MARITIME TRANSPORTATION RESEARCH AND EDUCATION CENTER
TIER 1 UNIVERSITY TRANSPORTATION CENTER
U.S. DEPARTMENT OF TRANSPORTATION

MarTREC
Maritime Transportation Research & Education Center

A Policy and Infrastructure Evaluation Model of Commodity Flows through Inland Waterway Ports

August 13, 2020 to August 13, 2022

PI: Sarah V. Hernandez, University of Arkansas
CO-PI: Sandra D. Eksioglu, University of Arkansas
Sanjeev Bhurtyal
Manzi Yves

August 2022

FINAL RESEARCH REPORT
Prepared for:
Maritime Transportation Research and Education Center

University of Arkansas
4190 Bell Engineering Center
Fayetteville, AR 72701
479-575-6021

ACKNOWLEDGEMENT

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 at the University of Arkansas.

Table of Contents

1 Project Description

1.1 Project Overview and Objectives

The purpose of this project is to guide strategic investment into port capacity through the development of a policy and infrastructure evaluation model of inland waterway commodity flows. This is accomplished by developing a two-stage stochastic optimization model to identify investment needs at the port level. The study builds on a growing body of research related to inland waterway port infrastructure investment. Quantifying capacity and throughput of inland waterway ports by commodity can identify opportunities for port infrastructure investment and economic development. Unfortunately, state-of-the-practice commodity flow data from shipper/carrier surveys don't provide port-level estimates. In this study, we develop a two-stage stochastic model that recommends port infrastructure investments and estimates commodity flow in such a way that the total of the investment cost and the expected transportation cost are minimized. These decisions are subject to uncertainty observed in the demand for different commodities shipped via water while considering rail and truck transportation costs and network availability. The use of a stochastic model, rather than a deterministic model, avoids underestimation or overestimation of the supply chain costs when uncertainly exists. This can eventually be used to direct and prioritize investment decisions, as well as develop effective transportation policies.

The specific objectives of this project are to:

I. **Develop a two-stage stochastic optimization model using publicly available data considering demand uncertainty, and**
II. **Apply the developed model to inland port terminals in Arkansas to prioritize port infrastructure investments following a specified budget.**

1.2 Motivation and Contribution

In relation to the objectives of the MarTREC research program, this project contributes to the areas of maritime and multimodal logistics management and infrastructure preservation by providing necessary data for effective (1) freight planning and travel demand modeling efforts and (2) mode shift analyses.

The proposed port throughput model is expected to close a critical gap in the ability of public sector decision makers to estimate port-level commodity flows and to evaluate policy and investment decisions regarding strategic investment in inland waterway ports. The resulting estimations of seasonal port-level commodity flows could be used to estimate freight-fluidity performance measures for project selection and prioritization by state and federal agencies (e.g., the Arkansas Department of Transportation (ARDOT)) and to highlight port facility improvement opportunities for private investors (e.g., regional economic development agencies or port operators).

As an example of an application of the work, consider the following scenario. Based on the current assignment model, relatively high amounts of agricultural products were found to be transloaded at ports along two sections of the McClellan-Kerr Arkansas River Navigation System (MKARNS), but there was only one port capable of handling such products. This finding suggests opportunities for investment in added commodity handling capabilities at select ports. Moreover, according to the US Department of Energy, Arkansas has 4.5 million tons of forest products and residues which can be used to generate renewable energy. But it is unclear if port infrastructure including grain elevators, silos, and other operational and storage equipment, along the MKARNS has the capacity to leverage this economic opportunity. By first using the proposed model to quantify port capacity as the tonnage by commodity type moved through each port during each month, it then becomes possible to evaluate opportunities for expanding capacity to accommodate future growth of key commodities like agriculture products and biomass, for example.

The research objectives highlighted above are in line with the marine transportation system priorities recommended by the U.S. Committee on the Marine Transportation System (CMTS). In particular, some of the recommendations highlighted by CMTS are: i) coordinate and apply big data analytics to reveal research gaps and overlap, foster potential collaboration, manage knowledge, and inform decision-making; ii) couple the newly-available vehicle probe data sets with more traditional freight data resources to quantify and contextualize travel times, dwell times, trip counts and other metrics; iii) create specific MTS system-scale performance indicators that relate to the freight flow network so they may be periodically updated and used for network calibration and validation; iv) develop and use decision support tools to identify nationally significant priority areas and project locations where agencies can leverage a variety of funding (U.S. Committee on the Marine Transportation System, 2022).

1.3 Scope

In the context of inland waterway transportation, more than 25,000 miles of U.S. inland waterways carry about 14% of all domestic freight, representing more than 600 million tons of cargo annually (American Society of Civil Engineers 2017). In particular, the methodologies developed for this project are applied to the Arkansas portion of the McClellan Kerr-Arkansas River Navigation System (MARNS), which consists of 308 miles of river, and contributes to the national economy with $4,535M in sales, $168M in business taxes, and 33,695 jobs (Nachtmann et al., 2015). Within the next 50 years, the net present value of sales, Gross Domestic Product (GDP), and generated taxes of the MKARNS are expected to be $232.5B, $111.3B, and $7.8B respectively (Oztanriseven et al., 2019). In 2017, the MKARNS transported 11.5M tons of goods, equivalent to 7.7 thousand barges, 443.9 thousand trucks or 115.4 thousand railcars, respectively. There are 43 freight port terminals are located along the Arkansas River (**Figure 1**), and 14 locks divide the river into 13 sections. Each lock chamber on the MKARNS is 110 feet wide by 600 feet long and can handle up to eight barges and a towboat. The U.S. Army Corps of Engineers (USACE) maintains a channel depth of nine feet on the MKARNS, allowing for barges to be loaded up to 1500 short tons (ODOT 2018).

Commodities transported on the MKARNS include iron & steel, fertilizers, petroleum products, minerals & building materials, grain (soybeans, wheat, and others), equipment and machinery, etc. (ODOT 2018). For this project, products transported on the MKARNS are grouped into eleven categories, merging the Lock Performance Monitoring System (LPMS) scheme and Arkansas Statewide Travel Demand Model (ARSTDM) scheme (**Table 1**). The temporal scope is one month, and spatial scope is individual port. LPMS data from 2009 to 2016 was used for this study.

