Spatial disaggregation of waterborne commodity flow by fusing truck GPS and lock performance data

Magdalena I. Asborno, M.S.[1], Dr. Sarah V. Hernandez[2]

PhD Student - University of Arkansas [1], Assistant Professor - University of Arkansas [2]

TRB Innovations in Freight Data Workshop
Irvine, CA | April 2019

I. BACKGROUND, OBJECTIVES AND APPLICATIONS

- Inland waterway ports are critical modal connectors along the multimodal freight transportation system.
- State-of-the-practice means of gathering commodity flows for ports are limited in their spatial disaggregation, temporal continuity, and multi-modal integration.
- At best, the United States Army Corps of Engineers (USACE) Lock Performance Management System (LPMS) makes publicly available monthly tonnages for commodities observed at USACE maintained locks.
- Commodity flow through inland waterway ports are not immediately available to public officials, since there are several ports located within each pair of consecutive locks.

Objective

Development of a transferable model for spatial disaggregation to port-level of inland waterway commodity flow.

Applications

1. Estimation of commodity-specific, multi-modal freight fluidity performance measures.
2. Support to location selection for waterborne freight transload facilities

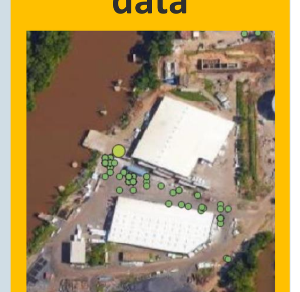
II. METHODOLOGY BRIEF & DATA PREPARATION

Methodology Brief

1. Spatial and temporal data fusion of USACE LPMS and ATRI Truck GPS data. Publically available LPMS data provides the required commodity dimension, while anonymous Truck GPS data provides spatial disaggregation.
2. Development of an optimization model that estimates the amount of each commodity transloaded at each port, to truck and rail.

Data Preparation

Truck GPS data



Sample anonymous GPS pings at a port

State-wide Truck GPS data sample: Number of trucks S on GPS dataset, per port i, and for sample period w ($S_{i,w}$)

Four 2-week sample periods: $S_{i,\,February}$, $S_{i,\,May}$, $S_{i,\,August}$, $S_{i,\,November}$

Truck Volume Coefficients*: $V_{v,Feb}\ V_{t,Q1}$, $V_{v,May}\ V_{t,Q2}$, $V_{v,Aug}\ V_{t,Q3}$, $V_{v,Nov}\ V_{t,Q4}$

$S'_{i,Q1} + S'_{i,Q2} + S'_{i,Q3} + S'_{i,Q4} = T_{i,Annual}$ Annual number of trucks per port

$V_{v,w}$: Spatial Coverage for sample period M, from Weigh-In-Motion data.
$V_{t,Q}$: Temporal Coverage extrapolation for quarter Q.

LPMS data

Lock & Dam

Tons of commodity j at locks L, for month t, per direction (U: upriver, D: downriver)

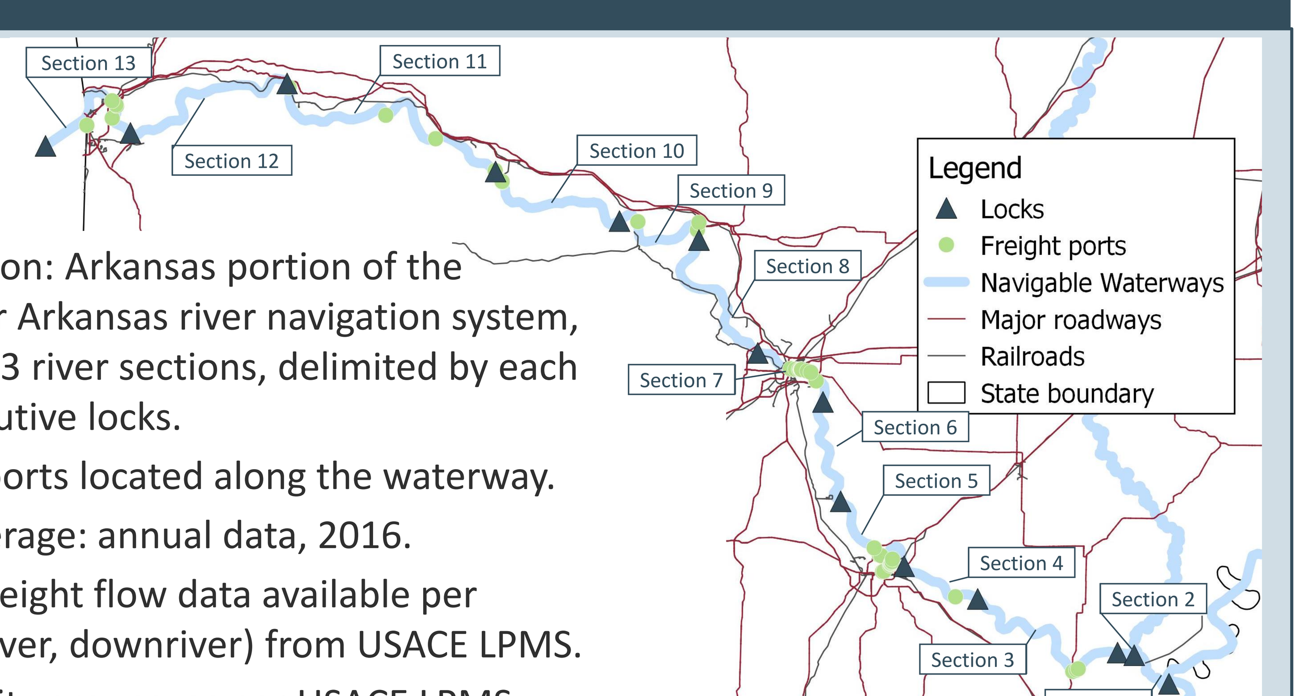
$\Delta L_{j,t} = |(L_2 - L_1)|_{D,j,t} + |(L_1 - L_2)|_{U,j,t}$ Difference in Tons of commodity j between consecutive locks L_1 and L_2, for month t.

