GIS-based identification and visualization of multimodal freight transportation catchment areas

Magdalena I. Asborn¹ · Sarah Hernandez² · Manzi Yves²

Accepted: 30 November 2020 / Published online: 2 January 2021
© The Author(s) 2021

Abstract

To estimate impacts, support cost–benefit analyses, and enable project prioritization, it is necessary to identify the area of influence of a transportation infrastructure project. For freight related projects, like ports, state-of-the-practice methods to estimate such areas ignore complex interactions among multimodal supply chains and can be improved by examining the multimodal trips made to and from the facility. While travel demand models estimate multimodal trips, they may not contain robust depictions of water and rail, and do not provide direct observation. Project-specific data including local traffic counts and surveys can be expensive and subjective. This work develops a systematic, objective methodology to identify multimodal “freight-shed” (or “catchment” areas) for a facility from vehicle tracking data and demonstrates application with a case study involving diverse freight port terminals. Observed truck Global Positioning System and maritime Automatic Identification System data are subjected to robust pre-processing algorithms to handle noise, cluster stops, assign data points to the network (map-matching), and address spatial and temporal conflation. The method is applied to 43 port terminals on the Arkansas River to estimate vehicle miles and hours travelled, origin, destination, and pass-through zones, and areas of modal overlap within the catchment areas. Case studies show that the state-of-the-practice 100-mile diameter influence areas include between 15 and 34% of the multimodal freight-shed areas mined from vehicle tracking data, demonstrating that adoption of an arbitrary radial area for different ports would lead to inaccurate estimates of project benefits.

Keywords Multimodal freight · Project prioritization · Economic impacts · Catchment areas · Geographic information system (GIS)

✉ Magdalena I. Asborn
masborn@ara.com

Sarah Hernandez
sarahvh@uark.edu

Manzi Yves
ym008@email.uark.edu

¹ Applied Research Associates, 119 Monument Pl, Vicksburg, MS 39180, USA

² Department of Civil Engineering, University of Arkansas, 4190 Bell Engineering Center, Fayetteville, AR 72701, USA

Introduction

In the context of public and private transportation infrastructure investment, projects compete for limited funding and resources. Thus, evaluation and prioritization of competing projects is critical, and of the methods available, all rely in part on estimation of benefits relative to costs. Calculation of benefits necessitates detailed estimation of the types and magnitude of project impacts associated with the project's area of influence (Hurtado et al. 2016; Weisbord and Reno 2009). Therefore, it is important to clearly and consistently define the extent, location, and characteristics of a project's area of influence.

The area of influence of a project affecting a multimodal freight facility can be defined as the region where the facility draws and delivers freight, or the connected origin–destination (OD) pairs served by the facility (Vadali et al. 2017). Despite general agreement on conceptual and qualitative identification of a project's economic area of influence (National Academy of Sciences, Engineering, and Medicine 2014; Vadali et al. 2017), little published work is available to define such areas for freight projects using data-driven approaches. Instead, planning agencies within state Departments of Transportation (DOTs) and at Metropolitan Planning Organizations (MPOs) rely on professional judgement to define the extent, coverage, and location of a project's area of influence (AASHTO 2015). As a result, there is a lack of a systematic method to determine areas of influence for competing projects. This leads to unfair competition among projects from different agencies and jurisdictions which may each use different and subjective means to define areas to measure project benefits. To ensure fair comparisons, the location, coverage and extent of the area of influence of different projects should be determined by following the same, systematic methodology.

State-of-the practice methods to identify the area of influence of a project consist primarily of selecting an arbitrary radial perimeter around the facility (Institute of Transportation Engineers 2010; NADO Research Foundation 2011; Carroll et al. 2017). For freight related projects, however, this method ignores complex interactions among freight supply chain components (e.g. truck, rail, water) and longer travel distances. Instead, the area of influence of a project affecting a freight facility can be better defined as the coverage and extent of freight movements generated by the facility (where “generation” is considered as production and attraction of goods), or the OD pairs and paths served by the freight modes using the facility. This has been referred to in this paper as the “freight-shed” or “catchment” area, borrowing from transit studies (Andersen and Landex 2008). Freight OD pairs may be obtained from project-specific data like stakeholder surveys or traffic counts and/or statewide travel demand models (TDM). Project-specific data can be time consuming to collect, expensive, and subjective. In the case of annual traffic counts, they may not be available at or near the project and do not provide data on freight paths. If a statewide freight TDM exists, the OD pairs and paths served by a freight facility might be estimated by performing a ‘select link analysis’ (Alliance Transportation Group and Cambridge Systematics 2015). However, although TDMs contain representative models of the roadway network, they often do not provide robust depictions of water and rail networks and are thus unsuitable for multimodal freight catchment analyses (Alliance Transportation Group and Cambridge Systematics 2015; Donnelly et al. 2018). For example, the statewide TDM in Arkansas (ARSTDM) contains a multimodal mode choice model, but only performs trip assignment for highway flows and not vessel flows because waterways are not part of the model network (Alliance Transportation Group and Cambridge Systematics 2015). This is a notable

limitation considering the key benefits of the Arkansas River to the state economy (Nachtmann 2015). In addition, network representation in TDMs often lacks the level of detail necessary to represent actual roadway geometry, i.e. port access roads may not be represented. Moreover, TDMs are predictive/forecasting tools, and do not constitute direct observation of existing flows. In contrast, the systematic method developed in this paper mines direct observations of truck and vessel tracking data to provide multimodal movements and patterns to and from freight facilities.

The lack of guidance regarding multimodal catchment area definitions can be attributed in part to the heterogeneity of the data used for this purpose. In a freight supply chain, the catchment area contains several modes, freight facilities, and industries, which would be better represented (and linked together) by spatially and temporally continuous data, such as historical truck and vessel paths. To overcome the limitations in defining catchment areas for freight facilities, this paper leverages two sources of “big data” to identify multimodal freight paths: truck Global Position System (GPS) and marine Automatic Identification System (AIS) data. The main contribution of this paper is a novel, automated geospatial data fusion method to identify the multimodal freight-shed of a project using ubiquitous vehicle tracking data. Conceptually, multimodal freight tracking data is used to characterize spatial patterns of freight intensity. In particular, GPS tracking data from trucks and marine vessels accessing freight facilities are mined to identify stops and to find complete paths which are then mapped to a high resolution, multimodal transportation network. The geographical coverage of the truck and vessel trips define the “multimodal catchment” area of the project. In this way, all competing projects are subject to identical criteria for catchment area definition, providing a common basis for funding prioritization. For new projects, where TDMs are relied upon for project evaluation, the methods in this paper provide a means to calibrate and validate TDMs so that they produce more accurate forecasts of multimodal catchment areas.

The methodology is applied to freight ports located on inland navigable waterways, although it can be applied to other infrastructure including bridges, railyards, and warehouses. Beyond benefit-costs analyses, quantitative definitions of multimodal catchment areas further support efforts to: (1) quantify multimodal performance measures, (2) visualize the extent of transportation impacts of extreme weather events (such as flooding), (3) estimate population exposures to pollutants or congestion effects induced by freight facilities by analyzing vehicle-miles and vehicle-hours travelled per mode, (4) identify areas of potential modal competition, and (5) support travel demand models which rely heavily on directly observed data for calibration.

Background

This manuscript develops a systematic, data-driven methodology to identify multimodal freight-sheds (or “catchment” areas) for a facility from vehicle tracking data, and demonstrates its applicability with a case study involving diverse freight port terminals. Thus, this section reviews international literature focusing on: catchment areas within the context of project evaluation and prioritization (“[Project evaluation, prioritization, and catchment areas](#)” section), vehicle tracking data (“[Vehicle tracking data characteristics](#)” section), and multimodal approaches to combine freight tracking data for different modes (“[AIS and truck GPS data fusion](#)” section).

