheric Administration. Accessed 15 March 2015, http://maps.ngdc.noaa.gov/viewers/nos_hydro/.

National Marine Electronics Association, 2015. “Standards (NMEA 2000, 0183).” National Marine Electronics Association. Accessed 26 April 2015, http://www.nmea.org/content/nmea_standards/nmea_standards.asp.

National Oceanic and Atmospheric Administration. 1967. “Descriptive Report: H08878.” Accessed May 8 2015. http://surveys.ngdc.noaa.gov/mgg/NOS/coast/H08001-H10000/H08878/DR/H08878.pdf.

National Oceanic and Atmospheric Administration. 2013a. "Charting and Geodesy." National Oceanic and Atmospheric Administration. Accessed August 10, 2013, http://www.noaa.gov/charts.html.

National Oceanic and Atmospheric Administration. 2013b. "Navigation Response Teams." National Oceanic and Atmospheric Administration. Accessed September 10, 2013, http://www.nauticalcharts.noaa.gov/nsd/nrt.htm.

National Oceanic and Atmospheric Administration, 2014. “Descriptive Report: H12617.” Accessed May 8, 2015. http://surveys.ngdc.noaa.gov/mgg/NOS/coast/H12001-H14000/H12617/DR/H12617_DR.pdf.

National Ocean and Atmospheric Administration, 2015a. “Los Angeles, CA – Station ID: 9410660.” National Oceanic and Atmospheric Administration. Accessed 8 March 2015, http://tidesandcurrents.noaa.gov/stationhome.html?id=9410660.

National Oceanic and Atmospheric Administration, 2015b. “NOAA Continually Updated Shoreline Product (CUSP).” National Oceanic and Atmospheric Administration. Accessed 9 March 2015, http://shoreline.noaa.gov/data/datasheets/cusp.html.

National Oceanic and Atmospheric Administration, 2015c. “Sewells Point, VA – Station ID: 8638610.” National Oceanic and Atmospheric Administration. Accessed 8 March 2015, http://tidesandcurrents.noaa.gov/stationhome.html?id=8638610.

National Oceanic and Atmospheric Administration, 2015d. “Water Levels – Station Selection.” National Oceanic and Atmospheric Administration. Accessed 22 February 2015, http://tidesandcurrents.noaa.gov/stations.html?type=Water+Levels.

NOAA Office of Coast Survey. 2012. *NOAA Hydrographic Survey Priorities, 2012 Edition*. National Oceanic and Atmospheric Administration. Accessed November 20, 2014. http://www.nauticalcharts.noaa.gov/hsd/NHSP.htm.

NOAA Office of Coast Survey. 2013b. "NOAA Survey Platforms." National Oceanic and Atmospheric Administration. Accessed August 10, 2013. http://www.nauticalcharts.noaa.gov/hsd/NHSP.htm.

NOAA Office of Coast Survey. 2013c. "United States Coast Guard." National Oceanic and Atmospheric Administration. Accessed August 10, 2013. http://www.nauticalcharts.noaa.gov/staff/uscg.html.

NOAA Office of Coast Survey. 2013a. *Arctic Nautical Charting Plan, February 15, 2013.* National Oceanic and Atmospheric Administration. Accessed 28 November 2014. http://www.nauticalcharts.noaa.gov/mcd/docs/Arctic_Nautical_Charting_Plan.pdf.

NOAA Office of Coast Survey. 2014. *Field Procedures Manual, 2014th ed.* National Oceanic and Atmospheric Administration. Accessed November 27, 2014. http://www.nauticalcharts.noaa.gov/hsd/fpm/2014_FPM_Final.pdf.

NOAA Office of Coast Survey. 2015. “Paper Charts (RNC and PDF).” National Oceanic and Atmospheric Administration. Accessed May 8, 2015.

Qian, Yao Yu, and Lin ChangChuan. 2011. "A Preliminary Scheme of the Online Monitoring System for the Ship Discharging Pollution at Harbor Based on AIS Information." *Procedia Engineering* 15: 1436-1440.

Port of Los Angeles, 2008. “Pacific L.A. Marine Terminal LLC Crude Oil Terminal Draft SEIS/SEIR.” Accessed 11 April 2015, http://www.portoflosangeles.org/EIR/PacificLAMarine/SEIR/1_%20Introduction.pdf

Port of Los Angeles, 2015a. “Facts and Figures.” Port of Los Angeles. Accessed 9 March 2015, http://www.portoflosangeles.org/about/facts.asp.

Port of Los Angeles, 2015b. “Timeline of Historic Events.” Port of Los Angeles. Accessed 8 March, 2015, http://www.portoflosangeles.org/history/timeline.asp.

Port of Virginia, 2015. “About.” Port of Virginia. Accessed 19 March 2015, http://www.portofvirginia.com/about/.

Sampath, Premalatha and David Parry. 2013. "Trajectory Analysis using Automatic Identification Systems in New Zealand Waters." *International Journal of Computer Information Technology* 2 (1): 132.

Schwehr, Kurt and P. A. McGillivary. 2007. "Marine Ship Automatic Identification System (AIS) for Enhanced Coastal Security Capabilities: An Oil Spill Tracking Application." IEEE, doi:10.1109/OCEANS.2007.4449285.

United States Coast Guard, 2013. “Welcome to Station Los Angeles Long Beach.” United States Coast Guard. Accessed 9 March 2015, http://www.uscg.mil/d11/staLALB/.

United States Coast Guard. 2014a. "Automatic Identification System Overview." United States Coast Guard Navigation Center. Last updated July 30, 2014. Accessed September 30, 2014. http://navcen.uscg.gov/?pageName=AISmain.

United States Coast Guard. 2014b. "Nationwide Automatic Identification System." United States Coast Guard Navigation Center. Last updated July 30, 2014. Accessed November 2, 2014. http://navcen.uscg.gov/?pageName=NAISmain.

United States Coast Guard, 2015. “Sector Hampton Roads.” United States Coast Guard. Accessed 15 March 2015, http://www.uscg.mil/d5/sectHamptonRoads/.

Vesselfinder.com, 2015. “Vessels Database.” Vessel Finder. Accessed 18 April 2015, http://www.vesselfinder.com/vessels.

Ward, Robert and Gilles Bessero. 2013. "Status Report on Hydrography and Mapping of the World's Seas, Oceans, and Coastal Waters." *United Nations Initiative on Global Geospatial Information Management*. Accessed August 10, 2013. http://ggim.un.org/docs/meetings/3rd%20UNCE/E-C.20-2013-10_IHO%20Land%20and%20Marine%20background%20paper.pdf.

Wood, Heather, 2012. "The Port of Virginia Infrastructure Update – Norfolk." Accessed 19 March 2015, http://www.norfolk.gov/DocumentCenter/View/1781.

World Meteorological Organization. 2014. "Global Observing System (GOS)." World Meteorological Organization. Accessed September 30, 2014. http://www.wmo.int/pages/prog/www/OSY/GOS.html.