$\frac{\Delta L_{j,t}}{f_j}$ Equivalent truckloads of difference of commodity j between consecutive locks L_1 and L_2, for month .

f_j = *Truck payload factor for commodity j*

$\sum_t \frac{\Delta L_{j,t}}{f_j} = \sum_t c_j^{s,t}$ Annual truckloads of change in commodity j between consecutive locks

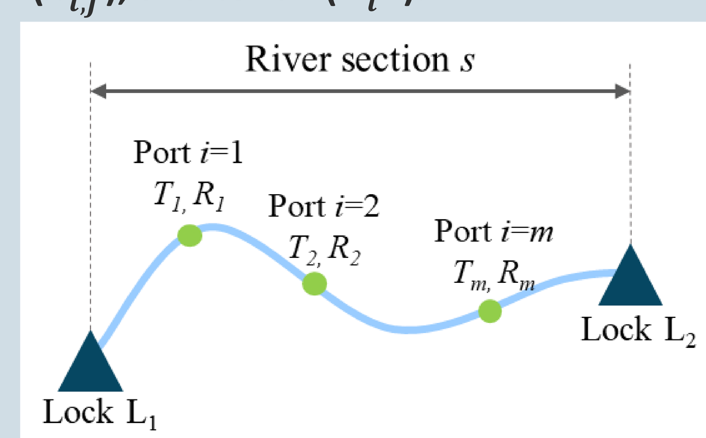
III. CASE STUDY: DATA & SCOPE OF WORK

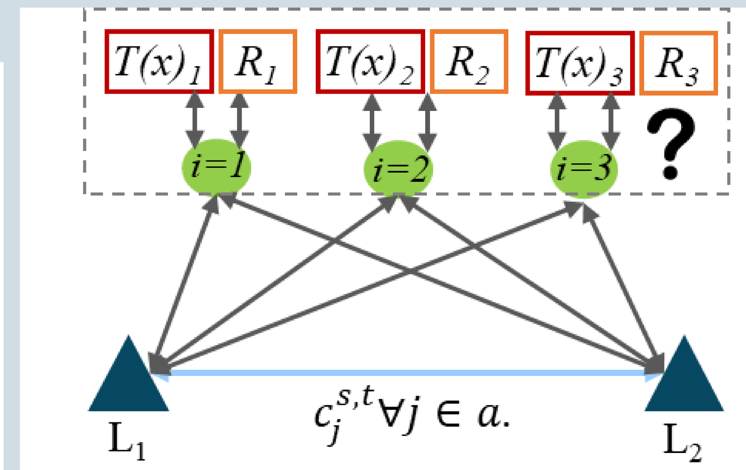


- Spatial definition: Arkansas portion of the McClellan-Kerr Arkansas river navigation system, consisting of 13 river sections, delimited by each pair of consecutive locks.
- There are 43 ports located along the waterway.
- Temporal coverage: annual data, 2016.
- Waterborne freight flow data available per direction (upriver, downriver) from USACE LPMS.
- Nine commodity groups, as per USACE LPMS