Project evaluation, prioritization, and catchment areas

To mitigate negative externalities of projected freight growth, physical and operational improvements to the multimodal transportation network and freight facilities are required. Three analytic methods are typically utilized to prioritize, compare, and select transportation projects: (1) benefit–cost analysis (BCA), (ii) economic impact analysis (EIA), and (iii) multi-criteria analysis (MCA) (National Academy of Sciences, Engineering, and Medicine 2014). BCA consists of quantifying project impacts as monetary units and distributing them over time to calculate the present value of all benefits and costs. The results are expressed as a net benefit (benefit minus cost), or as a benefit/cost ratio. A broader version of the BCA accounts for social impacts, including environmental impacts that affect non-travelers. In the EIA, project impacts are measured in terms of their effect on a region's economy. Quantitative measures include business output, job generation, net business income generation, household income, and GDP. In MCA, impacts can be measured either as quantitative indices or as qualitative ratings to portray relative importance. Thus, a broader range of positive and negative impacts may be considered for decision-making (National Academy of Sciences, Engineering, and Medicine 2014).

All three methods require the area of influence of the project to be defined. For instance, according to the Guide for Conducting Benefit–Cost Analyses (BCA) of Multimodal, Multijurisdictional Freight Corridor Investments (Vadali et al. 2017), the first step of a BCA is to define a project by: (1) the type of facility or location to be analyzed (whether it is a corridor, a modal or intermodal facility), (2) its area of influence, (3) the modes involved, and (4) the nodes involved (i.e. connections to freight network points such as ports, distribution centers, etc.). However, no methodology is proposed to identify areas of influence. Moreover, the Guide highlights the need to integrate data sources to perform BCA, because individual data sources do not address all modes (Vadali et al. 2017). Similarly, the EconWorks economic impact assessment tool allows transportation agencies to estimate economic impacts of diverse project types using past projects as case studies (AASHTO 2015). For each case study, economic impacts are measured in terms of number of jobs, sales, income, and investment. The size and location of the area where the economic impacts are calculated is critical for the study, and described as the “counties in which the project passes, or which are immediately impacted by the project” (AASHTO 2015). However, the selection of the area of influence is not a data-driven analysis; instead, it is based on professional judgement. Data-driven methodologies, such as the one developed in our paper, are objective; while professional judgement is subjective. Subjectivity leads to a higher degree of potential variability (and bias). Variability leads to inconsistencies in evaluating benefits of potential investments on diverse facilities that will compete for the same funds. For that reason, and considering the need to develop a systematic methodology applicable to diverse facilities with comparable results, the utilization of an objective, data-driven methodology like the one developed in this paper constitutes an improvement over subjective professional judgement. Also, considering the high rate of retirements at state DOTs and in the profession, the qualified assessments of seasoned planners and engineers may no longer be available (National Transportation Workforce Summit Summary of Results: Framework for Actions 2012). Thus, this work can fill a critical knowledge gap attributed to workforce retention issues.

Another BCA tool, the Freight module within the BCA tool Transportation Economic Development Impact System (TREDIS n.d.), enables users to define a project and identify affected freight flows and associated economic activities. Given a user-defined region,

TREDIS profiles the area's freight flow patterns, assesses the supply chain roles of those freight flows, and calculates how emerging economic trends may change future freight flows and investment needs. As output TREDIS allows visualization of county-level, not link or corridor level, freight flows. Arguably the most powerful tool available on the market, TREDIS does not provide network-based analyses and visualization of project-specific freight catchment areas.

Vehicle tracking data characteristics

Multimodal catchment areas defined in this paper are based on freight vehicle tracking data, specifically maritime AIS and truck GPS (Table 1). Both sources cover wide geographies, contain population-level data or exist as large samples, and are publicly available, either directly from government sources or through data sharing agreements with private data providers, making them viable and promising sources for data-driven catchment area identification as described in this paper. A statistical analysis of the datasets used in this work are presented in “[Case study: Catchment areas of port terminals on the Arkansas River](#)” section (Case Study).

Automatic identification system data (AIS)

AIS data (Table 1) is collected for navigational safety purposes (e.g., collision avoidance) and is required by the International Maritime Organization (IMO) for all passenger-carrying vessels and commercial vessels over 300 Gross Tonnage that travel internationally. Onboard navigation devices transmit location and characteristics of vessels in real time to receivers on shore, satellite, buoy, and other vessels (U.S. Coast Guard n.d.). In the U.S., AIS is mandatory along the Ohio River, between Mileposts 593 and 606, and in the Lower Mississippi River, up to Milepost 254.5 (Dobbins and Langsdon 2013). Even though AIS is not required in all U.S. inland waterways, most vessels use the AIS transponder (DiJoseph and Mitchell 2015). Vessel and voyage features entered to the database manually contain substantial errors and omissions. Broadcasting features, e.g., location (latitude and longitude), time stamp, speed over ground, course over ground, and heading, do not require manual intervention, thus contain few errors, and are used in this work for catchment area definition. In particular, AIS data has the ability to track a vessel's path with time stamps which is suitable to identify freight flows through inland navigable waterways. Although previous studies reconstructed vessel trajectories from AIS data (Zhang et al. 2018; Zhao et al. 2018; Graser 2019), they are limited in the lack of an inland waterways network setting, or in that movements are divided per day, masking the identification of trips. DiJoseph and Mitchell (2015) overcome the latter by linking consecutive AIS records together to generate paths on inland waterways; however, they did not fuse generated vessel paths with a defined network. The inability to map vessel data to a network precludes future integration of AIS data into multimodal, network-based models, such as state TDMs. In contrast, the algorithm applied in this work allows for the identification of trips defined by origin and destination (not duration), and matched to a defined intermodal network.

Truck GPS data

Truck GPS data consists of vehicle positioning data (latitude and longitude) broadcasted by onboard GPS devices. Spatial coverage in the US is almost ubiquitous (Short 2014).

Table 1 Vehicle tracking data characteristics

Dataset characteristics		Waterborne AIS data	Truck GPS data
Data collection in the U.S		U.S. Coast Guard	Private data collection entities such as truck preclearance programs and Electronic Logging Device (ELD) providers
Data elements		<i>Vessel</i> : Vessel name, length, width, MMSI, IMO, and call sign <i>Broadcasting</i> : Latitude, longitude, time stamp, speed over ground, course over ground, heading <i>Voyage</i> : Cargo, draft, status	Anonymous truck identifier, position (latitude and longitude), timestamp, heading, spot speed
Spatial coverage		All international and U.S. waterways	All U.S. territory
Temporal coverage/record frequency		Aggregated to 60-s intervals (for data storage and sharing purposes)	Every 30 s
Update frequency		Since 2009, data collected in real time, shared via annual updates	Since 2002, data collected in real time
Data storage format		File geodatabases (.csv) containing one month of data per time zone	Text files (.txt) containing a time window requested by the user
Data sharing scheme		Open source via www.MarineCadastre.gov	Publicly available through agreements with data collection entities
Limitations		Carried by tugs and tows on inland waterways, not barges	May lack representability of industries and small fleets

Private truck fleets typically record positioning data of their own trucks for security, route tracking, fuel cost, and other operational analyses. Data providers typically share anonymous (no identification of industry, operator, company, etc.) truck GPS data gathered from a sample of private fleets. Truck GPS data has been used for bottleneck identification, travel time analyses, border crossings, truck parking, hours of services tracking, etc., and is a valuable source of truck routing, destination analysis in urban settings, time-of-day usage, volume and speed data (Short 2014; Laranjeiro et al. 2019; Sharman and Roorda 2011). Truck GPS data covers every single road in the statewide network, while other truck data sources, such as static sensors like Weigh-in-Motion (WIM), inductive loop detectors, or temporary tube counters, are restricted to fixed and few locations. Like AIS data, it is necessary to employ geospatial fusion methods to map GPS traces to a defined transportation network. Methods for map matching and route identification for truck GPS data have been carried out in several prior studies (Camargo et al. 2017; Hashemi and Karimi 2014; Ciscal-Terry et al. 2016).

