Generation of Inland Waterway Trip Information using Automatic Identification System (AIS) Data

Submission Date: August 1, 2012

Word Count: 4943

James P. Dobbins, Ph.D. (corresponding author)
Affiliation: Vanderbilt University
Address: VU Station B-35 1831
Nashville, TN 37235-1831
Phone: (615) 322-0430
Fax: (615) 322-3365
Email: j.dobbins@vanderbilt.edu

Lindsey C. Langsdon
Affiliation: Vanderbilt University
Address: VU Station B-35 1831
Nashville, TN 37235-1831
Phone: (615) 322-2697
Fax: (615) 322-3365
Email: lindsey.c.langsdon@vanderbilt.edu

ABSTRACT

This research paper documents the analysis of archived AIS data in the Paducah, KY region to produce reliable inland waterway vessel trip data. There are currently few options when seeking such trip data, owing to confidentiality concerns; this absence impacts the quality of risk calculations. The combination of geographic information systems (GIS), relational databases, custom programming and data visualization tools are applied to extract meaningful vessel traffic information and detect events occurring within ports and waterways. The geographic configuration of the Paducah port area added a degree of difficulty to the generation of trip data. However, this was overcome by categorizing all trips into general river movements and calculating the total number of towboat trips transiting the area in through river movements or engaged in fleeting, docking or lockage operations. The AIS data was discovered to be of high quality, and capable of supporting a number of analyses. These include waterway and port congestion, hotspot identification, accident reconstruction (and near-miss investigation), and the impact of extreme weather on areal port and waterway traffic.

INTRODUCTION

A recent research effort into quantifying waterway accident rates was hindered by the absence of reliable and detailed vessel trip data (*1*). The researchers were able to determine the frequency of allisions, collisions and groundings on the U.S. inland waterways, but not the rates that would form the denominator in the risk calculation (e.g., groundings per million vessel trips). This paper describes an attempt to archive and use automatic identification systems (AIS) vessel position data to generate reliable trip information.

AIS is a technology that enables vessel to vessel, vessel to shore and shore to vessel communications. Shore stations include vessel traffic services (VTS), in charge of coordinating traffic within congested port and waterway areas. With an available geographical point and click interface, operators may obtain information about other vessels including call sign, vessel name, position, course, speed, size, cargo, destination and other relevant attributes. Most relevant to this study are the vessel name, course, speed and position (latitude and longitude). The coordinate data is GPS-based and the reporting interval is once every two to ten seconds (when underway). As a transponder-based technology AIS overcomes many limitations of conventional radar, including:

1. terrain blind spots – vessels around a bend are seldom visible using conventional radar,
2. vessel differentiation – vessels appear as features with attribute information in AIS (enabling vessels to be called directly over VHF radio) as opposed to simple blips in radar, and
3. complete vessel navigation integration – AIS has the capability to combine the entire vessel navigation picture into a single screen, displaying aids to navigation, other vessel positions, collision avoidance information (closest point of approach, etc.) and weather all within an electronic nautical chart.

AIS is standardized by the International Telecommunications Union (ITU) and adopted by the International Maritime Organization (IMO) (*2*). AIS carriage is required for vessels of 65 feet or longer, 300 gross tons or more, vessels on international voyages, among other categories (*3*). On the inland waterways AIS is only required in the following areas as per Table 161.12(c) from 33 CFR §161.12 (*4*):

1. Lower Mississippi River (within the VTS area of operations) up to 20 miles above Baton Rouge, LA at milepost 254.5)
2. Ohio River (within Louisville, KY VTS area of operations) between mileposts 593 and 606 when the McAlpine upper pool gauge is at approximately 13.0 feet or above.

OBJECTIVE

There are several maritime websites that broadcast AIS data received from networked antennas around the world. The majority of these websites broadcast the most recent vessel position information; essentially painting the picture of port and waterway vessel traffic conditions. Vessel agents and shipping company personnel often use subscription services receive updated information on vessel estimated times of arrival (ETA) and voyage status. The IMO recognizes the safety and security issues involved in broadcasting real-time vessel positions and "condemns the regrettable publication on the world-wide web, or elsewhere, of AIS data transmitted by ships." (*5*) Broadcasting an AIS feed is not a very capital-intensive operation: a VHF antenna, AIS receiver, serial connection to a PC, and a live internet connection are all that is required. Sites that broadcast AIS information have been around for several years, however, to date there is little documentation about efforts to process historical AIS data to generate better trip information on the inland waterways.