IV. MODEL DEVELOPMENT: A MULTI-COMMODITY ASSIGNMENT PROBLEM

- The river navigation system is broken down into as a series of river sections s. A river section s is defined as a stretch of navigable waterway located between each pair of consecutive locks (L_1, L_2).
- **The purpose is to identify, for a time window t and each river section s, the number of truckloads corresponding to each commodity j transloaded at each port i to truck ($x_{i,j}^{s,t}$), and to rail ($R_i^{s,t}$).**





Unknowns: $x_{i,j}, R_i\ \forall i \in s, \forall j \in a.$

- Problem formulation:

$$Minimize\ f(x,R) = \sum_t \sum_i (T(x)_i^{s,t} + \beta_i^s R_i^{s,t}) - \sum_t \sum_j c_j^{s,t} \quad \forall j \in a, \forall i \in s$$

Where:

$T(x)_i$ = freight transloaded from water to truck at port i.
R_i = freight transloaded from water to rail at port i.
β_i = coefficient to indicate whether port i has rail access.
c_j = Change on the amount of commodity j on river section s, during timeframe t.
t = temporal timeframe under analysis

Freight unit: equivalent truckloads.

Section	Port	Commodity $j = 1$	$j = 2$	$j = 3$	...	$j = n$	Truck	Rail	Row Subtotal
s	$i = 1$	$x^s_{1,1}$	$x^s_{1,2}$	$x^s_{1,3}$	...	$x^s_{1,n}$	$T(x)^s_1$	R^s_1	$(T(x)_1^s + \beta_1 R_1^s)$
s	$i = 2$	$x^s_{2,1}$	$x^s_{2,2}$	$x^s_{2,3}$	...	$x^s_{2,n}$	$T(x)^s_2$	R^s_2	$(T(x)_2^s + \beta_2 R_2^s)$
s	...	...	...	...	...	...	...	...	...
s	$i = m$	$x^s_{m,1}$	$x^s_{m,2}$	$x^s_{m,3}$	...	$x^s_{m,n}$	$T(x)^s_m$	R^s_m	$(T(x)_m^s + \beta_m R_m^s)$
Column Subtotal		c^s_1	c^s_2	c^s_3	...	c^s_n			$\sum_j c_j^s$

✓ The number of trucks T accessing each port i must match the sum of equivalent truckloads of all commodities transloaded at such port i (subtotal per row of C):

$$\sum_t T_i^{s,t} = \sum_t \sum_j \alpha_{i,j}^s x_{i,j}^{s,t} \quad \forall i \in s$$

- Conflictive Constraints:

✓ The sum of the amounts of commodity j transloaded between water and truck at all ports $i \in s$ must be smaller or equal to the amount of each individual commodity c_j observed in river section s (subtotal per column of C). The difference to match this inequality allows the model to assign commodities to Rail:

$$\sum_t c_j^{s,t} \geq \sum_t \sum_i \alpha_{i,j}^s x_{i,j}^{s,t} \quad \forall j \in a$$

✓ The percentage of trucks T accessing each port i with respect to the total trucks observed at river section s is:

$$\frac{\sum_t T_i^{s,t}}{\sum_t \sum_i T_i^{s,t}} \geq \frac{\sum_t \sum_j \alpha_{i,j}^s x_{i,j}^{s,t}}{\sum_t \sum_j c_j^{s,t} - \beta_i^s R_i^{s,t}} \quad \forall i \in s$$

✓ All values greater or equal than 0.

- Derived from Truck GPS data:

$\frac{\sum_t T_i^{s,t}}{\sum_t \sum_i T_i^{s,t}}$ $i=1...m,\ i \in s$	Percentage of trucks accessing port i with respect to all trucks accessing ports on section s.

- Derived from LPMS data:

$c_j^{s,t}$: Change in the amount c of commodity $j=1...n,\ j \in a$ (equivalent truckloads)

- Port data:

$\alpha_{i,j}^s = \begin{cases} 1\ if\ port\ i\ handles\ commodity\ j \\ 0\ if\ port\ i\ does\ not\ handle\ commodity\ j \end{cases}$	Commodities handled at each port.
$\beta_i^s = \begin{cases} 1\ if\ port\ i\ has\ access\ to\ the\ rail\ network \\ 0\ if\ port\ i\ does\ not\ have\ access\ to\ the\ rail\ natwork \end{cases}$	Whether a port has rail access or not.