AIS and truck GPS data fusion

The challenge in fusing truck and vessel tracking data is overcoming data heterogeneity in units of time, space, and context. Xu et al. (2017) developed a Generic Target Monitoring System (GTMS) to monitor multimodal vehicles, and tested it with AIS and truck GPS data collected at a seaport terminal. To overcome multimodal data heterogeneity, vehicle tracking data from different sources (i.e. truck, vessel) was converted to a uniform data format. A GIS web-based interface allowed users to visualize and analyze real-time and historical multimodal vehicle tracking data within a designated geographical area (Xu et al. 2017). Meyer-Larsen et al. (2015) combined real-time AIS and truck GPS data to improve the efficiency of logistics at container terminals. The system tracked container vessels positions from AIS data to estimate vessels estimated time of arrival (ETA) and compared it with the ETA manually entered by the vessel operator. The system automatically detected deviations between planned and scheduled ETA and communicated potential deviations in real time to port stakeholders (including truck operators), so they could schedule operations in response to vessels' delays (Meyer-Larsen et al. 2015). Monsreal et al. (2019) performed statistical analyses to determine vessels and truck activity correlations and causalities at coastal ports using AIS data from non-liquid carrying vessels acquired from a vendor, and truck GPS probe data from the National Performance Management Research Data Set (NPMRDS). The analysis was complemented with census and port administration datasets. The analyses produced coefficients representing changes in directional road traffic volumes corresponding to changes in import/export freight volume (measured in weight), and the time when those increments on road traffic were expected. For example, unloading of a vessel with 1000 TEU would increase traffic along an inland highway by approximately 500 trucks during the week the vessel arrives, and decrease by approximately 400 trucks two weeks later (Monsreal et al. 2019).

Overall, these studies were limited by the lack of: (1) a systematic, data-driven procedure to identify multimodal freight facility catchment areas, (2) network assignment procedures for AIS data, and (3) multimodal data fusion approaches applied to inland waterway transportation. In particular, the literature review suggests that previous attempts to combine AIS and truck GPS data were limited in their spatial and temporal scope to relatively small areas, such as port authorities, and time periods (i.e. less than a year), possibly due to the computationally expensive processes proposed on big data. The method presented in this paper overcomes

Fig. 1 Methodology flowchart

these limitations by characterizing spatial patterns of freight intensity that exceed the circular perimeter (radial buffer) typically used for catchment analysis by explicitly assigning truck and vessel flows to defined multimodal networks. Moreover, our methodology can be successfully applied to represent statewide multimodal annual freight activity from publicly available truck and vessel tracking data.

Methodology

The purpose of this paper is to develop a systematic, data-driven methodology to identify multimodal freight-shed (or “catchment” areas) for a project from publicly available vehicle tracking data. The “catchment” area of a multimodal freight facility is defined as the region where the facility draws and delivers freight, which can be visualized as the paths followed by vehicles, vessels, railcars, etc. accessing the facility. The methodology is organized in three steps: (1) data preparation, (2) data analysis, and (3) multimodal data fusion visualization and quantification (Fig. 1). These steps are automated using open source software. The method to prepare, process, and analyze vessel and vehicle tracking data, and to identify the multimodal freight catchment areas corresponding to each freight facility, was coded in Python 2.7. The visualization of catchment areas was modeled with the ModelBuilder module of geoprocessing software (Quantum GIS). Input data and results are stored in a PostgreSQL database administered via open source pgAdmin platform.

Step 1: Data preparation

Data preparation is necessary to: (1) reduce the AIS and GPS datasets, and (2) remove noise and deal with spatial and temporal inconsistencies of imperfect vessel and vehicle tracking data (quality control).

Data reduction

Data reduction is required to accelerate big data processing. Both GPS and AIS datasets are reduced by sampling location records (i.e., “pings” or latitude-longitude-timestamp points) with a frequency of 300 s as in (Pinjari et al. 2014), instead of the frequency of 30–60 s per mode provided with the unprocessed data (Table 1). In addition, for AIS data, AIS records with zero speed are removed, as in Fujino et al. (2018). Notably, this is not the speed s , calculated between two consecutive AIS records (Table 2), but the “point speed” reported for each record in the AIS database. Since records with zero point speed constitute a considerable portion of all AIS point location records (Osekowska et al. 2017), their removal considerably reduces computational time, without loss of representation of movement characteristics (Asborno and Hernandez 2020). By reducing the datasets, the density of location records (i.e. number of location records per area unit) decreases, and thus a broader geographical scope can be processed within a reasonable time.

Table 2 Consistency and relevancy (CR) algorithm parameters

Parameter/Description	CR threshold criteria	CR threshold values	
		Maritime AIS	Truck GPS
<i>s</i> Calculated speed for each pair of timewise consecutive records per vehicle/vessel	Reasonable maximum operational speed (El-Reedy 2012)	> 17.2 mph	> 81 mph
<i>a</i> Calculated acceleration rate for each pair of timewise consecutive records per vehicle	85th percentile of the average acceleration rate of heavy trucks	Not applicable	> 2.24 ft/s ²
<i>d</i> Euclidean distance of each position record from the transportation network	> 2 standard deviations of river width	Variable, based on river width	Not applicable
<i>C</i> Geographical coverage of all timestamped location records per vehicle/vessel (Length of the diagonal of the rectangular bounding box surrounding all records per vehicle/vessel)	Minimum distance between port authorities in the study area	< 1.2 miles	
<i>P</i> Number of timestamped location records per vehicle/vessel	As recommended by (Camargo et al. 2017)	< 20	

Data quality control

The purpose of this step is to remove low-quality records from the AIS and truck GPS datasets. Unprocessed AIS and anonymous truck GPS data may contain erroneous or irrelevant records due to transponder issues and mishandling, transmission obstructions, etc. Erroneous GPS records refer to records with unusual high speed (i.e. cargo vessels travelling at 70 mph), or located far from the transportation network (i.e. vessels far away from inland waterways or heavy trucks away from roadways). Irrelevant records come from vehicles that broadcasted less than 20 records within the reporting period. These records are removed by applying a consistency and relevancy algorithm (CR) to each of the datasets, following (Akter et al. 2018) for trucks and (Asborno and Hernandez 2020) for vessels. First, records without timestamp, latitude or longitude are removed from the dataset. Records with timestamp inconsistent with the temporal scope of the dataset are considered erroneous and thus discarded as well. Next, for each vehicle (or vessel), the CR algorithm iterates over the position records ordered by timestamp and evaluates them against the parameters in Table 2. Records below (or above) the threshold values in Table 2 are removed from the datasets. For example, if the calculated speed between two time-wise consecutive truck GPS records is more than 81mph, such records are removed. The remaining records are used for data analysis.

Step 2—Data analysis: stop identification and map matching

The purpose of this step is to reconstruct the vehicle and vessel paths observed from the GPS/AIS position data using mode-specific networks, also known as map-matching. Since the purpose of this manuscript is the systematic identification of multimodal freight catchment areas, the explanation of the map-matching algorithms are briefly explained here. The reader is directed to Akter et al. (2018) and Asborno and Hernandez (2020) for further details on the development of the map-matching algorithm for GPS and AIS data, respectively.

From the several map-matching algorithms available (Camargo et al. 2017; Hashemi and Karimi 2014), Camargo et al. is used because it has the advantage of wide applicability to multimodal data sets.¹ Briefly, the algorithm first identifies stops made by each vehicle. Then, stops are connected using shortest path algorithms to reconstruct vehicle trips as complete series of connected links. Akter et al. (2018) adapted the algorithm to statewide truck GPS samples, while Asborno and Hernandez (2020) adapted it to vessel movements on inland navigable waterways. The algorithm adaptation and parameters differ for each mode due to network and operational characteristics inherent to each mode. In particular, the roadway network typically presents several alternative routes to connect an origin to a destination, while the waterway network does not. As for operational characteristics, different speeds and acceleration rates associated to trucks and vessels play a role in the set-up of algorithm parameters. The following paragraphs summarize the stop identification and map-matching algorithm, highlighting the differences in its adaptation to each mode.

¹ Camargo's algorithm is written in Python 2.7 and openly available at https://github.com/pedrocamargo/map_matching.