Perez et al analyzed vessel traffic patterns in the Houston/Galveston nonattainment area using an AIS data archive (*6*). Researchers intended to use the AIS data to show the start and stop time of vessels to estimate commercial marine vessel emissions in Texas coastal waters. Though the researchers recognize the potential of AIS data, gap issues and carriage requirements (not all vessels were required to carry AIS) caused them to use aggregated trip data (by vessel type) published by the U.S. Army Corps of Engineers on a subset of the National Waterway Network (NWN).

The objective of this project was to establish the proof of concept of converting a 6-month archive of AIS data into reliable vessel traffic information. The archive was built using an AIS antenna and PC installed at a Paducah, KY river terminal. All AIS messages received from March 30, 2011 through September 30, 2011 were included in the archive, excluding 22 days due to a variety of computer outages. This data was then decoded and processed using custom software and geographic information system (GIS) models. The complex configuration at Paducah (the meeting point of the Ohio, Tennessee and Cumberland Rivers) added to the difficulty of generating useful trip data. Figure 1 contains a map of the Paducah port area. Note that the AIS positions show clustering of the main traffic patterns and the range of the receiver is approximately 17 miles. There are 4 locks within range of the receiver: Barkley Lock and Dam, Kentucky Lock, Smithland Lock and Dam, and Lock and Dam 52.

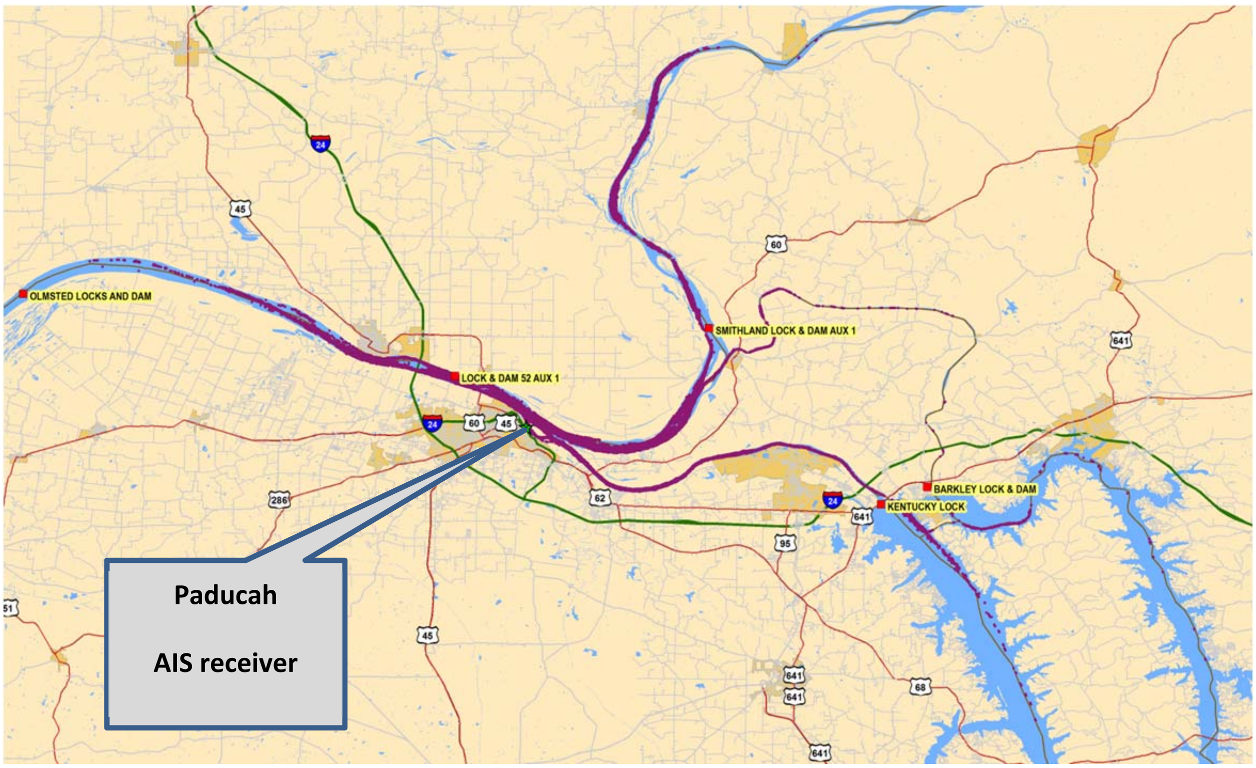

FIGURE 1 Map of Paducah, KY case study area.