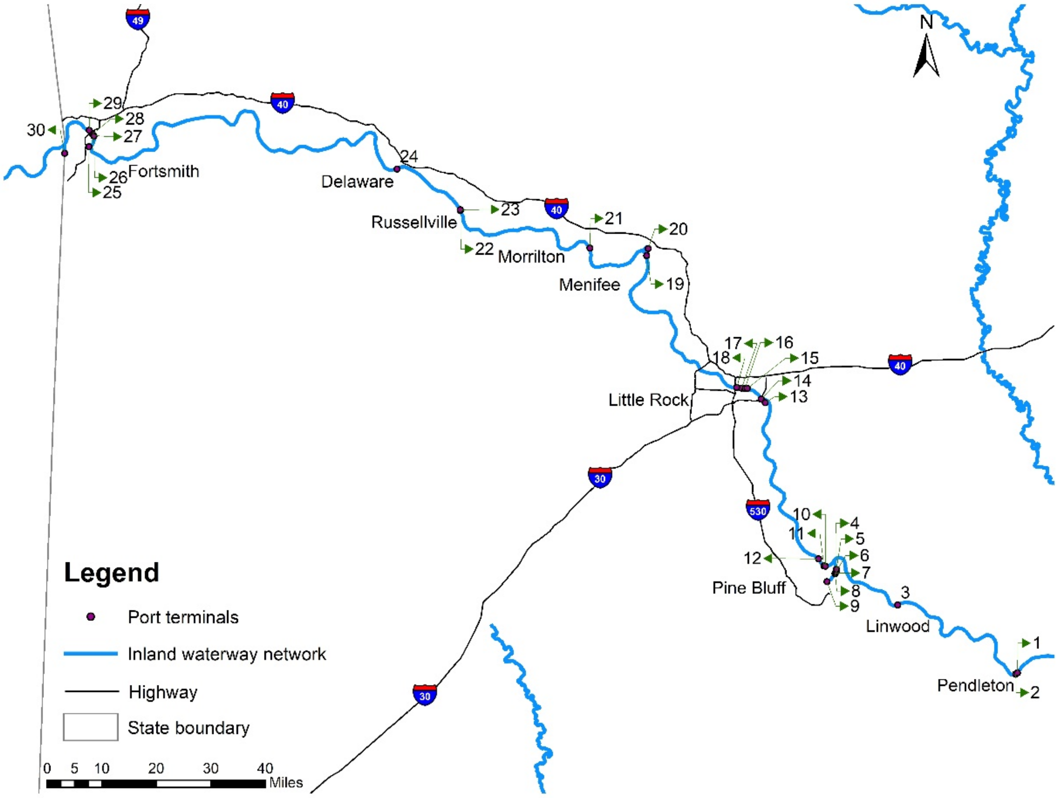

Figure 1. Arkansas portion of the MKARNS

Table 1. Commodity Classification

Commodity group
Agriculture and Food
Mining
Coal
Nonmetallic Minerals and Clay
Manufacturing
Lumber
Paper
Chemical
Petroleum
Primary Metal
Miscellaneous Mixed

1.4 Background

Understanding of freight supply chains is fragmented by the lack of data that connects freight modes. While it is possible to understand land side movements by truck using GPS data and movements along the navigable waterways using marine AIS data, only by examining where multiple modes meet is it possible to move towards more direct observation of the multi-modal supply chain. Besides mode specific datasets, of which little has been done to integrate, sources of publicly available multimodal supply chain data are limited to periodic surveys like the Commodity Flow Survey (CFS) conducted by the US Bureau of Transportation Statistics. To protect the privacy of private shippers/carrier respondents, the CFS aggregates the US into only 123 zones, of which Arkansas is represented as a single zone. This makes it challenging to disaggregate to port-level flows. Thus, it is critical to find other means of gathering port-level commodity flows that are more temporally continuous and spatially disaggregate and that maintain anonymity of shippers/carriers. This work builds on MarTREC Project 6008 by incorporating observed stochasticity of model input data gathered from publicly available datasets into the model framework. This allows representation of uncertainty in commodity volume data so that the model can be used to guide policy and investment.

Inland waterway ports are a critical part of the national freight network since they carry about half of the domestic US waterborne freight (U.S. Army Corps of Engineers, 2017). They contribute to the economy by moving commodities that are key to the Gross Domestic Product (GDP) of many states such as agriculture for Arkansas (University of Arkansas Division of Agriculture, 2021). There are both public and private port terminals in US inland waterway and commodity flow data collected by those inland waterway port operators are proprietary (Arkansas Waterways Commission, 2021) which makes it is challenging to understand ports' commodity throughput and capacity. Such data is valuable for infrastructure investment planning such as adding capacity to port servicing roads, waterway dredging schedules, and lock and dam maintenance programming. Furthermore, port level data enables port agencies and authorities to identify ports that need capacity expansion in terms of operational and storage infrastructures.

Publicly available statistics are published by the US Bureau of Transportation Statistics (BTS) via the Port Performance Freight Statistics. The program provides performance statistics for the top 25 ports based on overall cargo tonnage, 20-foot equivalent unit (TEU) of container cargo, and dry bulk cargo tonnage (Hu et al., 2021). However, the statistics are limited to the top 25 ports and only few inland waterway ports are included in the list. Additional data for gathering port commodity flows are the US Census Bureau's US Port Data (2021), Commodity Flow Survey (CFS) (U.S. Bureau of Transportation Statistics, 2021) and Waterborne Commerce Statistics from United States Army Corps of Engineers (USACE) (U.S. Army Corps of Engineers, 2021b). More recently the USACE has made available monthly commodity flow data at lock locations through the Lock Performance Monitoring System (LPMS) (U.S. Army Corps of Engineers, 2021a). However, data from such sources are limited in their spatial and/or temporal disaggregation which makes it challenging to collect port specific data for inland waterways. For example, the US CFS provides a periodic five-year snapshot of commodity flows through a survey that is expanded to represent the population using statistical sampling procedures. Notably, the LPMS data give monthly tonnage reports for locks but not ports, an many ports may be located between a pair of locks thus limiting the insights derived for single ports.

While collecting data at locks and dams may address inland waterway performance related questions, it alone does not help answer port specific investment questions. For example, data on commodity tonnage through a single lock may illuminate travel time delays across a given section of waterway.