Stop identification

Stops made by each vehicle or vessel are identified by iterating through temporally consecutive location records. A naïve approach to find stops would be to locate all zero speed records. However, for reduced datasets (such as the AIS sample, where zero-speed records have been removed), this is not an option. In datasets including zero-speed records (such as the GPS sample), assuming that a vehicle broadcasts a signal every few seconds even when stopped, several consecutive records with low or zero speed likely represent a single stop and the naïve method would thus overcount stops and misrepresent stop duration. The stop identification algorithm instead defines stop clusters based on the following parameters: (1) calculated speed (“stopped speed”), (2) stop duration, and (3) geospatial stop coverage.

The stopped speed is the maximum calculated speed between two timewise-consecutive records that may indicate that the truck or vessel is stopped (in conjunction with the other parameters). In reality, when a vessel or truck is stopped, their speed is zero. However, GPS and AIS data contains two problematic issues: (a) “floating stops” and (b) stop clusters. For (a) even though the vehicle is stopped, the GPS transponder is still active and several very low speed (non-zero) data points are produced. Thus, the stop speed threshold should be greater than zero to detect floating stops. For (b), the vehicle/vessel makes many small maneuvers to position itself to stop. Thus, for one stop, there may be many low speed (non-zero) records before and after the actual stop. Both for (a) and (b), these non-zero, stop-related data points are irrelevant to defining the stop and to defining the trips/journey of the vessel. The stop identification algorithm clusters these low speed data points together with the zero speed data points. This is a form of data reduction and is necessary for computational efficiency.

The stop duration parameter captures the minimum time (in seconds) between two time-wise-consecutive records that indicates that the truck or vessel are possibly stopped. In the event a vehicle/vessel broadcasts a single record while stopped, the minimum stop duration must be set as the minimum sampling frequency, otherwise the stop would not be identified by the algorithm. The geospatial stop coverage parameter represents the length of the diagonal of a bounding box containing all consecutive-stopped records.

For each mode, the algorithm parameters are identified by running several scenarios of different parameter values, varying one at a time, and comparing the scenario accuracy (i.e., number of stops correctly identified by the algorithm) with “groundtruthed” vehicle and vessel data. While manual, the groundtruthing approach need only be performed if one wished to recalibrate algorithm parameters. Thus, the data pre-processing steps are automatic, requiring no manual processing.

For each mode, the stop identification algorithm outputs a list of stops made by all vehicles, indicating: anonymous vehicle/vessel identification number, a generated stop identification number, time when the stop occurred, and its location coordinates (longitude and latitude).

Map-matching and trip identification algorithm

For each vehicle/vessel, the map-matching algorithm reconstructs trip paths as series of connected network links. Due to low sample frequency, a truck/vessel can traverse many links between records. Therefore, it is necessary to reconstruct the path, rather than just assign each record to the closest link, as the later would result in disjoint paths. When applied to truck GPS data, the algorithm first identifies all links of the network that are

Table 3 Key performance indicators of each freight facility, measured from its catchment area

Key indicators	Units of measurement
Freight catchment area size	Acres
Vehicles Miles Travelled (VMT) $VMT_{multimodal} = VMT_{truck} + VMT_{vessel}$	Aggregated miles for all vehicles; Percentage of VMT per mode
Vehicles Hours Travelled (VHT) $VHT_{multimodal} = VHT_{truck} + VHT_{vessel}$	Aggregated hours for all vehicles; Per- centage of VHT per mode
Population within the area ^a	Number of individuals
Number of business registered within the area ^b	Number of businesses
Number and location of unique TAZ as origin or destination of trips to/from each facility	Number of TAZs
Number and location of unique TAZ within the paths of trips to/from each facility	Number of TAZs

^aCan be stratified by demographic characteristics like income, race/ethnicity, etc.

^bCan be stratified by commodity type transported (e.g., NAICS code)

likely used by the vehicle as it travels between identified stops. Each record is associated with a network link if its location falls within a pre-defined “geoprocessing buffer distance” from the link. Then, disjointed links along the path are connected by finding the shortest path between the disjointed links. For the AIS data the map-matching algorithm identifies vessel trips by computing the shortest path between each pair of consecutive stops, skipping the geoprocessing buffer distance and thus, gaining efficiency. The streamlined approach for AIS is possible because the marine network is highly simplified, typically consisting of a single path with few or no alternative paths.

The map-matching algorithm outputs a sequenced list of network nodes visited by each vehicle, the time when the vehicle arrived and left each node, and its associated network link.

Step 3—Multimodal data fusion: visualization and quantification

The purpose of this step is to generate a visual map depicting the trips made by all vehicles to/from a given freight facility on a multimodal network. First, a complete inventory of freight terminals within the study area is needed. An initial inventory of freight facilities can be gathered from public sources such as the US DOT which maintains an inventory of port terminals for the U.S. inland waterways (U.S. Department of Transportation and Bureau of Transportation Statistics 2019). Since national inventories may be out of date or inaccurate, the following approach adapted from Joubert and Axhausen (2013) is used.

First, facilities that serve freight are extracted from a national database. Then, a bounding box delimiting the study area is defined. Lastly, the results of the stop identification algorithm (Step 2) are used to supplement the initial list of freight facilities by identifying clusters of stops not associated with the initial freight facility inventory locations (i.e. stop clusters located outside a buffer area around each port location). These clusters constitute potential locations of freight facilities (i.e. port terminals, docks, or loading/unloading areas) that were not included in the initial inventory and can be verified using aerial imagery. For the case study of Arkansas, three facilities are added to the initial inventory. Each facility is then assigned to an existing or new network node.

Fig. 2 Study area: 43 Freight port terminals on Arkansas River. Labels represent names of Municipalities. (Color figure online)

To visualize the vehicle trips to and from the inventoried freight facilities, first, for each mode, the results of the map-matching algorithm are queried (based on the network node associated with each facility) to select all vehicle and vessel trips that accessed the facility. To add geometry for visualization purposes, the trip data is joined to the corresponding mode-specific transportation network based on network link attributes (e.g., link ID). This also allows for estimation of Vehicle Miles Travelled (VMT) and Vehicle Hours Travelled (VHT) as the network link attributes include distance and free flow travel time. As a result, for each freight facility, the geometry of vehicle and vessel trips to/from the freight facility is produced and includes length (miles), duration (hours), origin, and destination (network node, TAZ, and freight facility ID when applicable). For visualization purposes, an arbitrary shaded area is traced around the trip paths (network links). In future work, the researchers will explore more robust means of visualizing the paths including use of kernel density functions to depict stop and trip concentrations along multimodal paths. Lastly, the two modes are super-imposed to visualize the multimodal catchment area.

To complement visual depictions of freight catchment areas, key quantitative indicators (Table 3) are calculated for each mode and by combining all modes (e.g., multimodal). The indicators constitute performance of the freight activity associated with each facility. The catchment area size, population, number of business within the area, and location of unique Traffic Analysis Zones (TAZs) within the path and serving as the origin or destination of trips to/from each freight facility are derived using statistical packages and modeling tools in GIS platforms. The VMT and VHT corresponding to all trips to and from each freight facility are calculated by aggregating the trip length (in miles) and duration (in hours) for all the trips with origin or destination in the freight facility.

Fig. 3 Distribution of processed AIS records per vessel

Fig. 4 Distribution of processed GPS records per truck

Case study: Catchment areas of port terminals on the Arkansas River

Scope

The methodology was demonstrated by applying it to 43 freight port terminals located on the Arkansas River (Fig. 2) with AIS and truck GPS data samples from 2016. The Arkansas River is a 308-mile stretch of navigable waterway that plays a key role in the national economy by connecting the heartland of the U.S. to the international markets via the Mississippi River, and contributes to the national economy with \$4.5 billion in sales, 34,000 jobs, and \$168 million in taxes (Nachtmann 2015). The total area of Arkansas is 53,187 square miles.