MOTIVATION FOR WORK

There are few options when searching for inland waterways trip data. The U.S. Army Corps of Engineers (USACE) Navigation Data Center (NDC) receives trip reports from barge carriers, including commodity, quantity, and origin-destination. The data released to the public are aggregated to prevent the disclosure of sensitive business information about carrier operating patterns and commodities. Publicly available tonnage data may be joined to the National Waterway Network (NWN), a GIS representation of U.S. navigable waterways (*7*). The tonnages are reported by river direction (upbound/downbound) and general commodity groupings, such as chemicals, petroleum products, waste and raw materials. The number of empty barge trips occurring on the waterway is not documented. A separate data source, "Waterborne Commerce of the United States," publishes trips and drafts data for waterway segments (*8*). It is possible using these data to determine the number of barges pushed upstream on the Ohio River in 2010 with a draft of 8 feet, for example. However, trips and drafts data is aggregated and reported over such long waterway segments (the entire Ohio River is reported as a single segment) that they are not useful as the denominator in a localized casualty rate calculation.

The USACE operates the Lock Performance Monitoring System (LPMS) and the Corpslocks (formerly known as OMNI) information system (*9*). The LPMS contains a vast repository of data on lockages (including commodity and directional information) which is useful as an indicator of vessel traffic and agricultural performance. Corpslocks provides information about lockages in near real-time and includes vessel arrival, locking and departure time information. Historical queries are supported and were used in the validation of this research. This included the number of lockages at each of the locks within the case study area. This data has limitations, as numerous vessels operate solely between locks and there are no locks on the busiest waterway in the United States, the Lower Mississippi River. Figure 2 shows a map of lock locations in the United States.

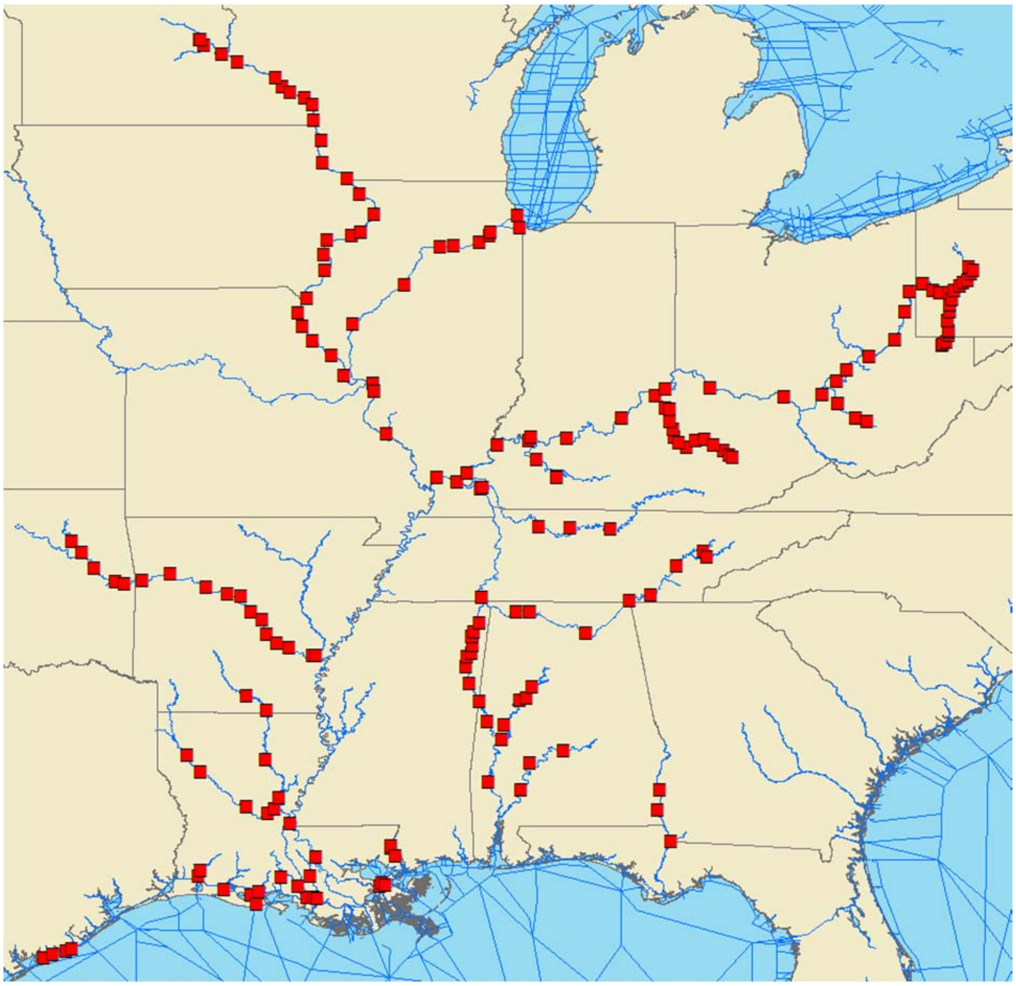

FIGURE 2 Map of USACE Locks in United States.

