

Preliminary Work on Use of Automatic Identification System (AIS) Data for Shoaling Detection

David L. Iverson
Ames Research Center, Moffett Field, California

Lilly Spirkovska
Ames Research Center, Moffett Field, California

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI Program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI Program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Report Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.
- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.

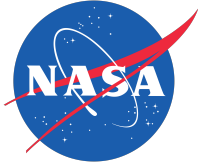
- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.
- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.
- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include creating custom thesauri, building customized databases, and organizing and publishing research results.

For more information about the NASA STI Program, see the following:

- Access the NASA STI program home page at <http://www.sti.nasa.gov>
- E-mail your question to help@sti.nasa.gov
- Fax your question to the NASA STI Information Desk at 443-757-5803
- Phone the NASA STI Information Desk at 443-757-5802
- Write to:
STI Information Desk
NASA Center for AeroSpace Information
7115 Standard Drive
Hanover, MD 21076-1320

NASA/TM-20260006244



Preliminary Work on Use of Automatic Identification System (AIS) Data for Shoaling Detection

David L. Iverson
Ames Research Center, Moffett Field, California

Lilly Spirkovska
Ames Research Center, Moffett Field, California

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035-1000

August 2026

Acknowledgments

We would like to thank our colleagues at the U.S. Army Engineering Research and Development Center (ERDC) for their generous technical and financial support of this study. We also thank our reviewers for their thoughtful comments and constructive feedback.

The use of trademarks or names of manufacturers in this report is for accurate reporting and does not constitute an official endorsement, either expressed or implied, of such products or manufacturers by the National Aeronautics and Space Administration.

Available from:

NASA Center for AeroSpace Information
7115 Standard Drive
Hanover, MD 21076-1320

National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312

Available electronically at <http://www.sti.nasa.gov>

Preliminary Work on Use of Automatic Identification System (AIS) Data for Shoaling Detection

David L. Iverson and Lilly Spirkovska
National Aeronautics and Space Administration
Ames Research Center
Moffett Field, California 94035-1000

Summary

This report describes a preliminary study undertaken in collaboration with the U.S. Army Corp of Engineers (USACE) to evaluate the utility of monitoring vessel traffic patterns on the Mississippi River using Automatic Identification System (AIS) tracking data with the goal of providing forewarning of growing shoals in the navigation channel. Activities discussed include mapping AIS position coordinates to a more favorable data space, discovering and extracting river dredging events from AIS data, and visualizing river traffic in the months leading up to dredging events. The brief study period allowed for traffic data evaluation at two sites on the river; further study could include development of metrics and user interface concepts. Initial evaluations showed promise for providing advanced notice of shoals or other traffic disruptions and are deemed worthy of follow-up in order to develop a robust and useful system to assist with river management. The intent of this report is to share details of these initial efforts with the aim of supplying a starting point and providing guidance for researchers revisiting and expanding on this potentially valuable effort. We provide several suggestions for future research. A prerequisite for any of these approaches is access to sufficient historical shoaling and dredging data that can be used as ground truth for model development and validation. We share a data-driven method we developed to infer dredging locations, likely spurred by shoaling, that enabled preliminary progress. To progress to a highly confident predictive system, it is critical to obtain a robust historical database of actual shoaling events. This will be vital to train and test any future model which aims to use indirect observations - such as AIS navigation paths, satellite imagery, etc. - in order to predict shoals at their formative stages.

Nomenclature

Acronyms

AIS	Automatic Identification System
DT-MIL	Deep Temporal Multiple Instance Learning
ERDC	Engineering Research and Development Center
IMS	Inductive Monitoring System
GIS	Geographic Information System
HDBSCAN	Hierarchical Density-Based Spatial Clustering of Applications with Noise
IENC	Inland Electronic Navigational Charts
KML	Keyhole Markup Language
LOMA	Lock Operations Management Application
MMSI	Maritime Mobile Service Identity
MPLD	Melvin Price Locks and Dam
USACE	U.S. Army Corp of Engineers
USGS	U.S. Geological Survey

1 Introduction

Hundreds of millions of tons of freight with an estimated value near \$100 billion are transported on the Mississippi River each year. (Ref. 1) Each barge on the river carries the equivalent of 16 rail cars or 70 semi-tractor/trailer loads, or more, of cargo while using less fuel and emitting far fewer pollutants than the alternatives. (Ref. 2) Given the importance of this waterway to U.S. commerce it is critical to minimize impediments to navigation such as naturally forming shoals that can cause obstructions throughout the course of the river. The U.S. Army Corp of Engineers (USACE) is charged with maintaining river navigation and spends hundreds of millions of dollars each year dredging the Mississippi to keep the traffic flowing smoothly. (Ref. 3, 4) Managing the dredging operations is a challenging task which can be ameliorated with advance information on locations of shoal formation activity. This information facilitates the dispatch of dredging resources in a timely and efficient manner. The supposition of this study was that savvy captains sailing the Mississippi River adapt their courses to avoid growing shoals. It has been observed that river captains are able to *read the river* and adapt their course to meet conditions. For instance, a study by Tabbert, et al. (Ref. 5) used vessel position data to demonstrate how captains adjust their approach to the Melvin Price Locks and Dam (MPLD) on the upper Mississippi River in Illinois depending on the flow rate and stage of the river, as shown in Figure 1. Our hypothesis was that vessel position data could be analyzed with data mining and anomaly detection techniques to discover changes in river traffic patterns in shoaling areas prior to formation of a disruptive shoal, thereby allowing USACE to dredge the shoal before it becomes a significant hazard to navigation. Even if there is no shoaling involved, unusual course deviations could be indicative of some other condition affecting the river that could also be useful information for river management.

We describe preliminary work toward detecting locations of shoal formation. The intent of this precursory work is to investigate possible methods and document both unsuccessful and promising avenues for future efforts. In Section 2, we describe the data sources, initial processing, and approaches for facilitating vessel course deviation detection. In Section 3, we describe efforts toward applying data driven anomaly detection and describe two visualization techniques that show promise for indicating shoal existence, and discuss possible metrics that could be derived to help alert USACE river managers of potential navigation issues. Section 4 summarizes the work and proposes future work.

2 Data Preparation

In this section, we discuss the steps to prepare the data into a form suitable for shoal formation analysis. We begin by describing the data source, Lock Operations Management Application (LOMA) Automatic Identification System (AIS) data (Section 2.1). We then provide an approach to map vessel position reports to facilitate recognition of navigation deviations in Section 2.2. This is required to, in effect, *straighten* the many twists and turns of the Mississippi River into a reference frame that allows insight into navigation deviations. Section 2.3 discusses an approach to infer where dredging may have occurred because that information was not directly available from sources we had access to but was necessary as *ground truth* for training data mining algorithms and for visualizing the precursor situation. Finally, in Section 2.4, we detail the specific data used for this study.

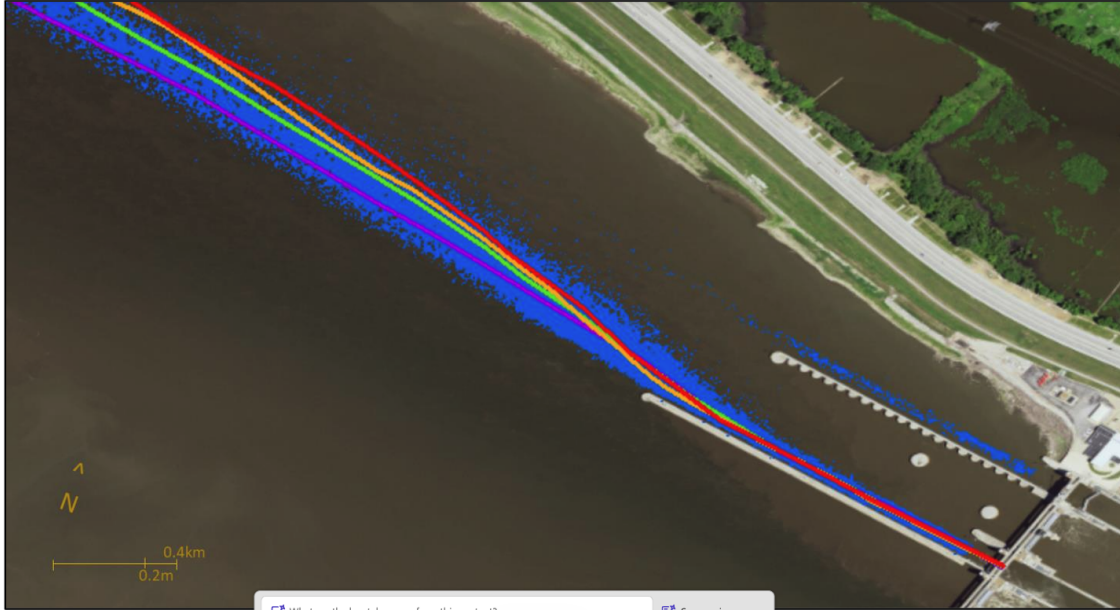


Figure 1.—The different colors depict the clustered paths of tow vessels entering the MPLD under different river discharge rate bands: purple: 25,000 cubic feet per second (cfs) – 75,000 cfs; green: 175,000 cfs – 225,000 cfs; orange: 225,000 cfs – 275,000 cfs; red: 275,000 cfs – 325,000 cfs. Blue points show all position reports (not clustered) for downbound vessels during the study period.(From (Ref. 5).)