Table 4 Stop identification and map-matching algorithm parameters

Delivery stop parameter / mode	Maritime	Roadway
Maximum Stopped speed s (km/h)	5.3	4.8
Minimum stop duration d (s)	300	300
Maximum stop coverage c (km)	5.0	0.3
Geoprocessing buffer size ($^{\circ}$)	Not applicable	0.001

Data preparation

AIS data

The un-processed AIS data in this case study consisted of 7,803,151 records, broadcasted every 5 min by a sample of 765 unique vessels observed along Arkansas waterways during 2016. This sample was extracted from the historic AIS dataset available for download at www.MarineCadastre.org (Office for Coastal Management 2018). Data reduction and quality control procedures removed 56% position records from the dataset, leaving 3,398,279 AIS records from 753 vessels for further data processing. The distribution of AIS processed records per vessel (Fig. 3) shows that the majority of vessels (21%) have less than 500 records. The average number of processed AIS records per vessel was 4513; the median was 2060; the standard deviation was 6436; and the range [23–46,378].

Truck GPS data

Truck GPS data acquired from a non-profit trucking industry research firm corresponds to 338,303,934 records broadcasted by approximately 307,942 unique trucks within a 10-mile buffer around the state of Arkansas, during four, nonconsecutive, 2-week periods (56 days in total) in 2016. The data spans all four quarters of the year (February, May, August, and November), allowing for a comprehensive representation of seasonal truck travel behavior (Sharman and Roorda 2011). Within Arkansas, the truck GPS data represents a sample of about 10% of the truck population, with minor variability across seasons and regions (Hernandez et al. 2019). The distribution of GPS processed records per vehicle (Fig. 4) shows that the majority of trucks (32%) have less than 100 records. The average number of processed GPS records per truck was 495; the median was 185; the standard deviation was 1039; and the range [21–23,264]. As the truck GPS data is a sample of the total truck population, temporal and volume expansion factors were implemented to estimate the annual population of trucks visiting each facility.

Other data

Information about the commodities handled by each port were gathered from the National Transportation Atlas Database (U.S. Department of Transportation and Bureau of Transportation Statistics 2019), supplemented by observation of storage areas on aerial imagery, and through a web-search of publicly available data about those facilities. Population data by census tract in Arkansas was obtained from the Census Bureau Topologically Integrated Geographic Encoding and Referencing (TIGER). Business location data was obtained from ESRI and consisted of a geocoded list of more than 200,000 establishments registered in

Fig. 5 Multimodal freight-shed (catchment areas) for freight port terminals located on the Arkansas River, 2016. Truck paths (red) were not available out-of-state. Green circles delimit an arbitrary 100-mile diameter area around each facility. (Color figure online)

Arkansas, including name, location, and North American Industry Classification System (NAICS) code, among others.

Methodology applied for data analysis and multimodal fusion

The methodology to identify the multimodal catchment areas was applied to the 43 freight port terminals on the Arkansas River, from the AIS and truck GPS data samples. First, the parameters adopted for the stop identification and map matching algorithms (Table 4) were obtained by comparing the number and location of stops and trips identified by the algorithm to control vehicles subjected to manual groundtruthing. This was the only “manual”

Fig. 6 Size of multimodal (purple) and mode-specific (dashed) freight-shed 2016. (Color figure online)

Fig. 7 Number of registered business within the multimodal freight-shed 2016. (Color figure online)

step in the process and is not required for model application, only for model calibration. Mode-specific algorithm parameters as described in “[Step 2—Data analysis: stop identification and map matching](#)” section, were identified independently for the GPS and AIS traces (Table 4) (Asborno and Hernandez 2020, Akter et al. 2018). Once the parameter values were found, the stop identification algorithm was applied to the statewide AIS and truck GPS datasets. Lastly, the results of the map-matching algorithm, namely trip paths made by trucks and vessels, were combined into a single map for multimodal visualization, and performance measures were derived.

Fig. 8 Population within the multimodal freight-shed 2016. (Color figure online)

Fig. 9 VMT and VHT of multimodal freight-shed 2016. (Color figure online)

Fig. 10 Unique TAZs as origin or destination and along trip paths to and from port terminals, 2016. (Color figure online)

Results

Four of the 43 ports in the case study were selected for detailed case study based on their commodity and location diversity. Commodities include diesel fuel (liquid bulk), food and farm, sand and other dry bulk quarry products, and steel structures and construction equipment. The freight-shed (Fig. 5) and corresponding quantitative measures (Table 3; Figs. 6, 7, 8, 9, 10) show the variability in scale and scope observed for each port with roadway

(red), maritime (blue), and multimodal (purple) catchment areas. For example, the catchment area for the port in North Little Rock (central Arkansas) reaches to the eastern and western borders of the state while the catchment area for the port in Van Buren reaches only to central Arkansas, about half the area of the North Little Rock port.

The multimodal catchment area represents the union of the two individual modes. Since the union of two sets is comprised of all elements in the two sets with common elements in the sets occurring only once, Fig. 6 is created by measuring the areas (in acres) shown in Fig. 5, and represented as horizontal dashes for roadway, vertical dashes for waterway, and grid dashes where the two single modes overlap. The height of each bar in Fig. 6 represents the multimodal metric corresponding to each port. A similar approach is followed to interpret Figs. 7 and 8, which represent the number of registered business, and population within the catchment areas, respectively, for the study year (2016). For example, Fig. 7 tells us that within the multimodal catchment area of the port in Pine Bluff shipping steel structures and construction equipment (Fig. 5b) there were 708 business registered in 2016; 82% of those businesses were within the path of both trucks and vessels (“overlap freightshed”); 88% of those business were located within the maritime catchment area; and 94% were located within the roadway catchment area.

The multimodal VHT and VMT associated to each facility (Fig. 9) is the aggregation of the VHT and VMT per mode. Figure 9 shows the percentages of multimodal VHT (and VMT, respectively) considering not only modes (red for truck, blue for vessels), but also the proportion that corresponds to trips made in corridors where both modes compete (diagonal dashes), or “overlap”. For example, for the port terminal in Van Buren (Fig. 5d), the multimodal VMT is 47,794 (Fig. 9); 46.9% of which occurs exclusively in the roadway catchment area (“roadway exclusively”, solid red), and 9.7% of which occurs in the maritime catchment area (“maritime exclusively”, solid blue). In these corridors, the two modes do not compete, i.e., the roadway and maritime catchment areas do not overlap (Fig. 5d). The remaining 42.3% of the multimodal VMT occurs on roadway network links that are located within the maritime catchment area (“roadway overlap”, dashed red); and 1.1% of which is observed in waterway network links within the roadway catchment area (“maritime overlap”, dashed blue). This information can be used to estimate and compare externalities, such as emissions generated by vessels and trucks in the region where both modes compete (discussed later in this paper).

The number and location of TAZs within the paths of multimodal trips to and from each freight facility (light purple) and comprising the origin and/or destination of trips (dark purple) (Fig. 10) differ by port. The TAZs used for this analysis correspond to the Arkansas State TDM. The 5849 zones within the state follow the boundaries of aggregated census blocks, while out-of-state zones match the Transearch data and Business Economical Area (BEA) districts, totaling 306 U.S. BEAs (without Arkansas) (Alliance Transportation Group and Cambridge Systematics 2015). Thus, TAZs within Arkansas represent areas much smaller than out-of-state TAZs.

Discussion and contributions

Catchment area comparisons and state-of-the-practice

The size and extent of each port terminal catchment area varies significantly by port (Figs. 5, 6), suggesting it would not be appropriate to adopt a generalized, arbitrary area.