METHODOLOGY

The AIS message is a binary format, and there are 27 different message types. All vessels are assigned a unique identifier known as the Maritime Mobility Service Identity (MMSI) number. For this research the messages of interest were position reports (type 1, containing latitude/longitude, speed, and course over ground) and the static/voyage data (type 5, containing vessel name and radio call sign). The MMSI number serves as the linkage between these two message types. All AIS messages received by the receiver were transcribed into text files via PC serial connection in four hour intervals. These text files were moved to a file server nightly for processing. Customized software routines translated and combined the binary messages into ASCII-formatted text files for use in the GIS. Sampling (selecting positions every 15 minutes) was deployed to reduce the size of the dataset, as most GIS software is not adept at handling large datasets. The 6-month archive measured 5.5 gigabytes (compressed binary data) and contained over 158 million messages.

A customized model inside the GIS ordered the position reports by MMSI number, then by date/time, before “connecting the dots” to convert a series of vessel positions into a daily track. Vessel positions were grouped by day, starting at 12:00:00 AM and ending at 11:59:59 PM. This grouping creates a line feature for each vessel/day combination and enables trip counting. Without such grouping, there could be no determination of when a vessel entered or left the case study area.

During the 6-month period, 473 unique towboats made a total of 10,247 trips through the case study area. These trips are displayed in Figure 3, along with every 5th mile marker. Trips that appear to traverse land, especially between Barkley Lock and Dam and the Ohio River and between miles 910 and 925 on the Ohio River are the result of the 15-minute sampling interval and the terrain. Between the receiver and these areas are large hills that occasionally prevent clear communication between waterborne AIS units and the Paducah antenna. This behavior was expected as AIS relies on line-of-sight VHF radio transmission. A GIS digital elevation model (DEM) analysis verified the existence of blind spots in these areas.