V. CASE STUDY RESULTS

- Software utilized for model development and results: IBM ILOG CPLEX Optimization Studio 12.8.0.
- Example results for section s = 3 :

ANNUAL TRUCKLOADS - TOTAL Port_id	Commodity 10-COAL	20-PETROL	30-CHEM	40-CRMAT	50-MANU	60-FARM	70-MACH	80-WASTE	99-UNKWN	Truck	Rail	Result Truck%	Objective Truck%
3001	798	-	9,556	10,877	3,332	6,258	-	-	-	30,822		71%	74%
3002	-	-	-	-	-	11,328	-	-	-	11,328		26%	26%
3003	-	1,039	-	-	-	-	-	-	-	1,039		2%	0%
Result S3	798	1,039	9,556	10,877	3,332	17,587	-	-	-	43,190	657		
Objective S3	798	1,039	9,556	10,877	3,332	17,587	175	228	254	43,846			

- Results post processing allows for discrimination of freight flow directionality:

ANNUAL TRUCKLOADS - UPRIVER Port_id	Commodity 10-COAL	20-PETROL	30-CHEM	40-CRMAT	50-MANU	60-FARM	70-MACH	80-WASTE	99-UNKWN	Truck	Rail
3001	339	-	3,358	1,331	3,086	218	-	-	-	8,332	
3002	-	-	-	-	-	395	-	-	-	395	
3003	-	621	-	-	-	-	-	-	-	621	
Upriver S3	339	621	3,358	1,331	3,086	613	-	-	-	9,348	145

ANNUAL TRUCKLOADS - DOWNRIVER Port_id	Commodity 10-COAL	20-PETROL	30-CHEM	40-CRMAT	50-MANU	60-FARM	70-MACH	80-WASTE	99-UNKWN	Truck	Rail
3001	459	-	6,198	9,546	246	6,040	-	-	-	22,490	
3002	-	-	-	-	-	10,933	-	-	-	10,933	
3003	-	418	-	-	-	-	-	-	-	418	
Downriver S3	459	418	6,198	9,546	246	16,974	-	-	-	33,841	512

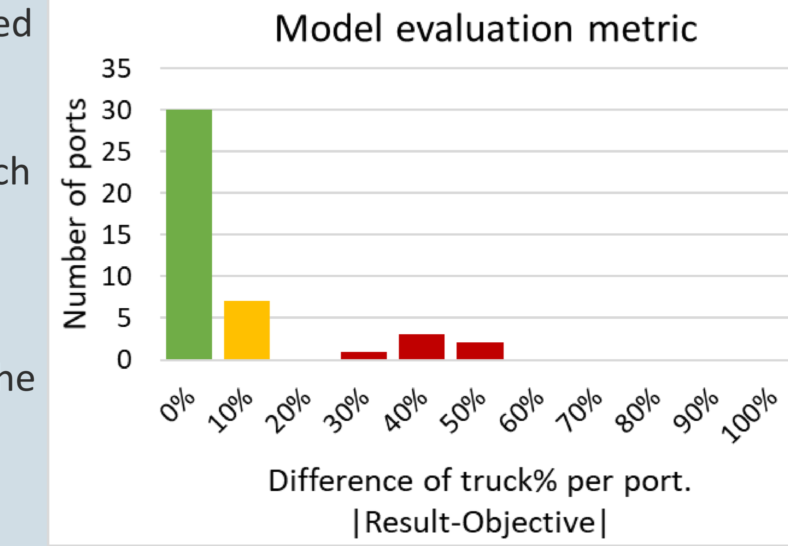
Commodity groups:

10-Coal, Lignite and Coke
20-Petroleum and Petroleum Products
30-Chemicals and Related Products
40-Crude Materials, Inedible, except Fuels
50-Primary Manufactured Goods
60-Food and Farm Products
70-All Manufactured Equipment and Machinery
80-Waste Material, Garbage, Landfill, Sewage Sludge and Waste Water
99-Commodity is Unknown

VI. DISCUSSION AND FUTURE WORK

- Multiple and conflictive conditions imposed to the optimization model cause the algorithm to find a solution with relaxed conditions. We utilize the relaxation of such conditions to evaluate the model.
- Model evaluation metric (EM): difference between the model results and the objective percentage of trucks per port. The lower EM, the better the model results.
- 86% of ports (37 out of 43) show EM < 20%.
- 75% of river sections (6 out of 8) show an average EM < 20%. Those 6 river sections gather 80% of the ports within the river navigation system.
- As for the amount of each commodity per section, the model results match 100% with the LPMS data.

Model evaluation metric



River sections	Number of ports	Average model evaluation metric per section
3; 4; 5; 7; 10	30	< 10%
13	6	< 20%
9; 11	7	< 40%
1;2; 6; 8	0	No ports. Algorithm not applicable.

- Limitations:
  - The proposed method is applicable to inland waterways with locks operated by USACE. USACE operates and maintains 291 Locks within the U.S.
  - Truck GPS data may not be representative of all industries or commodities. Representation is expected to improve in the future, as more companies are included in the sample frame.
- Future work: Adjust and apply the model to monthly results, identify paths of trucks observed at ports, incorporate AIS data, evaluate how much GPS data is needed.