2.1 Automatic Identification System (AIS) Data

AIS tracking is a maritime, radio-based broadcast system that automatically shares a vessel's identity, position, speed, and course in real time. Each large vessel on the Mississippi River carries an AIS transponder that broadcasts regular updates. Vessels equipped with Class A transponders automatically transmit information every 2 to 10 seconds while underway, depending on the vessel's speed and course, and every 6 minutes while anchored. (Ref. 6) A web search for *AIS data providers* will return several sources for historical AIS data. Many of these companies provide paid services including data and analysis tools. There is a no cost AIS data option at Marine Cadastre (<https://marinecadastre.gov>) as well. One disadvantage of many (perhaps all) of these services is they subsample the data to a point where important position details are lost and must be interpolated, a difficult task to do accurately given the river's many twists and turns.

In our initial studies, we used this lower rate subsampled data provided by Marine Cadastre and attempted to interpolate intermediate positions along the vessel tracks. A simple linear interpolation between position reports was insufficient. In more extreme cases a linear interpolation between two consecutive position reports that covered the vessel navigating a bend in the river would send the vessel straight through the point of land defining the bend. To overcome this, we implemented an interpolation that paralleled the *sailing line* (described below). This was an improvement, but was still inadequate.

Ultimately, we instead used high-rate data collected for the USACE LOMA. The LOMA project has deployed AIS receivers throughout the course of the Mississippi River to collect and distribute high-rate AIS data from vessels within receiver range to improve situational awareness for lock operators and nearby vessels. LOMA also maintains archives of this data for future

studies. These archives proved to be an invaluable data source for this study since they provide historical AIS data with sufficient resolution to accurately determine the detailed course of each vessel on the river without the need for interpolation. AIS data provides a wealth of information about vessels and their movements, not all of which were utilized for this study. Relevant AIS data that we considered includes:

1. Vessel Maritime Mobile Service Identity (MMSI): a unique nine-digit identification number
2. Navigation status: e.g., "at anchor", "under way using engine(s)", "not under command", etc.
3. Position in Latitude and Longitude (Lat-Long)
4. Rate of turn: right or left, from 0 to 720 degrees per minute
5. Speed over ground
6. Course over ground
7. True heading
8. Vessel name
9. Type of ship
10. Dimensions of ship
11. Draft of ship
12. Destination
13. Timestamp of when the data record was broadcast

The preliminary analyses relied primarily on position and timestamp reports, but the addition of other fields could enhance the overall results.

We used archived data from thirty-seven LOMA AIS receiver stations (locations shown in Figure 2). The data arrived in a raw compact packet bit field format; we used the Python module `pyais` to decode and extract relevant attributes. Specifically, we extracted and merged information from vessel description messages and vessel position reports to create augmented position reports that contain both vessel static information and dynamic position information.

The LOMA data contains several message types, transmitted at different rates, each with a subset of vessel information. These include static vessel information messages transmitted at about 6 minute intervals and position reports transmitted every 2 to 30 seconds for vessels in motion. Vessel description messages (message type 5) provide a timestamp, vessel's unique identifier (MMSI), callsign, ship name, ship type, dimensions, and draft. Position report messages (type 1 or 3, treated as one type since their interpretation is the same) provide timestamp, MMSI, latitude, longitude, speed over ground, course over ground, true heading, rate of turn, and navigation status. After decoding the messages using the `pyais` Python module we merged the two message types based on the shared MMSI into a combined record containing both vessel description and its timestamped position.

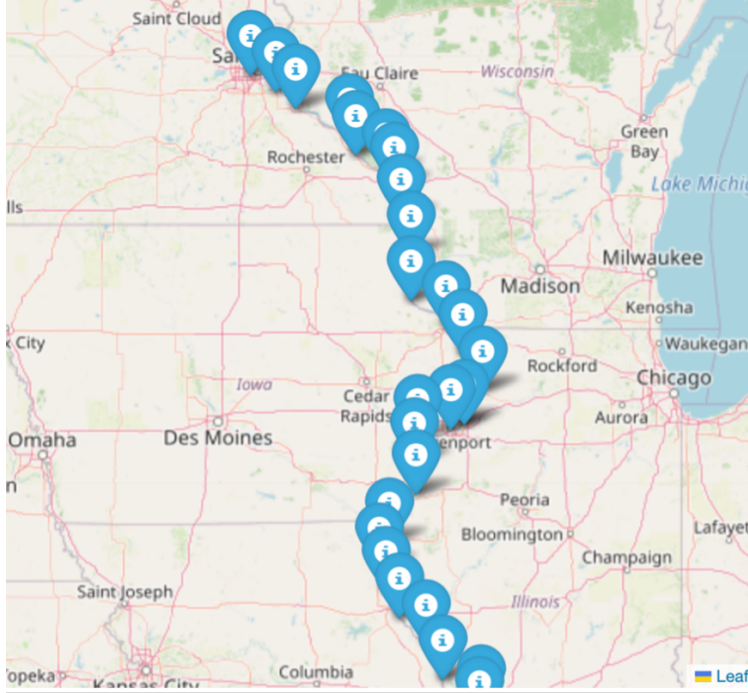


Figure 2.—LOMA AIS stations along the Mississippi River.

Because we were interested in dredging operations, we extracted the augmented position reports for ships of type dredge (AIS type 33) along with those of towing vessels (types 31 and 32) since they were the target of our course deviation analysis. Because the type code for dredge ships (33) is shared with underwater operations ships, we checked if the vessel name contained the word *DREDGE* or if its MMSI is on a list of MMSIs for known dredgers.

A final filtering step was to retain augmented position reports only for vessels operating on the Mississippi River since messages can reach an AIS receiver from hundreds of miles away under certain atmospheric conditions. We filtered out outliers from the initial AIS dataset in two steps. First, we filtered positions using a broad bounding box that contains the full east-west, north-south extent of the river. This eliminates spurious positions well away from the river (e.g., in the Pacific Ocean). We then narrowed the filter to eliminate positions that are more than 2 miles from the river.*

2.2 River Segmentation

The course of the Mississippi River is filled with twists and turns which can make it difficult to determine if deviations in a series of vessel latitude-longitude (lat-long) position reports are simply due to following the winding course of the river or avoiding navigational hazards. We therefore devised a method to map position reports, extracted as described in Section 2, to a new reference frame that effectively “straightens” the river to aid in characterizing typical traffic patterns and detecting diversions from those patterns. This involves referencing vessel positions as offsets to the left or right of a typical downstream course through the river.

One approach to defining this typical course for a given section, or *reach*, of the river is to

*Mississippi River coordinates from the headwaters to the Gulf of Mexico were obtained from <https://www.firstchurchofthemasochist.com/2023/08/river-coordinates-distances-mississippi.html>.

cluster traffic position reports through the reach and establish a navigation line through the centroids of the densest clusters. This was fairly effective, but often resulted in a somewhat jagged, unrealistic course line. A more effective approach was to use a navigational construct called the *sailing line*, as shown in Figure 3. This is a recommended course through the river designed to avoid shallow areas, obstructions, etc. In practice, most vessels do not follow this sailing line exactly, but it provides a reasonably continuous reference line through the river. One source for the coordinates of this sailing line is the USACE Inland Electronic Navigation Chart website (<https://ienccloud.us/>). In particular, the Keyhole Markup Language (KML) versions of these charts (https://ienccloud.us/ienc_kml.html) contain a “Recommended Track Line” block with the Lat-Long coordinates of the sailing line for the section of the river covered by that chart. The downloads from this website are compressed KMZ files, which can be extracted by reading them into Google Earth, searching for and selecting “RECOMMENDED_TRACK_LINE” from the Map Contents column, then selecting “Select Folder” and “Copy Features.” This places a copy of the sailing line coordinates in the clipboard which can then be saved to a file. Cautionary note: Sometimes coordinates will be out of order or loop back to the beginning of the line. It is important to plot the coordinates to be sure they form a contiguous line and edit as necessary. There may be other, easier to use sources for sailing line coordinates, but this was the best option we found during this study.

The sailing line provides a nice continuous route through the river reach, but we also needed a convenient reference for points along the sailing line. A commonly used reference is *river miles* which simply give the number of miles sailed upriver from a given point. In the case of the Mississippi there are river miles defined for the Lower Mississippi (starting South of New Orleans, LA, with mile zero at *Head of Passes* and proceeding North) and another set defined for the Upper Mississippi (starting with mile zero at the confluence of the Ohio River near Cairo, IL, and proceeding North). There are several online sources to obtain the lat-long coordinates of these mile markers. For instance, a search for “USACE IENC Distance Marks” results in a link to a Geographic Information System (GIS) map where you can extract those coordinates.