Table 5 Comparison of state-of-the-practice 100-mile diameter arbitrary area with catchment areas as identified in this paper

Port terminal location/commodity	Within 100-mile diameter area		Within AIS/GPS defined catchment area	Percentage of multimodal catchment area length within 100-mile diameter (%)
	Number of AADTT stations/locks	Multimodal catchment area length miles (roadway/waterway)		
Pine Bluff/ Food and farm products	175/5	724 (592/132)	4846 (4530/316)	15
Pine Bluff/ Steel structures and construction equipment	185/5	467 (335/132)	1371 (1104/267)	34
Little Rock/ Wholesale diesel fuel	454/5	1634 (1509/125)	5039 (4570/469)	32
Van Buren/ Gravel, sand, and other dry bulk quarry products	200/4	413 (259/154)	2113 (1948/165)	20
Average	254/5	810 (674/136)	3342 (3038/304)	25

Fig. 11 Roadway network truck data availability with relation to assumed 100-mile diameter freight port catchment areas. (Color figure online)

The difference between the arbitrary radial areas and the multimodal catchment areas derived from vehicle tracking data (Fig. 5) indicate the extent of the freight activity that would be ignored if arbitrary radial areas were utilized to estimate port activity. In particular, for the four ports presented in the case study, the state-of-the-practice 100-mile diameter areas only include between 15 and 34% of the multimodal freight catchment areas mined from vehicle tracking data (Table 5).

In the absence of the methodology proposed in this paper, the identification of catchment areas associated with freight facilities would rely on data collected from traffic count stations, surveys, or professional judgment. As a comparison, for the four case study ports, we compare the traffic counts available via locks on the waterways and Average Annual Daily Truck Traffic (AADTT) classification stations on the roadway network to the multimodal catchment areas observed from GPS and AIS (Fig. 11). The static traffic count data has several limitations. First, its spatial coverage is not ubiquitous as compared to AIS and GPS data. For example, on average, there are only five locks and 254 AADTT classification stations within the 100-mile diameter areas (Table 5). Second, AADTT and lock traffic counts do not track vehicles and thus do not provide origin, destination, and path data, masking the extent of the multimodal transportation system that would be impacted by an event occurring at a freight facility, and vice-versa. In contrast, the methodology developed in this paper individualizes freight trips specifically related to the facility, relying on the ubiquitous coverage of AIS and truck GPS data.

The freight catchment areas developed in this paper could not be visualized by relying solely on surveys or static traffic data. Even though truck trip paths (and thus, truck-shed) may be visualized from the output of a travel demand model, such models are based, in large part, on survey data as well. Waterway trip paths cannot be visualized from travel demand models that do not represent the navigable waterway network. In addition, travel demand models represent predictions or estimations of freight movement, not actual observations. In this context, vehicle tracking data provides a preferred alternative to the outputs

of state travel demand models to analyze multimodal freight catchment areas for project evaluation and prioritization. To the benefit of travel demand models, the catchment areas and associated pre-processing procedures presented in this paper could be used to better calibrate and validate travel demand models. Moreover, the ability to identify and visualize the multimodal freight-shed of different ports (or any other piece of infrastructure) based on the same data, methodology, and criteria provides a common basis for proper comparison, competition of funds, and resource allocation.

Identification of corridors with parallel modal routes and policy implication

Identification of multimodal freight catchment areas at each port can be used to develop policies and drive investment. For freight corridor planning, it can be observed from the case study that different modes (maritime and roadway) run along parallel shipping lines. For example, the corridors connecting Pine Bluff with Little Rock (Fig. 5a) and Dardanelle (Fig. 5b) present parallel multimodal routes, for which there may be opportunities for mode shifts, given appropriate cargo handling environments at the ports and underlying transportation cost structures. Parallel multimodal routes are also observed in the performance metrics quantified by this analysis (Fig. 6). For each port, the bigger the “overlap” area (grid dashes), the longer the corridor(s) where the modes may compete (provided cargo type handling). Some cargo types are more favorable for a modal shift than other cargo types (Mommens et al. 2020). While vehicle tracking data does not provide information on the type of cargo transported, a commodity-based analysis would support further conclusions regarding modal competition.

Cost–benefit ratios can be derived more systematically by observing mode specific and multimodal catchment areas and then used to prioritize modal shift policies (shift cargo from truck to vessels). In doing this, estimates of system usage by mode may provide a means to more accurately reflect costs and benefits of a project. Benefits, for example, may be associated with reduce roadway maintenance costs and/or emissions by shifting cargo from truck to barge (Braham et al. 2017).

In particular, VHT and VMT metrics (Figs. 9, 10) can be used to estimate and compare the emissions generated by vessels and trucks associated with each facility, and in the regions where the two modes compete (i.e. “overlap”). Notably, while trucks and vessels have different cargo capacities, emissions and roadway maintenance savings can be derived from the portion of multimodal VMT where the two modes compete, assuming a coefficient of freight equivalency per mode of 58 large semi-trucks per one barge of 1500 ton capacity (Kentucky Association of Riverports n.d.). When comparing two facilities, the greater the VMT associated with modal competition, the greater the opportunity to shift cargo from truck to barge, meaning that more benefits may be realized from targeted modal shift policies. For example, by observing Fig. 9b we may conclude that the most suitable candidate to target modal shift policies is the port in Pine Bluff, because it has the biggest proportion of VMT overlap.

The ability to estimate benefits at this level of spatial detail facilitates fair prioritization of possibly competing policy initiatives and investments, e.g., which port to target mode shift investments like additional loading docks or other capacity and handling capabilities. This paper provides a systematic method to identify and measure multimodal VMT and VHT in areas of modal competition. Moreover, the use of the same data sources to evaluate facilities located state-wide constitutes a sound basis for a comparison. To complement this

Fig. 12 Sample of detailed paths to/from a freight facility. (Color figure online)

analysis, the authors are exploring methods to estimate commodity ton-miles transported by mode and port, which would provide information to more detailed cost–benefit analyses.

Insight to port activity by region

The number and location of unique TAZs along the freight-shed associated with each port (Fig. 10), derived from multimodal vehicle tracking data, can be used to support long-range transportation planning, such as scenario planning. Scenarios simulating disruption of business in those zones might impact economic activity at the ports, and vice-versa. For example, a severe weather event such as a flood affecting a port in Little Rock (Fig. 5c), located in the center of the state, may have an impact on freight flows observed as far as Northwest Arkansas, encompassing a total area of 10,500 thousand acres (Fig. 6). While an event affecting traffic flows in Northwest Arkansas, such as an accident at a highway/rail crossing, may have an impact on the economic activity of a port located as far as Little Rock.

Furthermore, the freight-shed visualizations show the usage of the inland waterways for local, short trip deliveries (Fig. 5d). These visualizations provide evidence on the use of the inland waterways for domestic shipping of steel structures and sand (Figs. 5b and d, respectively), and the key role of the Arkansas River in connecting the U.S. Midwest with international markets for shipping farm products (Fig. 5a).

For local analysis, drayage and last mile trips can be observed from AIS and GPS data using the pre-processing steps (map-matching and stop identification) presented in this paper. For example, Fig. 12 is an extract of Fig. 5, showing a sample of the results at a scale that makes possible to visualize the very last portion of truck trip paths in an urban setting. Last-mile vehicle movements may only be observed by map-matching vehicle tracking data to a high-resolution network.

Extensions of applications

The methodology presented in this paper can be applied to (1) any region with available AIS and GPS data, and (2) to any transportation infrastructure component for which geo-spatial location is available, such as bridges, intermodal connectors, storage and warehousing facilities, rail crossings, etc., and (3) mode specific datasets, e.g., rail tracking data, if available. Moreover, the methodology could be applied to any area of interest, such as parcels corresponding to a specific land use, the location of a specific industry (i.e. an inland

petrol refinery, a forestry industrial area, etc.). Overall, the novel multimodal freight data analysis constitutes a sound basis to characterize spatial patterns of freight intensity to/from specific land-use parcels. Although it is likely that the parameters for the pre-processing steps are transferable given the non-unique characteristics of the datasets and the study region, future work will evaluate the transferability of parameter values for other segments of inland waterways and state-level truck GPS data. One of the challenges (and limitations) of working with AIS and GPS data is the lack of reliable information about the commodity carried. Thus, the authors are developing a model to quantify trip cargo by commodity. Such model would complement the work in this paper by discriminating vessel and vehicle trip paths by cargo volumes and types of products. Future work will also identify multimodal freight catchment areas by season. Lastly, the availability of maritime AIS, truck GPS, and open source software makes the methodology proposed in this work transferable to any other region.