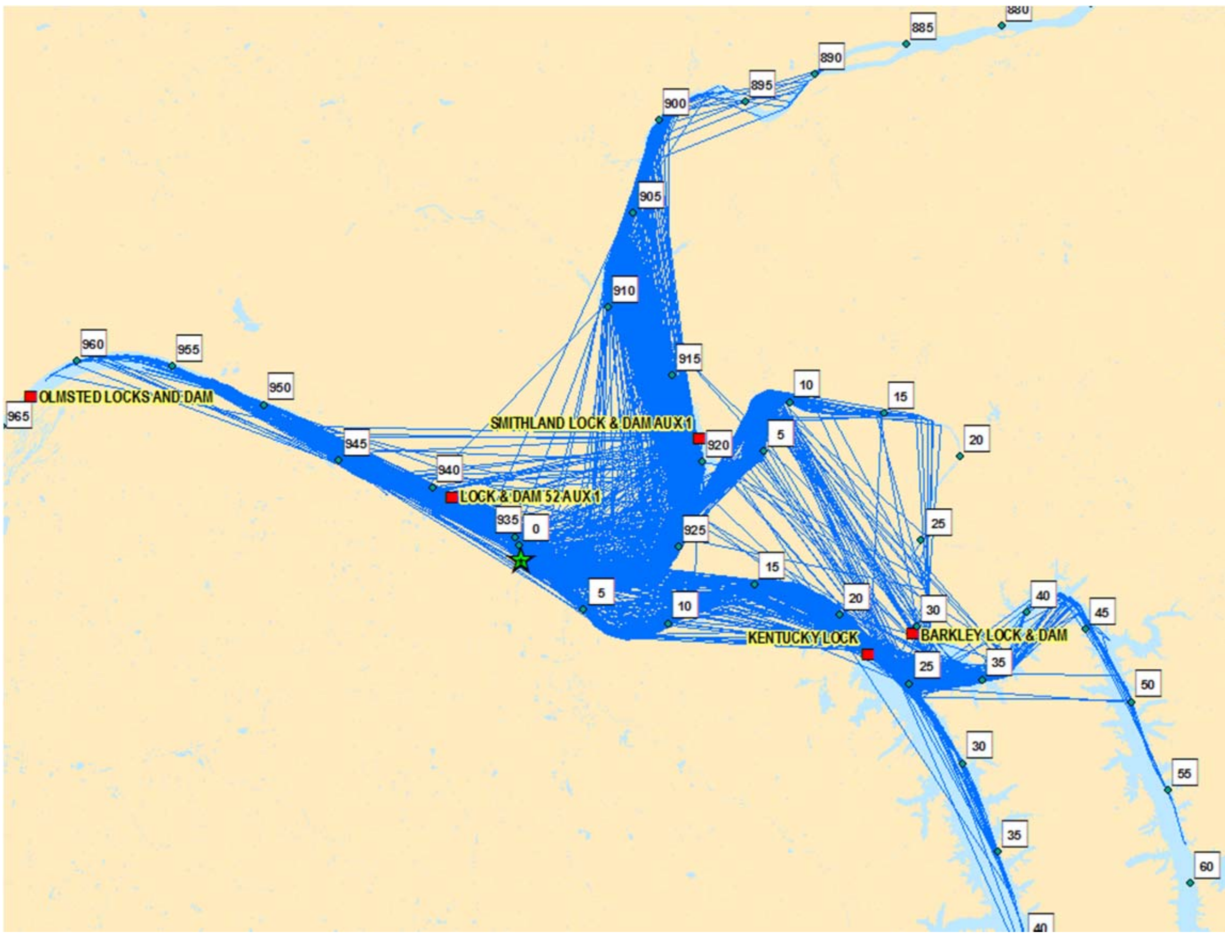

FIGURE 3 Trips and Mile Markers in Case Study Area at Paducah, KY.

The trips layer in Figure 3 could be used to accurately count the number of trips (and direction) occurring at a single point on any of the rivers in the case study area. However, a more useful analysis would be to understand the predominant trip patterns throughout the area. This was accomplished by determining the beginning and ending river and milepost for all trips. To support this analysis an area layer was generated with a polygon feature (a cell) at each milepost as shown in Figure 4. Junction areas such as the Port of Paducah were merged to better identify fleeting and docking events.

FIGURE 4 River cell layer used to locate trip begin and end locations.

Each river/milepost O-D pair was translated into one of 41 possible general movements. These general movements included values such as “Same cell” for fleeting and terminal operations (where the tow did not leave the cell during the day), “Up Ohio” to indicate an upstream, through movement on the Ohio River, “Ohio down to Paducah” for a movement that goes downstream on the Ohio and stops in the Paducah port area, “Down the Cumberland and Up the Ohio,” among others.

Detected stop events were studied using cluster analysis. This identified lockages, barge fleeting locations and active terminals. The AIS data was precise enough to differentiate lock chamber activity at a given lock, as shown by the identified lockages at Lock and Dam 52 in Figure 5.

FIGURE 5 AIS position data at Lock and Dam 52.

The results of a simple cluster analysis of all tow positions (where the speed was less than 1mph, indicating a stopped/fleeting/dock event) is shown in Figure 6. The color scale goes from green to red, indicating a larger amount of activity. Cells with little to no activity were excluded from this map. Using high-activity cells, fleeting areas, major terminals and lock chambers were digitized using an underlay of aerial photography, similar to the one shown in Figure 5. These locations were validated by comparing them with dock features described in the U.S. Army Corps of Engineers Navigation Data Center Port Series dataset, Master Docks Plus (*11*).