River miles are useful when referencing specific reaches of the river since they are nearly universally used by researchers, river managers, and sailors. However, we found that reference system too coarse for our purposes, so we devised a system to subdivide the sailing line with much finer resolution. It simply divides the sailing line into smaller segments using perpendicular transects; Figure 4 shows transects every 60 feet but other sizes might be as or more effective. Extending the transecting lines to the riverbank on either side outlines regions in the river, which we identified with consecutive numbers increasing downstream. Note that some of the transects may intersect mid-river due to tight turns in the river course. A program to map lat-long positions to segment numbers needs to take this into account as a given position could fall within multiple segments. In our case to resolve the ambiguity we selected the first segment encountered given the vessel’s direction of travel (lowest matching segment number for downstream travel, highest for upstream travel). Furthermore, to track lateral position a second parameter provided the vessel’s perpendicular distance from the sailing line in feet with a positive number on the right descending bank side (the riverbank to the right as a vessel proceeds downstream) and negative values on the left descending bank side (see Figure 5). In effect, this lat-long to segment-offset mapping straightens out the course of the river.

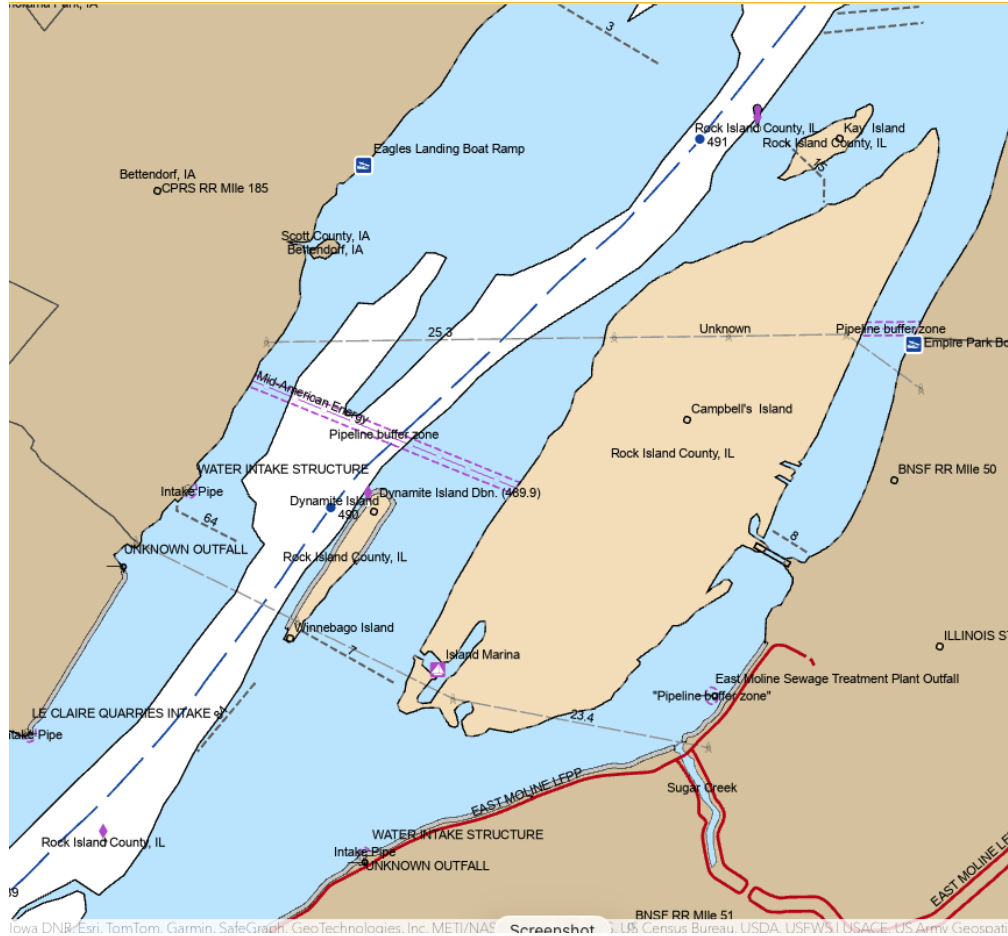


Figure 3.—Sailing line (dark blue dashed line, lower left to upper right) depicted on a Mississippi River navigation chart (source: USACE Inland Electronic Navigation Charts (Ref. 7))

2.3 Dredge Activity Identification

The dataset in the files resulting from the steps described in Section 2 contained augmented position reports for dredgers operating along the Mississippi River. The next step in the process was to identify areas where the dredgers were likely clearing shoals to maintain the navigation channel. We adopted this data driven historic shoal identification approach due to lack of access to records indicating when and where past shoaling had occurred.

Because dredging operations can continue for multiple days, we grouped positions for consecutive dates into one operation. For example, if a dredger operated on August 3 at 3pm to

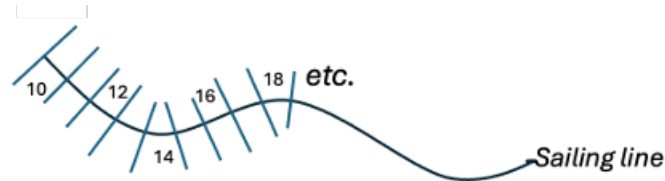


Figure 4.—Sailing line with numbered transect delineated segments to define indexed river regions.

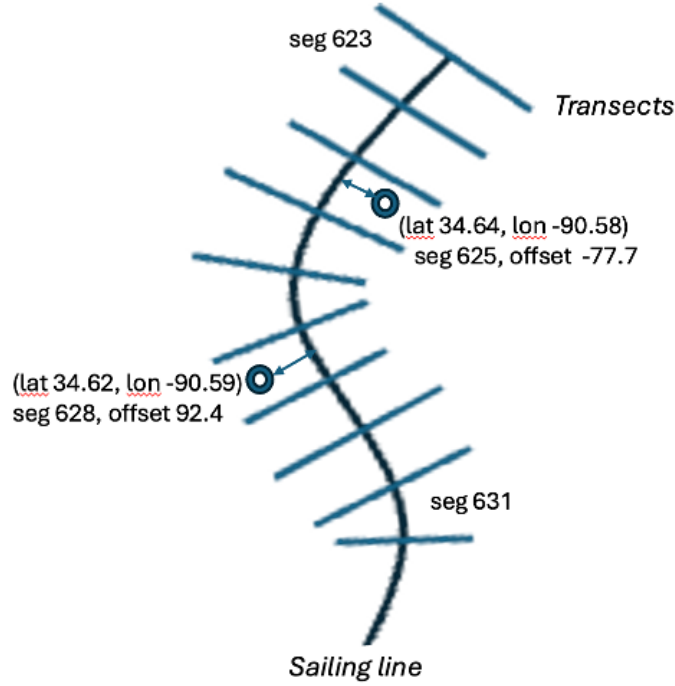


Figure 5.—Sailing line lat-long mapped to numbered segments and lateral position offset.

5pm, August 8 at 10pm to midnight, August 9 at midnight to 2am, etc., the operations were split into one operation on August 3 at 3pm to 5pm and one operation from August 8 at 10pm through August 9 at 2am. Each continuous time period was then individually processed to identify the location of dredging activity.

During an operating period, as defined by a series of positions over a continuous time period, a dredger may excavate in a small area, transit to a nearby area, excavate there, transit again, and so on. To select historical data sets for investigation of precursors to required dredging we needed to find locations and times of dredge activity in contiguous areas. Figure 6 shows the result of clustering dredger positions into distinct geographic areas. We used the HDBSCAN[†] clustering algorithm from the python `scikit-learn` (aka `sklearn`) package. Datasets were restricted to roughly 40,000 positions due to performance limitations of the development computer, a MacBook Pro with Apple M1 Max chip and 64 GB memory. Time periods with larger datasets were subsampled by keeping every n -th position with n defined by the size of the dataset divided by 40,000. HDBSCAN relies on distance between instances (positions, in this case) to determine clusters. We provided a precomputed Haversine great-circle distance (square) matrix for parameter metric. Additional non-default parameter values and their intent include:

`cluster_selection_epsilon` = 0.1 merge clusters closer than 0.1 km.

`min_cluster_size` = 50 minimum number of samples in a group for that group to be considered a

[†]HDBSCAN – Hierarchical Density-Based Spatial Clustering of Applications with Noise. The `scikit-learn` documentation states: “Performs DBSCAN over varying epsilon values and integrates the results to find a clustering that gives the best stability over epsilon. This allows HDBSCAN to find clusters of varying densities (unlike DBSCAN), and be more robust to parameter selection.”

cluster; groupings smaller than this size will be left as noise.

min_samples = 50 the parameter *k* used to calculate the distance between a point and its *k*-th nearest neighbor.

allow_single_cluster = *True* allow HDBSCAN to produce a single cluster, e.g., because the vessel appears to have operated in a single localized area during the time period.

The result was an individual file of AIS vessel positions for each cluster of operation for each dredger. These files identify sites of likely dredging operations which are then used to select times and regions of interest to train and test shoaling precursor detection techniques.