However, it does not provide a means to guide strategic decisions regarding how a limited monetary investment (perhaps less than the amount needed to improve a single lock's performance) can be best used to alleviate delays along the inland waterway system. In addition, data on commodity tonnage through a section of waterway between a pair of locks provide historical patterns of commodity flows for a series of ports (between the lock pair) but does not provide insight into the ports' throughput to capacity ratio. Thus, it would be difficult to strategically guide capacity expansion investments without further data on existing port-specific capacity or operational characteristics. An intuitive investment decision using the currently available data may be to expand the capacity of the largest port among a series of ports between two locks. However, such a decision may lead to a local optimal solution such as adding a new shipping berth to a port to alleviate loading. From a systems point of view, investments at other ports that have greater access to highways and railways, for example, may lead to global optimal solutions. Expanding port infrastructure requires large capital investment and tends to target long lifespans, e.g., 25 years or more. Therefore, it is imperative that port capacity expansion investments are neither under nor over invested.

Another facet of strategic decision making in this context regards the potential for growth or decline in port usage as a function of overall freight shipment demand. Thus, decisions about port infrastructure investments, like any transportation investment, should be evaluated for differing freight demand scenarios that reflect the unknown nature of economically driven trends seen for freight transport. The demand of commodities for waterway transportation is not known in advance, and since the demand impacts infrastructure expansion decisions, different scenarios of demand need to be considered while making investment decisions. This calls for an investment model that considers several scenarios and provides optimal inland waterway port infrastructure investment solutions when uncertainty is present.

The goal of this study is to provide a tool to guide strategic waterway transportation infrastructure investments decisions. These decisions are subject to uncertainty observed in the demand for different commodities shipped via waterways, railways and highways; as well as transportation costs and transportation network structure. To accomplish the goal, this study proposes a two-stage stochastic optimization (2-SOP) model that minimizes the total of port infrastructure investment cost and the expected commodity transportation cost. The 2-SOP model has two sets of decision variables, first and second stage variables. First stage decisions determine equipment and storage investments to increase capacity. These decisions are made prior to the realization of the stochastic parameter, commodity demand (tonnage). After the realization of uncertainty, second-stage decisions determine the flow of commodities via different modes of transportation (Liu et al., 2018). The use of 2-SOP model, rather than using a deterministic model, avoids underestimation or overestimation of the supply chain costs when stochastic parameters exist. In addition, 2-SOP problem mimics reality in which decisions such as infrastructure investments are made first and without full information about uncertainty.

2 Methodological Approach

The methodology presented in this section consists of three main steps: i) Problem formulation; ii) Solution Approach and iii) Model evaluation.

2.1 Problem Formulation

We consider a supply chain network, where the nodes represent counties that have a positive demand and/or supply for different commodities and ports; and the arcs represent the transportation paths for delivering commodities. Let $W = (N, A)$ denote this transportation network where N is the set of nodes and A is the set of arcs that connect nodes of the network. Set $N = J \cup I \; (J \cap I = \emptyset)$ represents the counties $J = \{1, 2, \ldots, |J|\}$, from where commodities are shipped and received via a set of land-side transportation modes $T = \{Truck, Rail\}$ and a set of ports $I = \{1,2, \ldots, |I|\}$, from where commodities are transported via barges. Set $C = \{1,2,3, \ldots, |C|\}$ represents the set of commodity groups transported along network W over a set time period $P = \{1,2, \ldots, |P|\}$.

We consider two transportation modes between each origin and destination pair. The first path uses rail, truck and waterways to deliver commodities. Rail and truck are used to deliver commodities to and from each port. Waterways are used to deliver commodities between ports. The second path uses rail and truck to deliver commodities to and from counties. As an example, consider a supply chain network consisting of three counties that supply commodities via rail, truck and waterways (**Figure 2**). Two ports receive the commodities from counties and ship them via barges to two destination ports. Finally at the destination port, the commodity is shipped to the destination county where there is corresponding demand for that commodity. Any unmet demand is satisfied from counties outside the state (external zones).

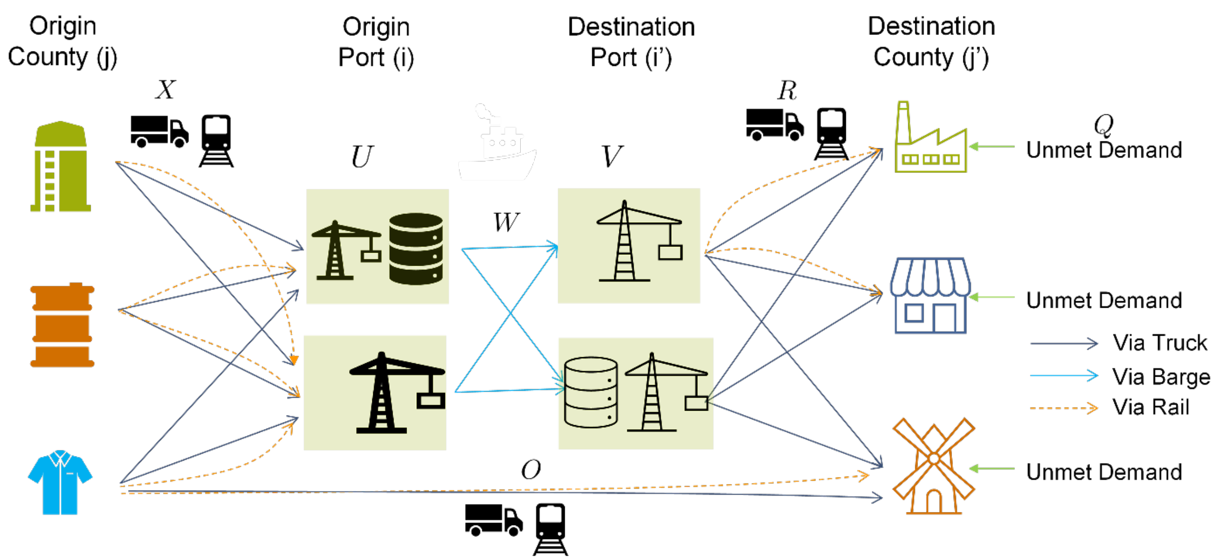

Figure 2. Network representation of supply chain

As a demonstration of the stochastic nature of commodity flows along the waterways and corresponding commodity demand for counties, consider the fluctuation in commodity demand along the Arkansas section of the MKARNS (**Figure 3**). Based on this non-linear and fluctuating demand variation over time, the demand for each commodity is considered stochastic in our model.