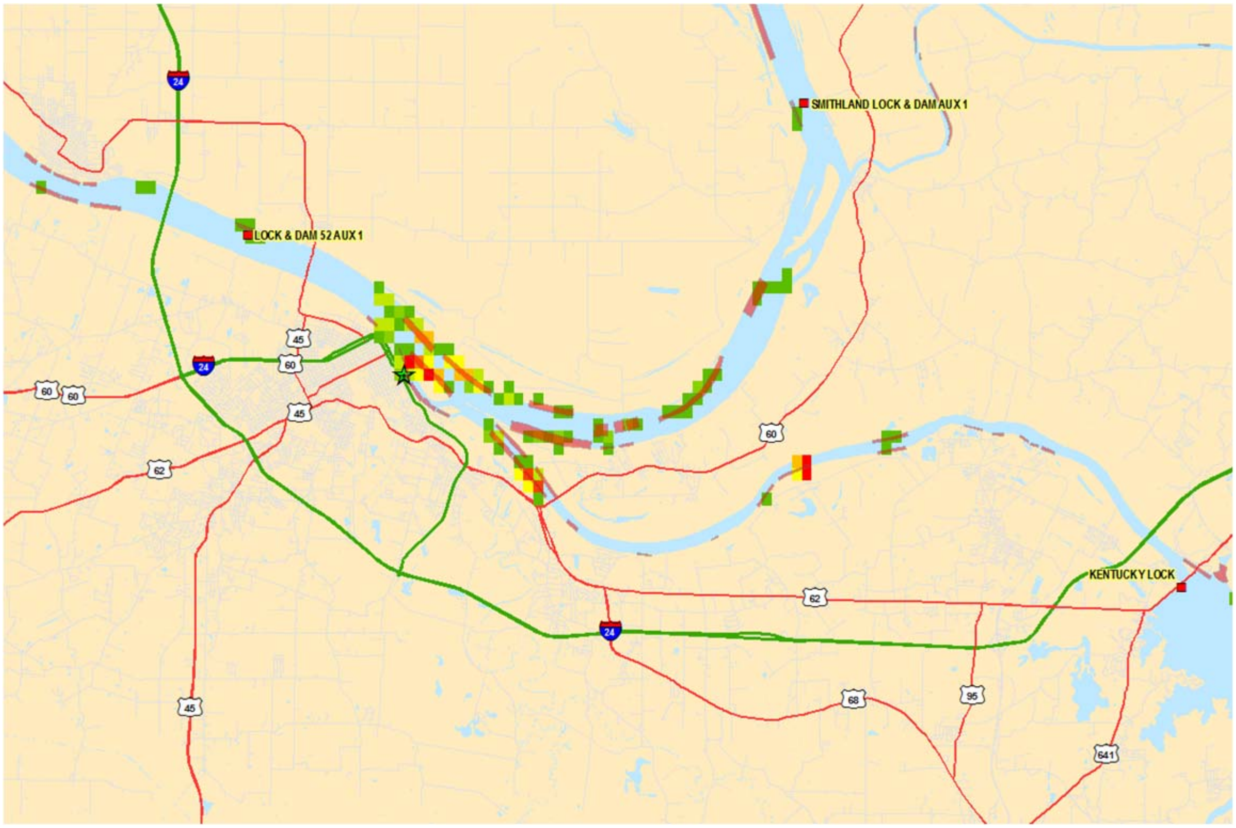

FIGURE 6 Analysis of stop events in the Paducah, KY region.

When dealing with time-stamped data such as the AIS archive used in this study, GIS software can animate the data in a playback of events. Animations were built that reenacted a "day in the life" of the case study area. However, the volume of data (vessel positions for every 10 seconds) was too much for a powerful workstation to smoothly process the data for more than an 8-hour window. Still, this functionality could be used in accident reconstruction, near-miss analysis and other detailed operational traffic examinations.

RESULTS

As the methodology section demonstrated, the data is of very high quality and could serve a number of potential uses. The combination of GIS, customized software and data visualization tools are all that is needed to convert the AIS data archive into valuable trip and port activity data.

Due to the 22 days of outages during the 6-month study period, the statistics gathered through the AIS analysis are incomplete when comparing to other data sources (such as the U.S. Army Corps of Engineers Lock Performance Monitoring System). A vessel trip for the purposes of this research is considered to be one day's track. If a fleeting towboat makes several runs across the river to pick up and drop off barges, the day's track counts as one trip. Vessel trips per day and the outage dates are shown in Figure 7. Note the peak and drop in trips in mid-May, a month that saw historic river levels and numerous waterway closures from Cairo, IL to Baton Rouge. This peak before the drop could have been due to river closure discussions, where tows were attempting to get through the area per-closure (either up the Ohio or up the Upper Mississippi River).