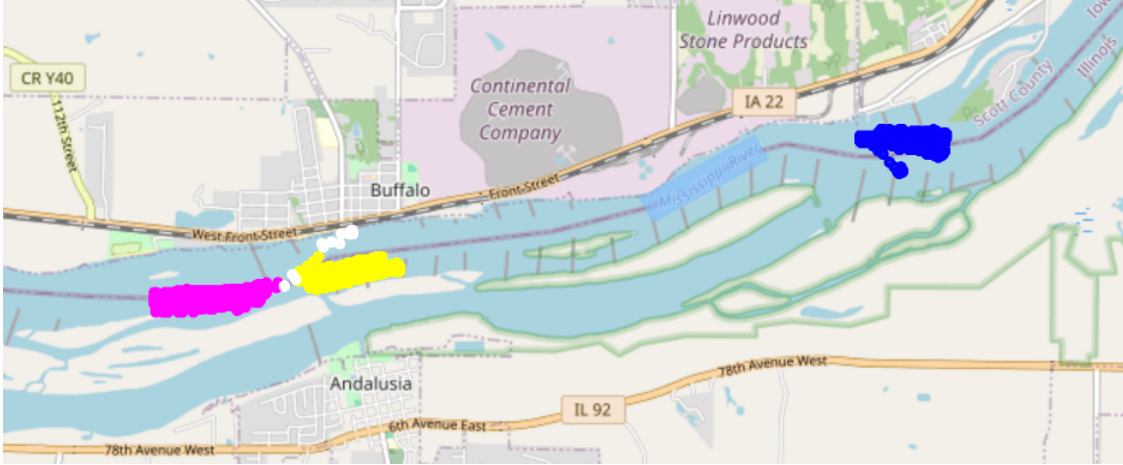


Figure 6.—Example of dredging operation clusters. Three clusters (magenta, yellow, blue) were extracted. White points represent ship transitioning; there are not enough nearby points to be considered a cluster and are presumed not to be indicative of dredging activity.

2.4 Study Data Specifics

Segment-offset mapping (Section 2.2) combined with dredge site detection (Section 2.3) provided the basis for our analyses. We selected river reaches of interest using dredge sites revealed by our clustering and availability of data on tow vessels in that area preceding and following the dates of the dredging. For the preliminary study, we concentrated on two areas – initially a reach near Vicksburg, MS, that was dredged in November 2024, and then a reach near Cordova, IL, dredged in October 2019.

The first step was to extract the sailing line through each reach from electronic charts, as described previously, and generate the 60-foot transect segments based on that line. This mapping covered about 10 river miles upstream and 10 river miles downstream of the target dredge site. River flow or stage data was collected covering 3 to 4 months before the dredge to about a month after the dredge. This river data is often available from nearby gauges and there are several online sources, such as the U.S. Geological Survey (USGS) National Water Information System Mapper at <https://apps.usgs.gov/nwismapper/>. This site allows the user to find water information resources available in a geographic area and download data sets for a given timespan. In this case, we are interested in surface-water sites, in particular stream gauges

that monitor flow or stage gauges that monitor river height. Not all sites will have nearby gauges, so flow must be inferred from upstream and/or downstream data. This flow/stage data is important since it can influence towboat captains who will often take a more direct course during higher flows. This could lead an automated shoal avoidance detection system to reach incorrect conclusions when the course deviations are influenced by differing river flows or stages rather than shoal avoidance. Partitioning the data analysis by flow rate, such as in the Tabbert, et. al. study (Ref. 5) (shown in Figure 1), could be useful in distinguishing shoal avoidance behavior from extraneous operational decisions. In the case of our two selected studies the flow was generally low throughout the study period, as might be expected when course deviations are influenced by shoaling (see Figure 7).

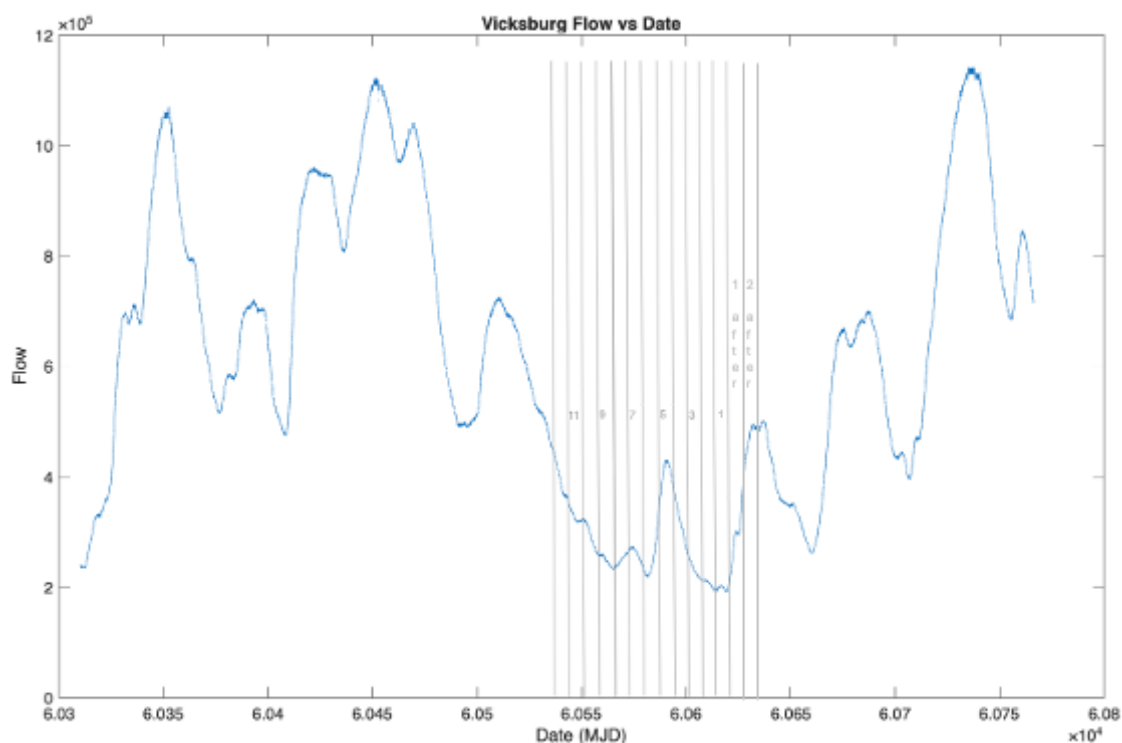


Figure 7.—River flow at Vicksburg, MS, gauge during study period (grey vertical lines).

On the advice of USACE subject matter experts, we concentrated on tow vessel (AIS type 31 and 32) transits through the reach of interest. A simple search through our LOMA data looking for vessel type 31 or 32 extracted the desired data records. These records were further divided by vessel identification (MMSI) and then temporally, looking for substantial time gaps in the AIS reports for the vessel and grouping records for each trip the vessel took through the reach based on those time gaps. Thus, we often ended up with multiple data files for a given vessel, each tracing a transit through the reach.

However, not all of the resulting records were appropriate for our purposes since many of the vessel tracks were local traffic and did not fully transit the reach. This was especially apparent in the Vicksburg traffic which had many vessels entering from the Yazoo tributary. We filtered out this extraneous traffic by ensuring that there was a continuous track between a point significantly

upstream and another point significantly downstream of the study area (regardless of direction of travel, as long as the overwhelming majority of the full course was either in an upstream or downstream direction). These tracks were then divided into upstream traffic or downstream traffic since we expected different shoal avoidance diversions depending on the direction of approach.

3 Analysis Techniques

3.1 Data Driven Anomaly Detection

Our initial approach for detecting shoaling areas utilized existing data driven anomaly detection tools. In particular, the cluster based Inductive Monitoring System (IMS) (Ref. 8) was considered a viable choice. IMS is a data driven system which clusters vectors of points representing nominal behavior and subsequently compares new, unseen points to those clusters, calculating a distance between the new point and the nearest nominal cluster (or, in a variant, an approximation of the average distance to the N nearest nominal points) to provide a measure of deviation from nominal. The original concept was to form AIS data from vessel tracks into vectors, group those vectors collected during non-shoaling periods into nominal clusters, then compare future tracks against these nominal representations. The vectors could represent a transit time series by concatenating data collected during a series of AIS updates. This data would not be limited to position data but could incorporate such parameters as speed over ground, course over ground, heading, or any other parameter deemed useful. In fact, an early test with limited data showed that such an approach could detect deviations from a nominal course, as illustrated in Figure 8. However, when applying this approach to a larger reach specific data set, IMS struggled to form nominal clusters that were sufficiently specific to provide robust detection of outlier vessel courses. The primary reason behind this was that vessel tracks tend to vary a great deal for a variety of reasons – beyond shoal avoidance, so there seems to be no particular nominal route through any specific section of the river. After a few trials we set aside this approach to investigate other options and did not have an opportunity to revisit, including after we had developed better normalization methods such as the segment-offset mapping. This approach, or other primarily data driven approaches, may still hold promise and have the added benefit of the ability to easily incorporate diverse parameters. After a few initial experiments, we pivoted to investigate other options. We believe it would be worthwhile to revisit this approach, or other primarily data driven approaches, with the better normalization methods. The ability to easily incorporate diverse parameters could prove beneficial.