Conclusion

Vehicle tracking data, namely from marine AIS and truck GPS, provide ubiquitous and consistent sources needed to identify multimodal freight paths to and from freight facilities and other specific land-use parcels. The methodology defines multimodal catchment areas depicted as unique and overlapping road and waterway paths to more accurately assess the area of influence of a freight facility. By matching vehicle tracking data to mode-specific networks and selecting the trips with origin or destination within bounding boxes surrounding a freight facility, the resulting freight paths constitute the facility's catchment area, or "freight-shed". The data processing procedures developed in this paper overcome challenges associated with the use of "big data" for relatively large (statewide) areas through the development of data reduction, stop identification, and map matching algorithms.

In this paper, the freight-shed can be interpreted to be bi-directional as the methodology centers on the identification of complete freight paths for a given facility and not on facility specific data collection through surveys or professional judgement. Thus, events affecting the transportation infrastructure within the catchment area of the facility are assumed to influence the use of the freight facility and vice versa. For example, an extreme flood that closes a port can be attributed to changes in roadway or waterway vessel traffic while road or waterway closures cause by a flood that lead to traffic delays could be attributed to shipment delays at the port. This would aid in cost-benefit assessment of resiliency-oriented projects, for example. Moreover, the identification and visualization of the geographic extent of multimodal freight catchment areas can be used to estimate population exposure statistics, such as exposure to emissions, by super-imposing census and business locations within the catchment areas.

Within the context of transportation infrastructure investment, several projects compete for a limited amount of resources. Prioritization of projects is often based on an estimation of project benefits relative to costs. To evaluate project benefits, it is important to understand the extent, location, and characteristics of a project's area of influence, or "catchment" area, which can be defined as the region where the facility draws and delivers freight, or the OD pairs and paths served by the facility. However, little has been written regarding systematic methods to identify multimodal catchment areas. State-of-the-practice methods to identify the area of influence of a facility consist of arbitrarily selecting a radial perimeter around the facility, ignoring complex interactions among freight modes

and supply chains. Alternatively, freight paths to and from a facility may be obtained from project specific data like surveys, which are not always comparable among projects, and/or from travel demand models which have imbalanced or non-existent multimodal network representations. The main contribution of this paper is a geospatial data fusion method to identify the multimodal freight-shed of a project by using increasingly ubiquitous vehicle tracking data. In this way, all projects are evaluated using the same method to identify their areas of influence, providing a common basis for proper comparison and competition of funds.

A case study to illustrate the value of identifying multimodal freight catchment areas highlights the differing size and shape of port freight-sheds, further supporting limitations presented by the naïve assumption of radial areas for freight facilities. Examples of inland waterway ports in Arkansas show that modes compete on the same freight corridors, presenting an opportunity for mode shift. Since the AIS and Truck GPS data are increasingly available worldwide, the methodology has wide applicability to broad geographies and facility types.

Acknowledgements The authors thank the Maritime Transportation Research and Education Center (Mar-TREC), a tier 1 University Transportation Center by the U.S. Department of Transportation, for sponsoring this paper under “Combining truck and vessel tracking data to estimate performance of inland ports”; the Arkansas Department of Transportation for acquiring the truck GPS data under “Truck activity analysis using GPS data” (TRC1702); and Taslima Akter for processing truck GPS data.

Author contributions Conceptualization: Sarah Hernandez, Magdalena Asbornio; Methodology: Magdalena Asbornio; Formal analysis and investigation: Magdalena Asbornio, Manzi Yves; Writing—original draft preparation: Magdalena Asbornio; Writing—review and editing: Sarah Hernandez, Magdalena Asbornio; Funding acquisition, Resources, and Supervision: Sarah Hernandez. All authors read and approved the final manuscript.

Funding This material is based upon work supported by the U.S. Department of Transportation under Grant Award Number 69A3551747130. The work was conducted through the Maritime Transportation Research and Education Center. The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The work is funded, partially or entirely, by a grant from the U.S. Department of Transportation’s University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Availability of data and material Automatic Identification System (AIS) data is available at <https://marin.ecadastre.gov/ais/>; Truck GPS data was acquired from a vendor.

Code availability https://github.com/pedrocamargo/map_matching (Camargo et al. 2017).

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- AASHTO: EconWorks case study development training, August 5. <https://planningtools.transportation.org/425/module-5---case-study-data-needs-and-sources.html> (2015). Accessed 27 Nov 2018
- Akter, T., Hernandez, S., Corro-Diaz, K., Chi, N.: Leveraging open source GIS tools to determine freight activity patterns from anonymous GIS data. Edited by AASHTO GIS for Transportation Symposium, pp. 55–69 (2018).
- Alliance Transportation Group and Cambridge Systematics: Arkansas statewide travel demand model documentation (2015)
- Andersen, J.L.E., Landex, A.: Catchment areas for public transport. WIT Trans. Built. Environ. Urban Transp. XIV **101**, 175–184 (2008). <https://doi.org/10.2495/UT080171>
- Asborno, M., Hernandez, S.: AIS map matching for freight characterization on inland waterways—working paper. Freight Transportation Data Research Lab. <https://sites.uark.edu/sarahvh/> (2020)
- Braham, A., Hernandez, S., Kent, J., Asborno, M., Smith, S.: Locating transload facilities to ease highway congestion and safeguard the environment. ARDOT. https://www.arkansashighways.com/TRC/TRC%20Reports/TRC1608_Locating_Transload_Facilities_to_Ease_Highway_Congestion_and_Safeguard_the_Environment.pdf. (2017)
- Camargo, P., Hong, S., Livshits, V.: Expanding the uses of truck GPS data in freight modeling and planning activities. Transp. Res. Rec. **2646**, 68–76 (2017). <https://doi.org/10.3141/2646-08>
- Carroll, M., Carlyle, C., Seman, M.: The economic and fiscal impacts of development near DART stations. Dallas Area Rapid Transit. <https://www.dart.org/about/economicimpact.asp> (2017)
- Chacon Hurtado, D., Bardaka, E., Yang, R., Gkritza, K., Fricker, J.D.: Transportation and economic development: an assessment of the SHRP2 C11 tools and their synergies with economic simulation models for the evaluation of wider economic benefits of transportation projects. In: ASCE International Conference on Transportation & Development. Houston, TX, June 26–29. <https://slideplayer.com/slide/10888234/> (2016)
- Ciscal-Terry, W., Dell’Amico, M., Hadjidimitriou, N.S., Iori, M.: An analysis of drivers route choice behaviour using GPS data and optimal alternatives. J. Transp. Geogr. **51**, 119–129 (2016). <https://doi.org/10.1016/j.jtrangeo.2015.12.003>
- DiJoseph, P., Mitchell, K.: Estimating vessel travel time statistics for inland waterways with automatic identification system data. In: Transportation Research Board 94th Annual Meeting. Washington, DC. <https://trid.trb.org/view/1339419> (2015)
- Dobbins, J., Langsdon, L.: Use of data from automatic identification systems to generate inland waterway trip information. Transp. Res. Rec. **2330**, 73–79 (2013)
- Donnelly, R., Upton, W., Knudson, B.: Oregon’s transportation and land use model integration program: a retrospective. J. Transp. Land Use **11**(1), 19–30 (2018). <https://doi.org/10.5198/jtlu.2018.1210>
- El-Reedy, M.: Chapter 5: Fabrication and Installation. In *Offshore Structures*, pp. 293–381 (2012). <https://doi.org/10.1016/B978-0-12-385475-9.00005-5>. Accessed 20 Jan 2020
- Fujino, I., Claramunt, C., Boudraa, A-O.: Extracting courses of vessels from AIS data and real-time warning against off-course. In: International Conference on Big Data Research. Weihai, China. <https://doi.org/10.1145/3291801.3291823> (2018)
- Graser, A.: Movement data in GIS. 01 26. <https://anitagraser.com/2019/01/26/movement-data-in-gis-19-splitting-trajectories-by-date/> (2019)
- Hashemi, M., Karimi, H.: A critical review of real-time map-matching algorithms: current issues and future directions. Comput. Environ. Urban Syst. **48**, 153–165 (2014). <https://doi.org/10.1016/j.compenurb.2014.07.009>
- Hernandez, S., Williams, B., Williamson, M., Akter, T., Diaz-Corro, K.: Truck activity analysis using GPS data. https://cpb-us-e1.wpmucdn.com/sites.uark.edu/dist/e/220/files/2019/08/TRC1702-Final-Report_v8-FINAL.pdf (2019)
- Institute of Transportation Engineers: Traffic impact analyses for site development. Washington, DC (2010).
- Joubert, J., Axhausen, K.: A complex network approach to understand commercial vehicle movement. Transportation **40**(3), 729–750 (2013). <https://doi.org/10.1007/s11116-012-9439-0>
- Kentucky Association of Riverports: Water transport benefits. http://www.kentuckyriverports.com/water_transport_benefits/ (n.d.)
- Laranjeiro, P., Merchán, D., Godoy, L., Giannotti, M., Yoshizaki, H., Winkenbach, M., Cunha, C.: Using GPS data to explore speed patterns and temporal fluctuations in urban logistics: the case of São Paulo, Brazil. J. Transp. Geogr. **76**, 114–129 (2019). <https://doi.org/10.1016/j.jtrangeo.2019.03.003>
- Meyer-Larsen, N., Müller, R., Köhler, T.: Optimisation of intermodal transport using satellite-based services. Operational Excellence in Logistics and Supply Chains. vol 379. <https://hiel.org/publications/2015/optimisation-intermodal-transport-using-satellite-based-services> (2015)