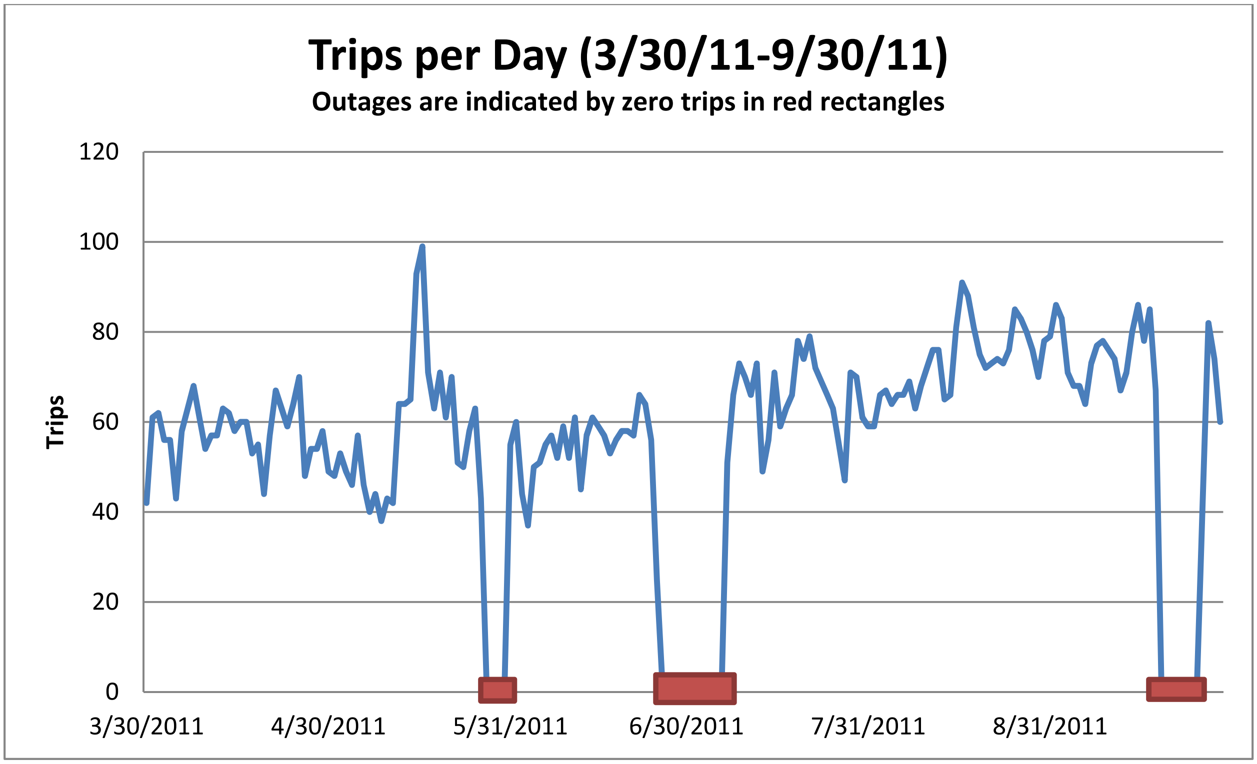

FIGURE 7 Number of trips per day (3/31/11-9/30/11).

Selected results from the analysis are displayed in Table 1. This includes a table showing the number of trips for all possible general movements by month (having at least 25 trips in any month). Of interest is the dip in the number of trips in the month of May 2011. The drop in overall trip traffic and increase in the number of same cell trips (that could indicate idling, or docked, tows) could be attributed to the prevailing river conditions.

TABLE 1 Number of Trips by General River Movement (Minimum 25 Trips per O-D Pair)

Movement Description	April	May	June	July	August	September	Total
Same Cell (fleeting)	609	800	447	619	792	588	3,855
Up OH	254	152	187	243	338	274	1,448
Down OH	224	146	187	189	352	249	1,347
TN Down	53	41	36	51	82	38	301
Paducah to Up OH	40	43	42	42	62	35	264
OH Up to Paducah	47	34	34	42	58	47	262
TN Down, OH Down	54	32	39	39	46	35	245
OH Down to Paducah	47	25	40	39	57	36	244
TN Up	33	30	32	43	75	31	244
Paducah to Down OH	47	31		35	72	47	232
OH Up, TN Up	45	40	25	29	39	29	207
TN Down to Paducah	36	34		38	50		158
Paducah to Up TN	42		31	39	32		144
TN Down, OH Up			28		31		59
Down Ohio, Up TN					30		30
CU Down, OH Down			26				26
OH Up, CU Up					25		25

As previously stated, AIS carriage is not required in the case study area. The trips identified in this study only include vessels transiting the case study area with operational AIS receivers. An informal survey of towboats operating with AIS receivers in the Paducah area was possible by comparing the number of lockages captured through the AIS data with the official number of lockages reported by the Corps of Engineers in the OMNI Reports information system (now known as Corpslocks). At each lock, the actual number of lockages was found to be three times the number of AIS-detected lockages.