3.2 Evolution of Traffic Patterns Over Time

Our next investigations provided more favorable outcomes. In essence, we developed visualization methods to reveal changes in traffic patterns during the months and weeks leading up to a dredging period, the thought being that quantifying changes in those patterns could provide a system to forewarn USACE resource managers that river variations, very possibly shoaling, are occurring in a particular area. We investigated two approaches – a curve fit of vessel courses through a reach and a count of transits and dwell times through lateral sections of a transect segment. Though we did not have time to develop and test actionable metrics, these approaches showed visual changes in traffic patterns and seem amenable to quantification. We recommend testing each approach on other dredging events in different river reaches and on traffic in sections of river that did not experience disruptions. If the predictive premise holds for one or

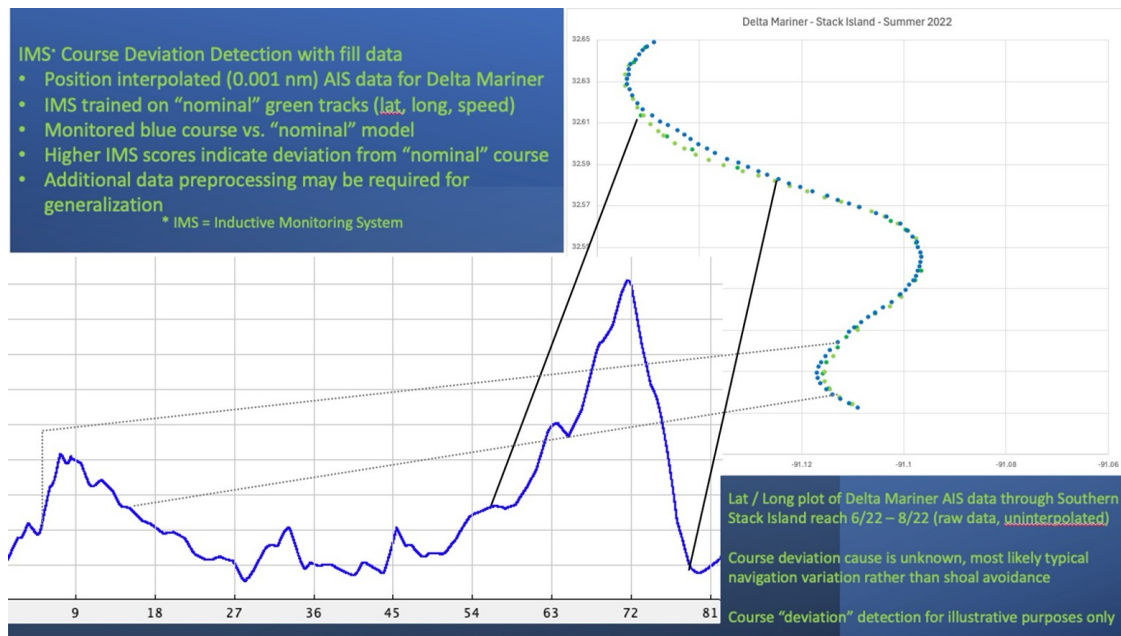


Figure 8.—IMS course deviation detection of a vessel passing through the Stack Island reach using vectors formed from latitude, longitude, and speed parameters.

both approaches, the next step would be to develop and test metrics that could be used to alert river managers. The following discussion suggests some high level metrics.

The inspiration for these visualization concepts was density histogram plots of traffic counts through a reach covering lower Mississippi River miles 769 – 770, just upstream of Memphis, TN, which experienced shoaling during the Spring of 2023. That shoal was cleared in mid-April 2023, by the Dredge Hurley. In the graph shown in Figure 9, the bars show the share of vessels that passed through a roughly one-tenth mile long section of river at different distances to the left and right of the sailing line. A visual inspection shows increased traffic toward the left descending bank in the months prior to dredging and a shift in traffic toward the right descending bank after the shoal was cleared. Notably, there is a large mud flat on the right descending bank in this section of river which likely expanded and impinged on navigation prior to dredging. These plots seem to indicate vessel captains steering clear of that growing shoal in the months prior to dredging and returning to a more direct course near the right bank after dredging. All transit counts were collected during similar, moderate river flow. Note - these plots used lower frequency AIS data obtained from Marine Cadastre.

3.2.1 Visualization 1: Vessel Course Curve Fitting

Because the initial data clustering efforts were not successful in extracting a typical vessel course through a reach, we tried a more direct approach of fitting lines and curves to the segment-offset coordinates of vessel LOMA/AIS position reports. This has the advantage of providing a well-defined path through the reach but, unfortunately, does not conveniently support including other potentially useful parameters, such as speed or heading, like clustering or similar data driven approaches could. Future work could explore if there is a hybrid approach that takes advantage of both.

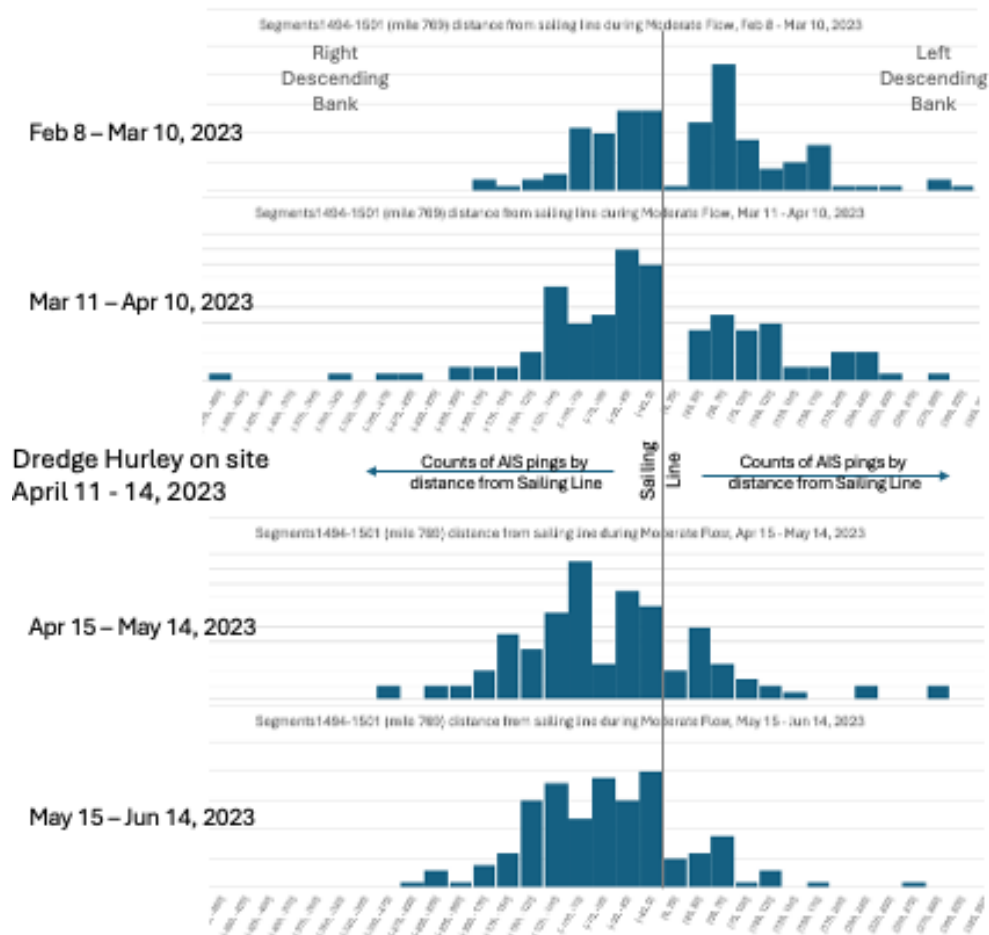


Figure 9.—Density histogram of transit counts through a river section upstream of a shoal for months preceding and following dredging of that shoal.

Since the segment-offset mapping effectively straightens out the course of the river, we first tried a linear fit to vessel position reports. To accomplish this, we collected AIS reported position coordinates for all tow vessels transiting the reach of interest in a particular direction (upstream or downstream travel) over given time periods prior to the dredging operation (and also after the dredging operation to assess the impact of clearing the navigation channel) similar to the transit count density histogram plots described previously (Figure 9). However, to ensure more evenly balanced data sets between time periods, rather than using data from a specific number of days or weeks, we measured time in transits. That is, a time period would correspond to, say, 15 vessels passing through the reach in the downstream direction regardless of how many days elapsed during those transits. In general, the traffic levels appeared regular enough that N transits covered a roughly similar number of days throughout the study period, though this should be confirmed with each data set since significantly diverging time scales may influence the results. Once the data was partitioned in this manner, a linear fit algorithm was used to determine a typical straight line course through the reach during that time period. Comparison of the lines from each period prior to the dredge did show changes in traffic patterns, but did not seem to show useful, consistent trends. In fact, in most instances, the quality of the linear data fit was

fairly low (e.g., high RMS error).

Even though the segment-offset coordinate mapping appeared to straighten the river, the vessel captains did not necessarily take a straight route through any given reach, hence the poor quality of the linear data fitting. To better capture the curving vessel courses, we hypothesized that a curved data fit is perhaps more appropriate than a linear fit. Indeed, applying a second order curve fit to the vessel position data seemed to better capture the traffic variations over time. Figure 10 shows variations in the traffic near Vicksburg, MS, in weeks prior to dredging activity. While the small sections of the curves shown here may not provide an intuitive path through the reach (in fact, they may better represent the distribution of transit paths), they do show distinct variations in the general traffic patterns as time grew closer to the dredge period. We theorize that these changes over time might be attributed to vessel captains altering their courses to avoid the growing shoal.