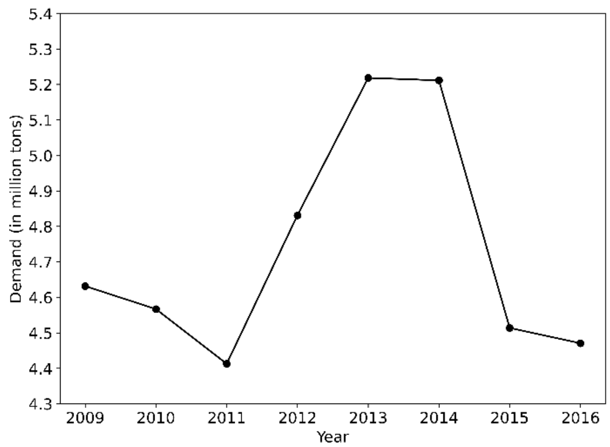

Figure 3. Commodity demand via waterways in Arkansas section of MKARNS

The notations used in the model formulation are summarized as follows:

<u>Sets:</u>

- I', I'' ports of origins/destinations of shipments ($I = I' \cup I''$)
- J', J'' counties of origins/destinations of shipments ($J = J' \cup J''$)
- P periods in the planning horizon, $p \in P$
- C commodity groups, $c \in C$
- E equipment, $e \in E$
- F storage facilities, $f \in F$
- S scenarios, $s \in S$

<u>Problem parameters:</u>

- ω_s is probability of occurrence of scenario s
- κ_e is the cost of equipment e
- ι_f is the cost of storage facility f
- α_{ij}^t is the unit cost of transporting commodity via truck between county $j \in J'$ and port $i \in I'$ (in \$/ton)
- α_{ij}^r is the unit cost of transporting commodity via rail between county $j \in J'$ and port $i \in I'$ (in \$/ton)
- a_{ik} is the unit cost of transporting commodity via barge from port $i \in I'$ to $k \in I''$ (in \$/ton)
- h_c is the unit inventory holding cost of commodity c
- μ is the unit penalty cost for unmet demand (in \$/ton)
- l_{jm}^t is the unit cost for transporting commodity via truck between county $j \in J^{\wedge\prime}$ and $m \in J''$ (in \$/ton)
- l_{jm}^r is the unit cost for transporting commodity via rail between county $j \in J'$ and $m \in J''$ (in \$/ton)
- β_{ec} takes the value 1 if equipment e can process commodity c, and takes the value 0 otherwise
- Γ_{fc} takes the value 1 if commodity c can be stored in storage facility f, and takes the value 0 otherwise
- δ_i takes the value 1 if port $i \in I$ has railway access, and takes the value 0 otherwise
- γ_j takes the value 1 if county $j \in J$ has railway access, and takes the value 0 otherwise

- B_{fe} is the minimum ratio of number of storage facility f to number of equipment e
- q_{jcps} is the supply availability of commodity c in county $j \in J'$ in month p in scenario s (in tons)
- d_{jcps} is the demand of commodity c in county $j \in J''$ in month p in scenario s (in tons)
- ζ_{fc} is the normalized tonnage of commodity c for inventory in storage facility f
- Λ_{ec} is the normalized tonnage of commodity c for processing in equipment e
- l_f is the storage capacity of storage facility f
- k_{if} is the existing number of storage facility f at port $i \in I$
- m_e is the processing capacity of equipment e
- n_{ie} is the existing number of equipment e at port $i \in I$

Decision Variables:

- Y_{if} is the number of storage facility f installed at port $i \in I$
- Z_{ie} is the number of equipment e installed at port $i \in I$
- X^t_{jicpes} is the tonnage of commodity c shipped via truck from county $j \in J'$ to port $i \in I'$ and processed using equipment e during period p in scenario s
- X^r_{jicpes} is the tonnage of commodity c shipped via rail from county $j \in J'$ to port $i \in I'$ and processed using equipment e during period p in scenario s
- U_{icpfs} is the tonnage of inventory of commodity c in origin port $i \in I'$ in storage facility f in month p in scenario s
- W_{ikcpes} is the tonnage of commodity c processed using equipment e and transported by barge from ports $i \in I'$ and $k \in I''$ in period p
- V_{icpfs} is the tonnage of inventory of commodity c in destination port $i \in I''$ in storage facility f in month p in scenario s
- R^t_{ijcpes} is the tonnage of commodity c processed using equipment e and transported from port $i \in I''$ to county $j \in J''$ via truck in month p in scenario s
- R^r_{ijcpes} is the tonnage of commodity c processed using equipment e and transported from port $i \in I''$ to county $j \backslash in\, J''$ via rail in month p in scenario s
- O^t_{jmcps} is the tonnage of commodity c shipped from county $j \in J'$ to county $m \in J''$ via truck in month p in scenario s
- O^r_{jmcps} is the tonnage of commodity c shipped from county $j \in J'$ to county $m \in J''$ via rail in month p in scenario s
- Q_{jcps} is the tonnage of commodity c shortage in county $j \backslash in\, J''$ in month p in scenario s

Our proposed mathematical model is defined as follows,

$$(WSN): \min \sum_{i\, \in I, e\, \in E} \kappa_e Z_{ie} + \sum_{i\, \in I, f\, \in F} \iota_f Y_{if} + E\left(H\left(X, \tilde{d}\right)\right) \qquad 1$$

subject to:

$$\sum_{i \in I} \sum_{e \in E} \kappa_e Z_{ie} + \sum_{i \in I} \sum_{f \in F} \iota_f Y_{if} \le b \qquad 2$$

$$Z_{ie} \in Z^+, \qquad \forall i \in I, e \in E \qquad 3$$

$$Y_{if} \in Z^+, \qquad \forall i \in I, f \in F \qquad 4$$

Function (1) represents the objective function that consists of the total of first stage costs (capacity expansion) and the expected second stage cost (transportation costs). The first term of function (1) represents the total cost of installing new operational (loading/unloading) equipment and the second term represents the total cost of installing new storage facilities. Constraint (2) ensures that the total cost does not surpass the total available budget allocated for investments to improve the inland waterway system. Constraints (3) and (4) determine the integer variables related to the decisions on storage facilities and operational equipment, respectively. Let $\mathcal{Y}=\{Z_{ie}, Y_{if} \mid i \in I, e \in E, f \in F\}$ represent solutions to problem (1) - (4). For a given value $\bar{\mathcal{Y}} \in \mathcal{Y}$, and a realization d of random demand $\tilde{d}$, the following is the formulation of second stage problem $H(\bar{\mathcal{Y}}, d)$.