- Mommens, K., van Lier, T., Macharis, C.: Multimodal choice possibilities for different cargo types: application to Belgium. *Res. Transp. Bus. Manag.* (2020). <https://doi.org/10.1016/j.rtbm.2020.100528>
- Monsreal, M., Kruse, J., Hun Kang, D., Park, E.S.: Developing and applying an analysis methodology to identify flow generation influences between vessel and truck shipments. 10.5281/zenodo.3165953 (2019)
- Nachtmann, H.: Regional economic impact study for the McClellan–Kerr Arkansas river navigation system. https://martrec.uark.edu/news-and-events/mkarns_for_web.pdf (2015)
- National Academy of Sciences, Engineering, and Medicine: Development of Tools for Assessing Wider Economic Benefits of Transportation. Washington, DC: The National Academy Press. <https://doi.org/10.17226/22502> (2014)
- National Association of Development Organizations Research Foundation: Inland ports: linking freight and regional economic development. <https://www.nado.org/inland-ports-linking-freight-and-regional-economic-development/> (2011)
- National Workforce Summit Summary of Results: Framework for action. National Transportation Workforce Summit, Council of University Transportation Centers. Washington, D.C., April 24–26, 2012.
- Office for Coastal Management: Vessel Traffic Data. 04. <https://marinecadastre.gov/ais/> (2018). Accessed 01 2019
- Osekowska, E., Johnson, H., Carlsson, B.: Maritime vessel traffic modeling in the context of concept drift. *Transp. Res. Proc.* **25**, 1457–1476 (2017). <https://doi.org/10.1016/j.trpro.2017.05.173>
- Pinjari, A. R., Zanjani, A. B., Thakur, A., Irmania, A. N., Kamali, M., Short, J., Pierce, D., Park, L.: Using truck fleet data in combination with other data sources for freight modeling and planning. Florida DOT. <https://www.fdot.gov/docs/default-source/research/reports/FDOT-BDK84-977-20-rpt.pdf> (2014)
- Sharman, B.W., Roorda, M.J.: Analysis of freight global positioning system data. *Transp. Res. Rec. J. Transp. Res. Board* **2246**, 83–91 (2011)
- Short, J.: Truck GPS data for tracking freight flows. Edited by Transportation Research Board. Developing Freight Fluidity Performance Measures, pp. 35–38. <http://onlinepubs.trb.org/onlinepubs/circulars/ec187.pdf> (2014)
- TREDIS: TREDIS—Products. <https://www.tredis.com/products/tredis-freight> (n.d.). Accessed 4 Dec 2018
- U.S. Coast Guard: Automatic identification system. <https://www.navcen.uscg.gov/?pageName=AIMSmain> (n.d.). Accessed 11 2019
- U.S. Department of Transportation and Bureau of Transportation Statistics: National transportation atlas database—ports. Washington, DC. <https://data-usdot.opendata.arcgis.com/datasets/ports> (2019)
- Vadali, S., James Kruse, C., Kuhn, K., Goodchild, A.: NCFRP report 38. Guide for conducting benefit-cost analyses of multimodal, multijurisdictional freight corridor investments. doi:<https://doi.org/10.17226/24680> (2017)
- Weisbord, G., Reno, A.: Economic impact of public transportation investment. American Public Transportation Association. <http://onlinepubs.trb.org/onlinepubs/tcrp/docs/TCRPJ-11Task7-FR.pdf> (2009)
- Xu, K., Zheng, H., Li, Y., Yue, L.: Comprehensive monitoring system for multiple vehicles and its modeling study. *Transp. Res. Proc.* **25**:1824–1833. <https://www.sciencedirect.com/science/article/pii/S2352146517304520> (2017)
- Zhang, L., Meng, Q., Xiao, Z., Xiuju, Fu.: A novel ship trajectory reconstruction approach using AIS data. *Ocean Eng.* **159**, 165–174 (2018). <https://doi.org/10.1016/j.oceaneng.2018.03.085>
- Zhao, L., Shi, G., Yang, J.: Ship trajectories pre-processing based on AIS data. *J. Navig.* **71**, 1210–1230 (2018). <https://doi.org/10.1017/S0373463318000188>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Magdalena I. Asborno received her PhD degree in Civil Engineering with a specialization in multimodal transportation planning and systems from the University of Arkansas (UofA), Fayetteville. She holds a M.S. from the Energy Superior Institute, Spain, and a B.S. in Civil Engineering from the University of La Plata, Argentina. In 2020, Dr. Asborno joined Applied Research Associates, Inc (ARA). As a Senior Consultant, Dr. Asborno manages projects for the US Army Corps of Engineers, focusing on waterborne freight and its interaction with the multimodal transportation network. Her research interests include data fusion, geographic information systems (GIS), machine learning tools, commodity-based freight network modeling, and project prioritization. Before joining ARA, Dr. Asborno added value to UofA's Freight Transportation Data Research Lab, and to the development and construction of public works world-wide for more than 15 years.

Sarah Hernandez received her PhD in Civil and Environmental Engineering with a specialization in transportation systems engineering from the University of California, Irvine. She holds a M.S. from the University of California, Irvine (Irvine, California) and a B.S. from the University of Florida (Gainesville, Florida). Dr. Hernandez joined the University of Arkansas Civil Engineering Department in August 2015. She teaches graduate classes on transportation planning and transportation data analysis. Dr. Hernandez is the faculty advisor for the student chapter of the Institute of Transportation Engineers (ITE) and a member of ITE. Her research focuses on new and advanced technology applications in transportation systems engineering and is centered on developing tools and methods to collect and analyze freight and commercial vehicle operations data for long range freight planning. She serves as a member of the Transportation Research Board's Standing Committee on Freight Transportation Data and is the Committee Research Coordinator for the Workforce Development and Organizational Excellence.

Manzi Yves is a researcher at the Freight Transportation Data Lab at the University of Arkansas. His research focuses on the examination of inland waterway vessel movement data with GIS tools to predict docking locations, supported by the Maritime Transportation Research and Education Center. He is currently pursuing his bachelor's degree in civil engineering. He is a member of the National Society of Black Engineers, and of the student chapter of the Institute of Transportation Engineers, where he volunteered to share his experience as an intern at the Arkansas Department of Transportation. He is the recipient of the Arkansas Academy of Civil Engineers scholarship, the Samuel T. and Mary I. Blair scholarship, and the Blanche Bledsoe Rosecrans & Clarence J. Rosecrans, Sr. Memorial scholarship.