Through examination of the data, several observations were made. Though the AIS message structure has fields for estimated time of arrival (ETA), destination and cargo, these fields were almost never populated. Additionally, it appears the call sign and name of the vessel can be manually overridden, judging by the presence of typos in each field.

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

This effort has established proof of concept that archived AIS data may be used to perform historical trip analysis and that the data may also be used for a number of other studies. The combination of custom software/programming, GIS, enterprise-grade relational databases and data visualization is uniquely capable of processing the data and visualizing trends and overall statistics. This area of study is ripe for further exploration and a number of recommendations are presented below.

A more sophisticated means of detecting and delineating trips should be researched. Instead of cutting trips at midnight, a data structure that accommodates counting of trips that begin and end before and after midnight, respectively, should be explored. Line thinning techniques within the GIS should also be applied to keep the size of the data manageable.

The geographic configuration of the Paducah area, with 3 rivers, 4 locks and a canal, added a degree of difficulty to the conducted analysis. The full potential of the data may be realized by examining AIS data in a standard river location, preferably in an area of mandatory AIS carriage. Such analysis could include real-time translation of the data, as opposed to historical analysis. Adding another antenna at another location (on a different river) and “tracing” trips throughout the inland waterway system is an especially daunting challenge. Relative levels of waterway congestion could be established using GIS point density routines.

Analytics and data visualization tools are especially adept at processing and visualizing the data. These types of software, along with data warehousing/data mining relational database management systems, should be explored further with AIS data. Particularly, enterprise relational databases could be used to perform the binary to ASCII translation, eliminating a point of failure by removing the need for custom software outside the database.

The effect of flood conditions during the month of May 2011 was especially interesting to observe. With a networked group of AIS receivers and archived data, the influence of extreme weather conditions on vessel traffic could be quantified.

ACKNOWLEDGMENT

This research was funded by the Intermodal Freight Transportation Institute (IFTI) at the University of Memphis. Additional funding support was provided by the Ingram Barge Company.

REFERENCES

1. Dobbins, J. and Jenkins, L. "Estimating Coastal Maritime Risk Using Geographic Information Systems." *Transportation Research Record*, Transportation Research Board, National Research Council, Washington, D.C., 2011, Volume 2222/2011, Pages 17-24.

2. U.S. Coast Guard Navigation Data Center. U.S. Coast Guard AIS Regulations: Now and Proposed. http://www.navcen.uscg.gov/pdf/AIS/USCG_AIS_Regs_Current-Proposed_v2.pdf. Accessed June 2012.

3. International Maritime Organization (IMO), International Convention for the Safety of Life at Sea (SOLAS). Regulation 19, Chapter V (Safety of Navigation). http://www.imo.org/ourwork/safety/navigation/pages/ais.aspx London, United Kingdom. Accessed 2012.

4. Code of Federal Regulations, Title 33: Navigation and Navigable Waters. Section 161.12 Vessel operating requirements.

5. International Maritime Organization (IMO). *Automatic Identification Systems: AIS Transponders*. http://www.imo.org/ourwork/safety/navigation/pages/ais.aspx. Accessed June 2012.

6. Perez, H.M., Chang, R., Billings, R. and Kosub, T.L. Automatic Identification Systems (AIS) Data Use in Marine Vessel Emission Estimation. Presented at the 18th Annual International Emission Inventory Conference. Baltimore, 2009.

7. U.S. Army Corps of Engineers (USACE) Navigation Data Center (NDC). *National Waterway Network (NWN)*. Website: http://www.ndc.iwr.usace.army.mil/data/datanwn.htm. Accessed July 2012.

8. U.S. Army Corps of Engineers (USACE) Navigation Data Center (NDC). *Waterborne Commerce of the United States*. Website: http://www.ndc.iwr.usace.army.mil//wcsc/wcsc.htm. Accessed May 2012.

9. U.S. Army Corps of Engineers (USACE) Navigation Data Center (NDC). *Lock Use, Performance and Characteristics*. Website: http://www.ndc.iwr.usace.army.mil/lpms/lpms.htm. Accessed May 2012.

10. U.S. Army Corps of Engineers (USACE) Navigation Data Center (NDC) – Lock Performance Monitoring System. *Corpslocks Information System*. Website: http://corpslocks.usace.army.mil. Accessed June 2012.

11. U.S. Army Corps of Engineers (USACE) Navigation Data Center (NDC) – *Ports and Waterways Facilities Dataset*. Website: http://www.ndc.iwr.usace.army.mil//ports/ports.asp. Accessed June 2012.