$$\mathrm{H}(\bar{\mathcal{Y}}, d) = \min \sum_{c\in C}\sum_{p\in P}\Bigg(\sum_{i\in I'}\sum_{j\in J'}\sum_{e\in E}\left(\alpha_{ij}^{t} X_{jicpes}^{t} + \alpha_{ij}^{r} X_{jicpes}^{r}\right) + \sum_{i\in I'}\sum_{f\in F} h_c\, U_{icpf} + \sum_{i\in I'}\sum_{k\in I''} a_{ik}\, W_{ikcpe} + \sum_{i\in I''}\sum_{f\in F} h_c\, V_{icpf} + \sum_{i\in I''}\sum_{j\in J''}\sum_{e\in E}\left(\alpha_{ij}^{t} R_{ijcpe}^{t} + \alpha_{ij}^{r} R_{ijcpe}^{r}\right) + \sum_{j\in J'}\sum_{m\in J''}\left(l_{jm}^{t} O_{jmcp}^{t} + l_{jm}^{r} O_{jmcp}^{r}\right) + \sum_{j\in J''} \mu * Q_{jcp}\Bigg) \quad 5$$

subject to:

$$\sum_{m\in J''}\left(O_{jmcp}^{t} + O_{jmcp}^{r}\right) + \sum_{i\in I'}\sum_{e\in E}\left(X_{jicpe}^{t} + X_{jicpe}^{r}\right) \le q_{jcp} \forall j \in J', c \in C, p \in P \quad 6$$

$$Q_{mcp} + \sum_{i\in I''}\sum_{e\in E}\left(R_{imcpe}^{t} + R_{imcpe}^{r}\right) + \sum_{j\in J'}\left(O_{jmcp}^{t} + O_{jmcp}^{r}\right) = d_{mcp} \forall\, m \in J'', c \in C, p \in P \quad 7$$

$$\sum_{j\in J'}\sum_{e\in E}\left(X_{jicpe}^{t} + X_{jicpe}^{r}\right) + \sum_{f\in F} U_{icp-1f} = \sum_{f\in F} U_{icpf} + \sum_{k\in I''}\sum_{e\in E} W_{ikcpe} \,\forall i \in I', c \in C, p \in P \quad 8$$

$$\sum_{i\in I'}\sum_{e\in E} W_{ikcpe} + \sum_{f\in F} V_{kcp-1f} = \sum_{f\in F} V_{kcpf} + \sum_{m\in J''}\sum_{e\in E}\left(R_{kmcpe}^{t} + R_{kmcpe}^{r}\right) \forall k \in I'', c \in C, p \in P \quad 9$$

$$\sum_{c\in C}\left(\Lambda_{ec \sum_{j\in J''}(R_{ijcpe}^{t} + R_{ijcpe}^{r})} + \sum_{k\in I''}\left(W_{ikcpe} + W_{kicpe}\right) + \sum_{j\in J'}\left(X_{jicpe}^{t} + X_{jicpe}^{r}\right)\right) \le m_e\left(n_{ie} + \bar{Z}_{ie}\right) \quad \forall i \in I, p \in P, e \in E \quad 10$$

$$\sum_{c\in C} \zeta_{fc}\left(U_{icpf} + V_{icpf}\right) \le l_f\left(k_{if} + \overline{Y_{if}}\right) \forall i \in I, p \in P, f \in F \quad 11$$

$$U_{icpf} = 0 \;\forall i \in I', c \in C, p \in \{0\}, f \in F \quad 12$$

$$V_{icpf} = 0 \;\forall i \in I'', c \in C, p \in \{0\}, f \in F \quad 13$$

$$X_{jicpe}^{t}, X_{jicpe}^{r}, U_{icpf}, W_{ikcpe}, V_{icpf}, R_{ijcpe}^{t}, R_{ijcpe}^{r}, O_{jmcp}^{t}, O_{jmcp}^{r}, Q_{jcp} \in R^{+} \;\forall i \in I', k \in I'', j \in J', m \in J'', c \in C, p \in P, f \in F \quad 14$$

Function (5) minimizes the total supply chain costs. These costs include shipping cost via truck, rail, and barges, inventory cost, and a penalty cost for any unmet demand. Constraints (6) ensure that the amount of commodity shipped does not surpass the quantity of supply available. Constraints (7) capture the shortage of commodity in cases when demand exceeds supply. Constraints (8) and (9) are the flow balance constraints for origin and destination ports. Constraints (10) ensure that the commodity handled at origin and destination ports does not exceed the port capacity. Constraints (11) ensure that the total inventory at a port does not exceed the inventor holding capacity of that port.

Constraints (12) and (13) assign initial inventory at the destination and origin ports, respectively as zero. Constraints (14) are the non-negativity constraints.

2.2 Solution Approach

The computational burden from solving the problem formulated in (1) - (14**Error! Reference source not found.**) warrants a use of varied solution approaches. This section discusses Benders Decomposition algorithm and techniques to accelerate its convergence rate.

2.2.1 Benders Decomposition Algorithm

The uncertainty in commodity demand requires examination of a number of scenarios for a robust network design. We approximate the distribution of our stochastic demand via a discrete distribution. Let S represent the discrete set of demand realization and let $\omega_s \forall s \in S$ represent the corresponding probabilities.

The solution to 2-SOP problem can be computationally expensive based on the size of the problem given by $|I|, |J|, |C|, |P|$ and $|S|$. To overcome this computational burden, Benders decomposition algorithm (Benders, 1962), widely used to solve large-size, mixed integer linear problems, is employed. In this method, the original problem is decomposed into two subproblems: an integer master problem (MP) (15) to (16) and a $|S|$ linear subproblem (18). The MP along with an auxiliary variable and optimality cut provides an approximation of the original problem.