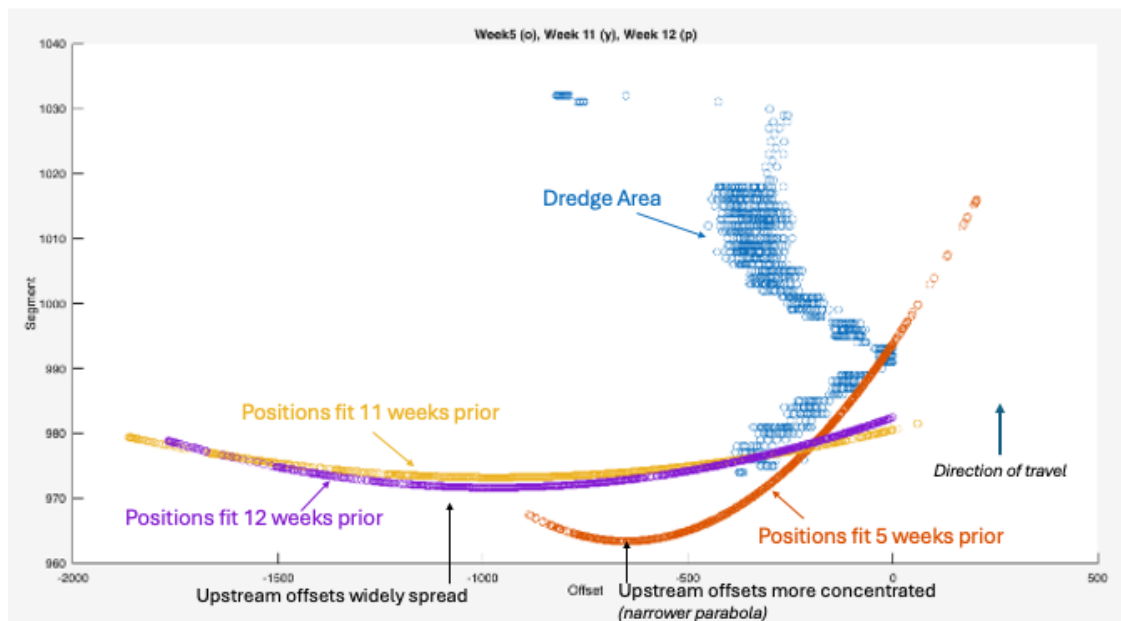


Figure 10.—Second order curve fits of South bound (downstream) vessel positions in the weeks leading up to dredge activity near Vicksburg, MS. The coordinates of the dredge ship during dredge activity, which occurred over four consecutive days, are shown with blue circles.

In more familiar geometric terms, a second order curve is simply a parabola, which can be characterized by a focus, vertex, and directrix. These features provide potential metrics to evaluate changes in the second order fit over time. Trends in changes in relative positions of the foci and vertices could indicate significant traffic pattern alterations in response to river variations, such as a growing shoal. While we were unable to fully explore such trends in this precursory work, a handful of observations did reveal some potentially useful metrics. The Vicksburg, MS, case showed some clear patterns in migration of foci over time; Figure 11 shows the focus change 12, 11, and 5 weeks prior to dredging and Figure 12 shows the change 9, 8, and 7 weeks prior to dredging. As dredge time approached, the foci of curves fit to Southbound traffic tended to be closer to the parabola curve itself – indicating a narrower parabola/approach corridor – and also shifted in offset toward the sailing line. A similar study in a different reach

near Cordova, IL, also showed changes in curve fit parabola over time (in both Southbound and Northbound traffic), though the trends in foci positions, while still present, were not quite as obvious as those seen in the Vicksburg example. Nonetheless, noticing changes in these curves over time could be a good indicator of traffic pattern changes. Further application of this curve fitting approach to data collected near other dredge sites could provide more insight and allow metric development and refinement.

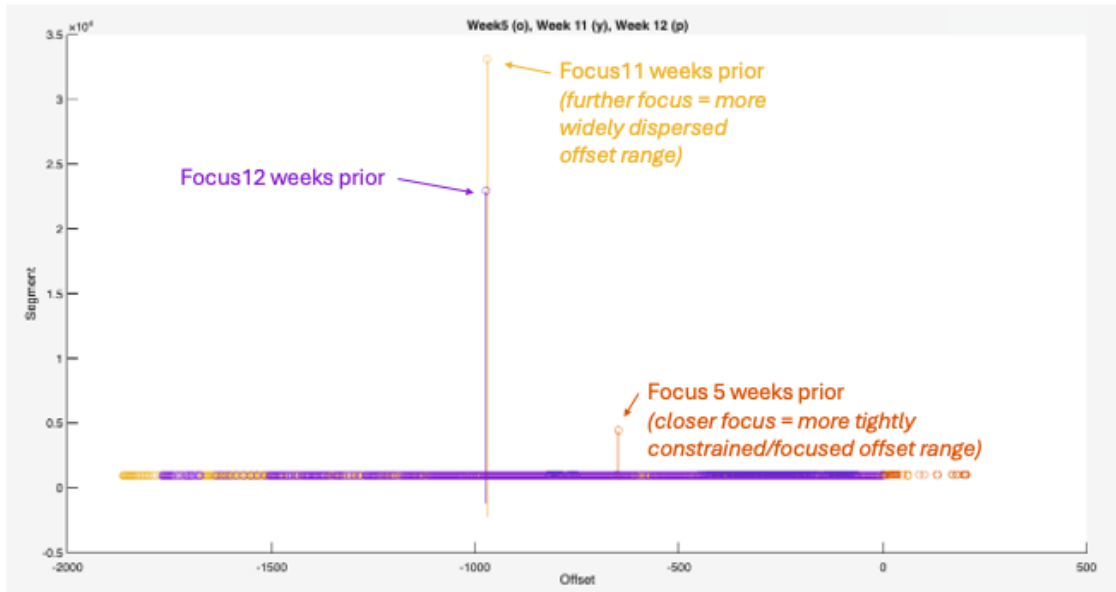


Figure 11.—Change in focus over time for second order curve fits of South bound (downstream) vessel positions in the weeks leading up to dredge activity near Vicksburg, MS.

3.2.2 Visualization 2: Transit Counts and Dwell Times Through Segments

Perhaps the most promising results came from an updated version of the transit count density histogram plots described previously. This technique used LOMA position data to count the number of tow vessel transits through 75-foot lateral subdivisions of a given transect segment over a period of time, where time was defined, as above, by a fixed overall number of transits. In both of our study areas, we saw a definite shift over time in transits, with lateral transit positions moving away from the eventual dredge area, as shown in Figure 13. A 3-D heat map plot, shown in Figure 14, of transit counts visualizes the phenomenon well. Note in Figure 14 the shift over time of initially widely dispersed transits away from the eventual dredge area (center of X-axis) and toward the left descending bank (right side of figure). Figure 15 shows an overhead view of the Figure 14 plot. In this view, it is apparent that the traffic migrated toward the left descending bank, with a fairly sharp inflection from widely dispersed to more concentrated away from the dredge site at about 2.5 months before the dredge. (It is notable that the more concentrated transit counts correspond to the more tightly constrained orange parabolas shown in Figure 11 and Figure 12.) This type of transit count pattern was also repeated in our study area near Cordova, IL, providing about a one-month warning in advance of that dredge activity.

Performing this type of transit analysis in several segments approaching known shoaling areas

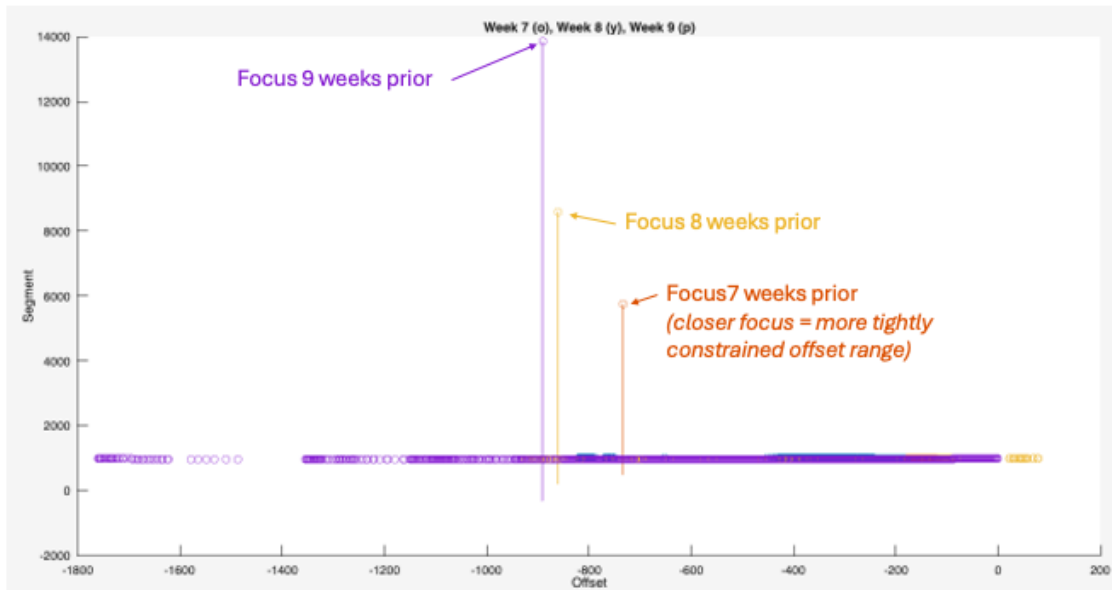


Figure 12.—Change in focus over time for second order curve fits of South bound (downstream) vessel positions in the weeks leading up to dredge activity near Vicksburg, MS.