Benders decomposition is an iterative procedure. Let $\bar{\mathcal{Y}}$ be the solution of MP (15) to (16). For the given $\bar{\mathcal{Y}}$, $|S|$ subproblems are solved, one for each realization d_s of the stochastic demand $\tilde{d}$. The solutions to the subproblems are used to develop feasibility and optimality cuts that are added to the MP. These cuts ensure that, if the current solution $\bar{\mathcal{Y}}$ of the MP is not feasible or optimal to the original (WSN) problem, this solution is excluded from the feasible region and will not be used in other iterations of Benders algorithm. In each iteration of the algorithm, a lower bound and an upper bound are generated. The lower bound is given by objective value of $(M - WSN)$, and the solution of $(M - WSN)$ and $\left(S - WSN(s)\right)$ provides an upper bound for the original problem (WSN) This is continued until the relative gap between the lower and upper bound converge to a given threshold. The following model (15) to (16) is the MP.

$$(M - WSN): min \sum_{i \in I e \in E} \kappa_e Z_{ie} + \sum_{i \in I f \in F} \iota_f Y_{if} + \sum_{s \in S} \omega_s \theta_s \qquad 15$$

Subject to: (2-4)

$$\theta_s^n \geq \sum_{j \in J} \sum_{c \in C} \sum_{p \in P} (\nu_{jcps} q_{jcp} + \xi_{jcps} d_{jcps} + \sum_{i \in I} \sum_{e \in E} \sum_{p \in P} \pi_{ipes}\, m_e\left(n_{ie} + Z_{if}\right) + \sum_{i \in I} \sum_{f \in F} \sum_{p \in P} \sigma_{ipfs}\, l_f\left(k_{if} + Y_{if}\right) \forall s \in S, n = \{1, \dots, N'\} \qquad 16$$

where N' is the current number of iterations.

The Benders decomposition solves $(M - WSN)$ iteratively and in each iteration n, let $\bar{\mathcal{Y}}^n$ represent the corresponding solution. Let Φ^n be the objective function value of model (15) to (16) obtained at the n^{th} iteration. Let ϕ^n, given by equation (17), be the corresponding cost of infrastructure investment.

$$\phi^n = \sum_{i \in I e \in E} \kappa_e Z_{ie}^n + \sum_{i \in I f \in F} \iota_f Y_{if}^n \qquad 17$$

Given $\bar{\mathcal{Y}}^n$, following scenario-based subproblem $\left(S - WSN(s)\right)$ is solved for each scenario $s \in S$.

$$\left(S - WSN(s)\right) : min \sum_{c \in C} \sum_{p \in P} \left(\sum_{i \in I'} \sum_{j \in J'} \sum_{e \in E} (\alpha_{ij}^t X_{jicpes}^t + \alpha_{ij}^r X_{jicpes}^r) + \sum_{i \in I'} h_c\, U_{icpfs} + \sum_{i \in I'} \sum_{k \in I''} \sum_{e \in E} a_{ik}\, W_{ikcpes} + \sum_{i \in I''} h_c\, V_{icpfs} + \sum_{i \in I''} \sum_{j \in J''} \sum_{e \in E} (\alpha_{ij}^t R_{ijcpes}^t + \alpha_{ij}^r R_{ijcpes}^r) + \sum_{j \in J'} \sum_{m \in J''} \{ (l_{jm}^t O_{jmcps}^t + \} l_{jm}^r O_{jmcps}^r) + \sum_{j \in J''} \mu * Q_{jcps} \right) \quad 18$$

Subject to: (6) – (14)

Constraint (**Error! Reference source not found.**) is the scenario specific optimality cut added to ($M - WSN$) in each iteration of the Benders decomposition algorithm. $\nu_{jcps} \forall\, j \in J'$, $\xi_{jcps}\ \forall\, j \in J''$, υ_{icps}, $\chi_{icps}\ \forall\, i \in I''$, π_{ipes}, σ_{ipfs} are dual variables for constraints (6) to (10) respectively. Let $\mathcal{X}^n(s)$ denote this solution for scenario s at n^{th} iteration. Constraint (7) makes the scenario-based subproblem $\left(S - WSN(s)\right)$ always feasible for any values in first-stage decision variables. This constraint makes sure that the demand is satisfied either by the supply from counties within the state or via other states. For this reason, we do not need to add feasibility cuts to the MP. The dual problem of $\left(S - WSN(s)\right)$ is shown in Appendix A1.

Let Θ_s^n be the objective value of function s^{th} $\left(S - WSN(s)\right)$ for scenario s. We calculate Θ^n as follows

$$\Theta^n = \sum_{s \in S} \omega_s\, \Theta_{(s)}^n$$

The pseudo-code of the Benders Decomposition algorithm is shown in Algorithm 1.

Algorithm 1: Benders Decomposition Algorithm

```
Initialize ϵ. Set n ← 1, LB^n ← −∞, UB^n ← +∞, abort ← false
while abort = false do
    Solve (M − WSN) to obtain Φ^n, ϕ^n and 𝒴^n
    if Φ^n > LB then
        LB^n ← Φ^n
    end
    For all s ∈ S, Solve (S − WSN(s)) to obtain Θ_s^n and 𝒳^n(s)
    if UB^n > Θ^n + ϕ^n then
        UB^n ← Θ^n + ϕ^n
    end
    if (UB^n − LB^n)/UB^n ≤ ϵ then
        abort ← true
    else
        Add cut (Error! Reference source not found.) to (M − WSN)
        n ← n + 1
    end
end
return UB, 𝒴^n and 𝒳^n(s)
```

2.2.2 Methods to Accelerate Benders Decomposition Algorithm

The Benders Decomposition algorithm is known to be extremely slow and computationally expensive. Côté and Laughton (1984), Magnanti and Wong (1981), McDaniel and Devine (1977), Poojari and Beasley (2009), Rei et al. (2009), and Saharidis et al. (2010) have developed and successfully implemented acceleration techniques to improve the convergence rate of Benders Decomposition. In our study we add Knapsack inequalities and Pareto-optimal cuts to enhance the convergence rate of the algorithm. The details of the proposed techniques are provided in later parts of this section.

Knapsack inequalities:

Adding knapsack inequalities can improve the convergence rate of the algorithm by reducing the solution space of $(M - WSN)$, thus, reducing the time it takes to solve the problem (Santoso et al., 2005). In addition, state-of-the-art solvers, such as GUROBI and CPLEX, can extract valid inequalities from knapsack inequalities. Let LB^n denote the lower bound obtained in n^{th} iteration of the algorithm. Hence, following knapsack inequalities are added to $(M - WSN)$ in the $(n+1)^{th}$ iteration to accelerate the convergence rate of Benders algorithm:

$$LB^n \leq \sum_{i \in I e \in E} \kappa_e \, Z_{ie} + \sum_{i \in I, f \in F} \iota_f \, Y_{if} + \sum_{s \in S} \omega_s \, \theta_s \qquad 19$$

Pareto-optimal cuts:

The subproblems $\left(S - WSN(s)\right)$ are capacitated transportation problems. The transportation problem is degenerate in nature (Ahuja et al., 1988), that is, it has multiple optimal solutions and each solution generates optimality cuts of different strength. Hence, the solution of the subproblem should be chosen in such a way that it produces the strongest cuts. Magnanti & Wong (1981) found that adding Pareto-optimal cuts to the MP improves the convergence rate of the Benders Decomposition algorithm. The generation of Pareto-optimal cuts proposed by (Magnanti & Wong, 1981) requires solving two subproblems, one associated with solution of the MP and another associated with core points. A point $y \in ri(Y^c)$ is a core point of Y, where $ri(S)$ and S^c are the relative interior and convex hull of set $S \subseteq R^k$ respectively (Papadakos, 2008). Let, $\bar{\mathcal{Y}}^n$ be the solution of the MP at n^{th} iteration and $\bar{\mathcal{Y}}^o$ ={ $Z_{ie}^{o,n} = 0, Y_{if}^{o,n} = 0 | n = 1, i \in I, e \in E, f \in F\}$ be the set of initial core points. The subproblem to generate Pareto-optimal cuts is given as:

$$\begin{aligned}\left(MWS - WSN(s)\right): max \sum_{j \in J} \sum_{c \in C} \sum_{p \in P} \sum_{s \in S} \left(\nu_{jcps} q_{jcps} + \xi_{jcps} d_{jcps}\right) \\ + \sum_{i \in I} \sum_{e \in E} \sum_{p \in P} \sum_{s \in S} \pi_{ipes} \, m_e (n_{ie} + Z_{ie}^o) + \sum_{i \in I} \sum_{f \in F} \sum_{p \in P} \sum_{s \in S} \sigma_{ipfs} \, l_f \left(k_{if} + Y_{if}^o\right)\end{aligned} \qquad 20$$

subject to: (24) - (38)

$$\begin{aligned}\sum_{j \in J} \sum_{c \in C} \sum_{p \in P} \left(\nu_{jcps} q_{jcps} + \xi_{jcps} d_{jcps}\right) + \sum_{i \in I} \sum_{e \in E} \sum_{p \in P} \pi_{ipes} \, m_e (n_{ie} + Z_{ie}^n) \\ + \sum_{i \in I} \sum_{f \in F} \sum_{p \in P} \sigma_{ipfs} \, l_f \left(k_{if} + Y_{if}^n\right) = \Theta^n(s) \; \forall s \in S\end{aligned} \qquad 21$$

Since this technique relies on the solution of the subproblems, Papadakos (2008) proposed a methodology to generate sub problem independent Pareto-optimal cuts. Papadakos (2008)showed that by using different core points at each iteration, constraint (21) could be disregarded. Pareto-optimal cuts generated by this approach are commonly known as modified $Magnanti - Wong\ pareto - optimal\ cuts$. The core points are updated in every iteration as follows:

$$Z_{ie}^{o,n+1} = (1-\lambda)Z_{ie}^{o,n} + \lambda Z_{ie}^{n} \quad 22$$

$$Y_{if}^{o,n+1} = (1-\lambda)Y_{if}^{o,n} + \lambda Y_{if}^{n} \quad 23$$

Papadakos (2008) and Mercier et al. (2005) empirically showed that the value of λ = 0.5 gives the best result. Pseudo code of the accelerated Benders Decomposition with Pareto-optimal cuts is provided in Algorithm 2.

Algorithm 2: Benders Decomposition Algorithm with Knapsack Inequalities and Pareto-optimal Cuts

Initialize ϵ, $Z_{ie}^{o,n}$, $Y_{if}^{o,n}$. Set $n \leftarrow 1$, $LB^n \leftarrow -\infty$, $UB^n \leftarrow +\infty$, abort $\leftarrow$ $f$alse

while $abort = false$ **do**

  For all $s \in S$, Solve $\left(MWS - WSN(s)\right)$ to obtain Θ_s^n and $\mathcal{Y}^n(s)$

  Add cuts (**Error! Reference source not found.**) to $(M - WSN)$

  Solve $(M - WSN)$ to obtain Φ^n, ϕ^n and $\mathcal{Y}^n$

  if $\Phi^n > LB$ **then**

    $LB^n \leftarrow \Phi^n$

  end

  For all $s \in S$, Solve $\left(S - WSN(s)\right)$ to obtain Θ_s^n and $\mathcal{X}^n(s)$

  if $UB^n > \Theta^n + \phi^n$ **then**

    $UB^n \leftarrow \Theta^n + \phi^n$

  end

  if $\frac{UB^n - LB^n}{UB^n} \leq \epsilon$ **then**

    $abort \leftarrow true$

  else

    Add cuts (**Error! Reference source not found.**) to $(M - WSN)$

    Add Knapsack inequalities (19) to $(M - WSN)$

    Update core points using equations (22) and (23)

    $n \leftarrow n + 1$

  end

end

return UB, $\mathcal{Y}^n$ and $\mathcal{X}^n(s)$

3 Computational Study and Managerial Insights

This section discusses the solution quality of the proposed algorithm, a case study application of the model to the Arkansas Section of the MKARNS, and performance evaluation of stochastic solutions.

3.1 Performance Evaluation

We solved a variety of problems to determine the quality of the proposed algorithms (**Table 2**) and compare various solution approaches: (1) Gurobi solver, (2) Benders algorithm, (3) Benders with Knapsack inequalities, (4) Benders with Knapsack inequalities and Pareto-optimal cuts. We use the following stopping criteria to terminate the algorithm: (i) optimality gap $\leq 1\%$, (ii) number of iterations ≥ 500 and (iii) algorithm run time $\geq$ 12,600 seconds. The experiments are carried out on a Windows 10 PC with an Intel Core i7 3.2 GHz processor and 32 GB of RAM. Results of the experiments are summarized by the algorithm run time in seconds $t(s)$, the optimality gap $\epsilon(\%)$, and the number of iterations (n) at the time with the stop criteria is met (**Table 3**).