(from both upstream and downstream) could be useful in alerting river managers to growing shoals. Developing a numeric traffic shift metric would enhance and simplify application of the visualization shown here. A measurement of changes to the distribution of transits over time across each segment is an obvious approach. Perhaps a distribution oriented measure, such as Kullback-Leibler divergence, would be useful. The grid based nature of the graph may also lend itself to a neural network approach. Or, more geometrically, something like looking at the slope of a line fit to the changing counts in given block over time, then comparing how that slope changes to the changes seen in the other blocks might provide insight. Intuitively, one would expect decreasing slopes in blocks covering the eventual dredge area as traffic decreased and increasing slopes in blocks away from the dredge area.

A complimentary analysis examining dwell time in a segment also showed promise. This

Segment divided into 75' wide blocks, count transits of each block

Segment 990 transit counts ~10 weeks prior	0	1	1	2	4	0	3	0	2	1	2	1	0	0	1	0
Segment 990 transit counts ~6 weeks prior	0	0	0	1	0	1	0	2	2	0	1	1	2	2	6	1
Segment 990 transit counts ~1 week prior	0	0	0	0	0	0	0	1	2	2	0	1	1	0	12	0

| Jadwin Dredge Area |

Figure 13.—Segment transit counts through 75-foot blocks in weeks leading up to dredge activity near Vicksburg, MS.

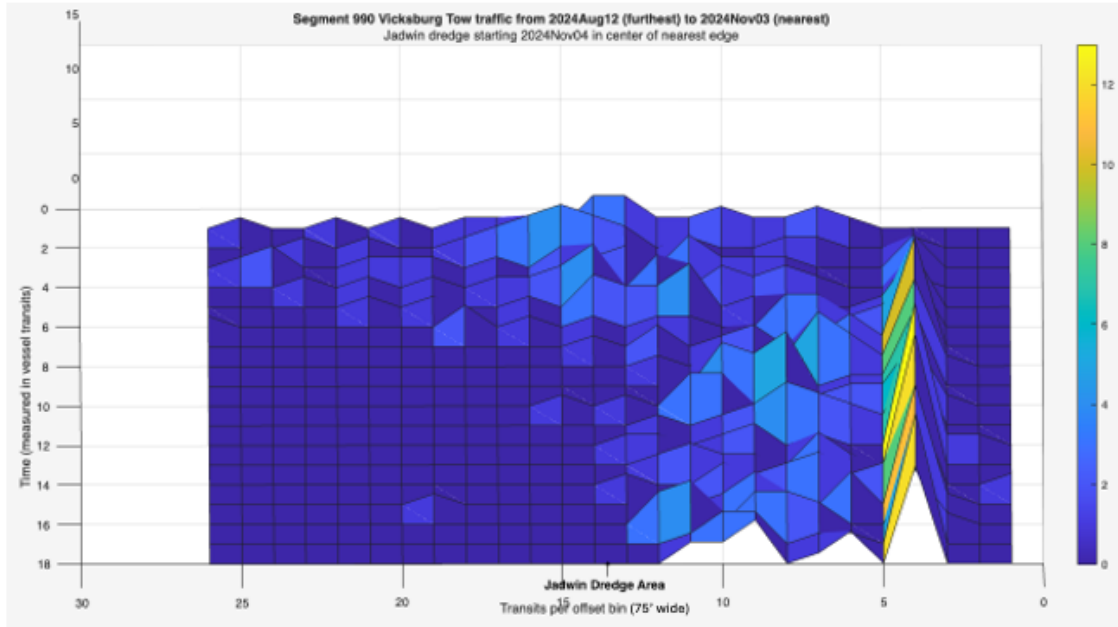


Figure 14.—3-D plot of Segment 990 transit counts (Z-axis) through 75-foot blocks (X-axis) in weeks (Y-axis) leading up to dredge activity near Vicksburg, MS. Earlier widely dispersed transit counts are shown toward top of the plot with more recent counts concentrated away from the dredge area shown toward the bottom of the plot. Dark blue squares indicate a count of zero transits in that segment subdivision during the time period, lighter colors are higher counts.

analysis measured the amount of time vessels took to transit a particular segment. Counts of vessels that took a particular amount of time to transit a segment during a particular week showed that in many cases the transit time for a segment increased, i.e., the vessels moved more slowly through the segment, as the dredge period neared, as shown in Figure 16. Perhaps this is a sign of captains approaching a hazardous shoal area more cautiously? These counts can be made simply by subtracting the timestamp of the position report when a vessel enters a segment from the timestamp when it enters the next segment, rounding or truncating to the nearest second, and tallying how many transits took 1 second, how many took 2 seconds, etc. A heatmap plot of the dwell times, as in Figure 17, shows similar tendencies to the transit count plot shown in Figure 15 and could likely use similar metrics to the transit count analysis with an emphasis on detecting increasing dwell times. A dwell time analysis performed on the Cordova, IL, reach showed similar results. Combining the two features (transit counts and dwell times) could add an extra dimension to the analysis not available from position reports alone.

4 Future Work and Conclusion

This technical memo was written to detail our preliminary efforts at analyzing Mississippi River vessel traffic patterns using AIS data to detect possible precursors to shoaling along the navigation path. While our analyses showed promise, our progress was limited by two main factors: our study period was much too brief to fully explore the possibilities, and we lacked reliable shoaling event ground truth to develop a well vetted, highly confident, robust analysis

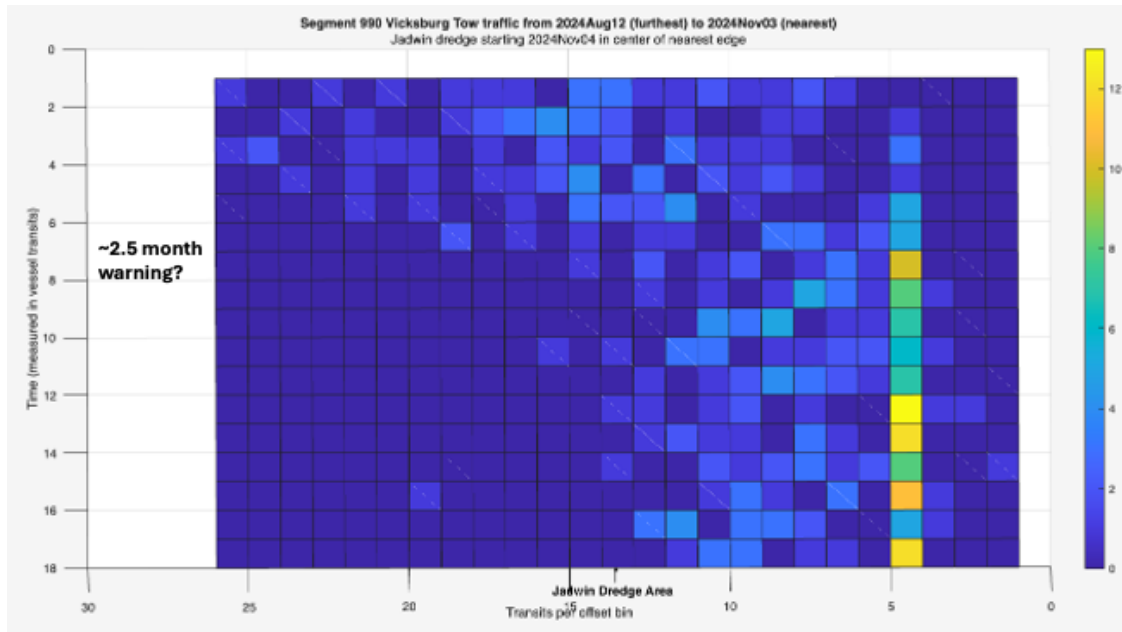


Figure 15.—An ‘overhead’ view of Figure 14 that shows a noticeable shift in traffic transit counts approximately 2.5 months prior to the dredge period.

system. As such, the intent of this write-up was to share our initial efforts in order to provide a starting point and guidance for researchers revisiting and expanding on this task. We stress that to make meaningful progress on modeling and method development, it is vital to obtain a diverse, wide-ranging database of times and locations of historical shoaling events throughout the course of the river along with detailed LOMA data covering those periods. Although we describe a method to infer the location of likely shoaling events, a reliable system must be based on substantial and certain ground truth.

Our first suggested follow-on activity would be to apply the analyses (or similar analyses)

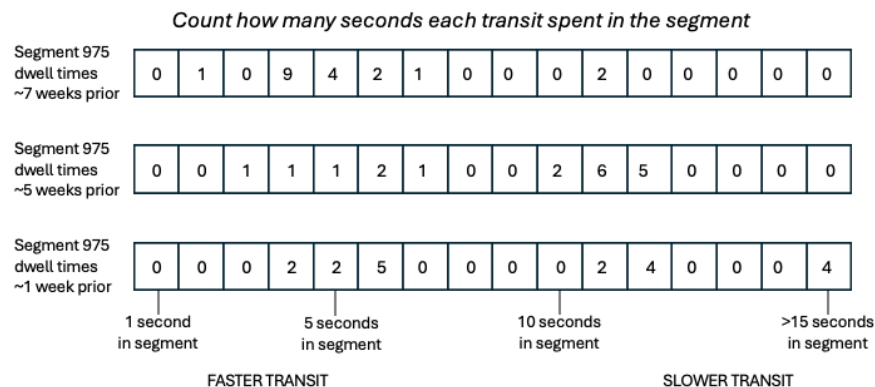


Figure 16.—Dwell time – number of vessels that spent a given number of seconds transiting through a transect segment during a given ‘time period’ showing counts of faster transits on the left and slower on the right. Transit times through segments on approach to the dredge area tended to increase as dredge time neared.

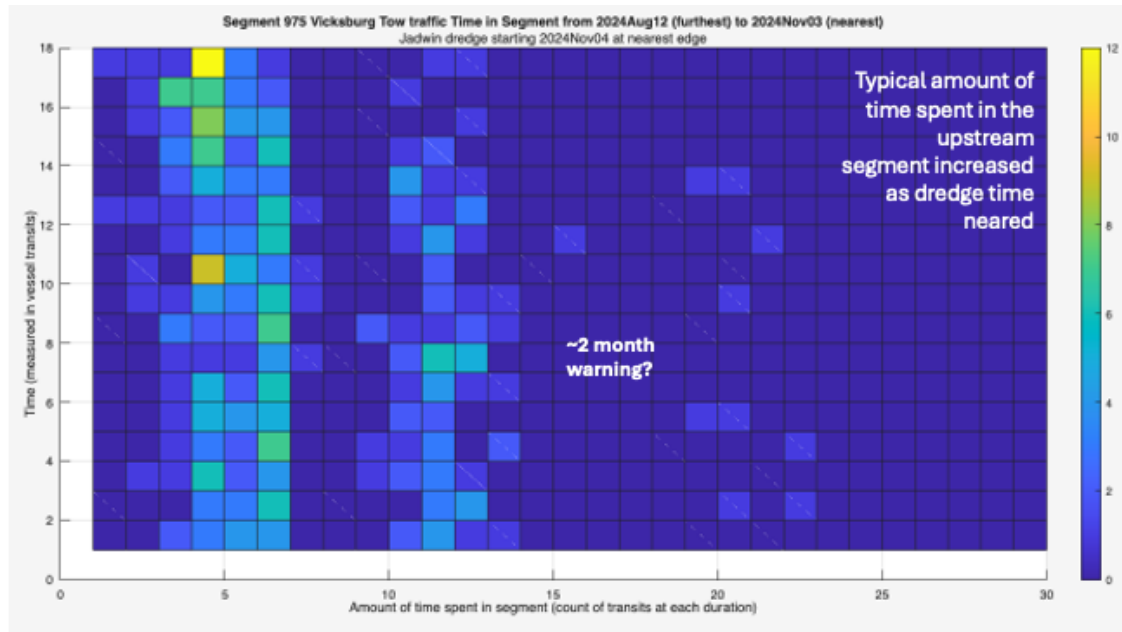


Figure 17.—Dwell time heat map – number of vessels that spent a given number of seconds transiting through a transect segment during a given ‘time period’ showing counts of faster transits on the left and slower on the right. Dark blue squares indicate zero transits of that duration in that segment during the time period, lighter colors are higher counts.

outlined in this report to other reaches of river, both those that have experienced shoaling events and those that have remained clear for navigation. The two case studies discussed in this report showed promise for providing advanced notice of shoaling; additional study is required to confirm that the techniques generalize. If either approach withstands further testing and scrutiny, then metrics should be developed that could signal changes in traffic patterns that indicate possible shoaling or may be useful indicators for other river management.

Some possible metrics were discussed in earlier sections. With regard to data curve fits, measuring changes over time in the curves (e.g., changes in focus and vertex locations) and noticing what kind of changes occur prior to shoaling could be fruitful. For transect segment transit counts and times, shoaling may be indicated by decreasing transit counts in some bins and increasing counts in others as well as shifts in transit times in a segment. Multiple segments in both the upstream and downstream approaches to known or suspected shoaling sites should be monitored. Keep in mind that each reach has different characteristics, so meaningful trends in one area may have a different character than meaningful trends in another. Applying metrics to historical data will help establish meaningful trends for that particular reach.

The techniques spelled out in this report represent only a few possibilities. Our limited study precluded further exploration. Perhaps other data driven approaches could be successful? Our efforts at cluster-based anomaly detection were abandoned early on due to time limitations, but there may be alternate or hybrid data-driven anomaly detection routes with better results. We were also unable to try neural network or deep learning approaches. One deep learning tool that may be a good fit is Deep Temporal Multiple Instance Learning (DT-MIL), a NASA developed system that “learns” decisions made by experienced captains and compares future approaches or

transits to the methods those experts used. (Ref. 9) Decisions made by experienced captains under nominal river conditions might differ from those made when navigation hazards are building.

The framework set up for our studies, mapping AIS lat-long positions to segment-offset coordinates, was quite useful and could form the basis for future investigations. The winding nature of the Mississippi River makes traffic pattern studies difficult, so such a “straightening” of the river may be required. Furthermore, knowledge of past shoaling events is vital to discovering shoal predicting patterns in the data. There may be archived records of shoaling or dredging activity in some river reaches, but we were unable to access those records so our method of discovering dredge activity in the AIS/LOMA data proved invaluable. Also, as is often the case, the more detailed data available through the LOMA system was far more useful than lower fidelity AIS data obtained from public sources. Fortunately, the Mississippi River is well instrumented for LOMA data collection with good coverage in most reaches. Collaboration with USACE engineers and researchers provided access to the tremendous LOMA archive.

Finally, a user interface to present traffic pattern alerts or trends to river managers will be essential. We were unable to investigate interface designs but envision some form of graphical interface to visualize and highlight traffic pattern changes. In addition, some sort of daily or weekly summary of notable traffic trends, similar to the NASA aviation “Morning Report” system (Ref. 10), could be an effective method to alert operators and managers to significant activity which could then be reviewed using the graphical interface.

The Mississippi River is a vital U.S. waterway which carries billions of dollars of freight each year. Navigation disruptions are very costly, so timely management and dredging of shoals in the navigation path is essential. There is a great deal of vessel data available through the AIS/LOMA data system, and our early results show promise that the data can be utilized to provide forewarning of shoaling or other river traffic disruptions. Such forewarning would help U.S. Army Corps of Engineers personnel more efficiently and effectively manage their limited resources to maintain river channels and minimize traffic flow disruptions, thus augmenting the valuable services they provide to our country.

References

1. PBS News Special: Tipping Point - The Mighty Mississippi. <https://www.pbs.org/newshour/nation/watch-tipping-point-the-mighty-mississippi-a-pbs-news-special>, July 15 2025. Accessed: 2026-07-13.
2. U.S. Army Corps of Engineers Mississippi Valley Division: 05_Navigation_WEB. https://www.mvd.usace.army.mil/Portals/52/docs/05_Navigation_WEB.pdf, May 2019. Accessed: 2026-07-13.
3. Johnson, C.: Nonstop dredging kept Mississippi River open this year, but moving sand creates its own problems. <https://wisconsinwatch.org/2023/12/mississippi-river-wisconsin-minnesota-sand-dredging/>, December 14 2023. Accessed: 2026-07-13.
4. Schexnayder, C.: Corps projects to deepen Mississippi River heads toward final phase. <https://waterwayscouncil.org/media/in-the-news/article/2022/02/corps-projects-to-deepen-mississippi-river-heads-toward-final-phase>, February 7 2022. Accessed: 2026-07-13.
5. Tabbert, C.; Vest, J.; Rhoads, A.; Myers, D.; Lauth, T.; Brauer, E.; Wallace, J.; Gordon, D.; and Kress, M.: AIS Data Case Study: River Level and Vessel Approach Variation at Melvin Price Locks and Dam in St. Louis District. 2020. ERDC/CHL CHETN-IX-51.
6. Kontas, V.: What kind of information is AIS-transmitted. <https://support.marinetraffic.com/en/articles/9552860-what-kind-of-information-is-ais-transmitted>, May 7 2025. Accessed: 2026-07-13.
7. USACE: Inland Electronic Navigational Charts. <https://ienccloud.us>, 2001. Accessed: 2026-07-14.
8. Iverson, D. L.: Inductive system health monitoring. *International Conference on Artificial Intelligence*, 2004.
9. Janakiraman, V. M.: Explaining aviation safety incidents using deep temporal multiple instance learning. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 406–415.
10. Maille, N. P.; and Statler, I. C.: Comparative analyses of operational flights with AirFASE and the morning report tools. 2009. NASA/TM-2009-215379.