Table 2 Test Case Size

Problem Nr.	$\lvert I\rvert$ *Ports*	$\lvert J\rvert$ *Counties*	$\lvert C\rvert$ *Commodity*	$\lvert P\rvert$ *Periods*	$\lvert S\rvert$ *Scenarios*	Problem Nr.	$\lvert I\rvert$ *Ports*	$\lvert J\rvert$ *Counties*	$\lvert C\rvert$ *Commodity*	$\lvert P\rvert$ *Periods*	$\lvert S\rvert$ *Scenarios*
1	15	45	5	6	4	9	30	75	11	9	12
2	15	45	5	6	8	10	30	75	11	9	16
3	15	45	5	6	10	11	30	75	11	12	4
4	15	45	5	6	12	12	30	75	11	12	8
5	15	45	5	6	16	13	30	75	11	12	10
6	30	75	11	9	4	14	30	75	11	12	12
7	30	75	11	9	8	15	30	75	11	12	16
8	30	75	11	9	10						

Overall, Gurobi outperforms the alternative solution approaches in terms of run time, optimality gap, and number of iterations in all experiments except for the cases (8)-(10) and (12)-(15). This warrants the implementation of Benders decomposition algorithm. Although standard Benders decomposition does not have memory issues, it fails to converge to solutions with smaller than 1% optimality gap within the given time limit for medium and large sized problems (problems (6) - (15), **Table 3**). This finding justifies the use of an acceleration techniques to improve the convergence rate of the algorithm. With the addition of Knapsack inequalities, improvement in terms of solution time, optimality gap and number of iterations is observed for majority of the cases. However, like the standard Benders decomposition, it is not successful in finding solutions with smaller than 1% optimality gap within 12,600 seconds (3.5 hours). We observe a significant improvement in terms of solution quality and run time when we add both Knapsack inequalities and Pareto-optimal cuts to the MP. Benders decomposition algorithm with Knapsack inequalities and Pareto-optimal cuts provides solutions of high quality within the proposed threshold run time.

Table 3 Comparison of solution approaches

Case	Gurboi		Benders			Benders + KI			Benders + KI + PO		
	$t(s)$	$\epsilon(\%)$	$t(s)$	$\epsilon(\%)$	n	$t(s)$	$\epsilon(\%)$	n	$t(s)$	$\epsilon(\%)$	n
1	**5**	0.9	40	0.51	21	38	0.45	19	26	0.95	7
2	**20**	0.07	160	0.99	38	135	0.98	31	119	0.28	15
3	**25**	0.16	378	0.99	64	349	1	55	131	0.85	13
4	**36**	0.22	185	0.97	29	151	0.98	24	144	0.92	12
5	**52**	0.09	161	0.98	20	161	0.97	20	157	0.99	10
6	**955**	0.51	12,600	5.09	179	12,600	4.78	135	2,960	0.81	30
7	**4,063**	0.46	12,600	4.96	118	12,600	4.51	99	5,784	0.88	30
8	-	-	12,600	4.52	95	12,600	4.41	88	**5,442**	0.96	23
9	-	-	12,600	3.24	87	12,600	2.76	80	**6,702**	0.48	24
10	-	-	12,600	2.15	66	12,600	2.02	63	**7,421**	0.89	20
11	1,732	0.77	12,600	5.48	157	12,600	5.2	135	5,232	0.8	36
12	-	-	12,600	5.33	88	12,600	5.1	83	**8,820**	0.98	31
13*	-	-	12,600	5.41	72	12,600	4.94	67	**10,245**	0.94	29
14	-	-	12,600	3.77	61	12,600	3.42	58	**10,407**	0.85	25
15	-	-	12,600	2.70	46	12,600	2.70	46	**12,586**	0.69	23

* indicates a representative case size for a real-world problem

3.2 Case Study: McClellan-Kerr Arkansas River Navigation System

The 2-SOP problem is applied to the Arkansas section of the MKARNS. This 308-mile inland waterway system has 13 locks and 43 freight ports.

3.2.1 Data Description

3.2.1.1 Commodity demand and supply data

The US Army Corps of Engineers (USACE) produces the Lock Performance Management System (LPMS) report which records monthly commodity volume (in tons) passing through each lock. The LPMS classifies commodities into nine groups and several subgroups. Data is available for the years 2009 through 2016. This source was used for creating demand scenarios in the 2-SOP model.

The Arkansas Statewide Travel Demand Model (Alliance Transportation Group and Cambridge Systematics, 2012) maintained by ARDOT reports amounts (in tons) of commodities shipped between counties. The total amount of commodity shipped from a county is considered to be the supply for the (shipment) origin county. Similarly, the total amount of commodities received by a county is considered to be the demand for that (shipment) destination county. The commodity grouping in the Statewide Travel Demand Model and LMPS differed. We reorganized commodities into 11 groups that align with the subgroups reported by LMPS.

3.2.1.2 Transportation cost data

The transportation cost (in $/ton-mile) for truck and rail are derived from the ARDOT Travel Demand Model (Alliance Transportation Group and Cambridge Systematics, 2012), Bureau of Transportation Statistics (2010) and Surface Transportation Board (2003). The barge transportation cost (in $/ton-mile) is derived from the data reported by (United States Department of Agriculture, 2001) for major cities along waterways. However, the data is not reported for cities along the Arkansas River. Thus, rates from the surrounding region in St. Louis and Cairo-Memphis were averaged to estimate the costs for the Arkansas river. Transportation cost data used in our study are shown in **Table 5** in Appendix A2.

3.2.1.3 Port capacity data

The National Transportation Atlas Database (NTAD), published by the United States Department of Transportation BTS (2021) contains information about the location, commodity handling equipment, storage facility, and road/railway connection of terminals at US coastal, Great Lakes, and inland ports. Data about the number of equipment units and storage facilities are gathered from this dataset. This data is supplemented and updated using information obtained from ports' websites and satellite images. The data for port processing and storage capacity used in our study are shown in **Table 8** and **Table 9**, respectively, in Appendix A2.

3.2.1.4 Infrastructure cost

Costs for equipment (i.e., crane, conveyor, hooper, forklift) and storage facilities (i.e., warehouse, storage tank, paved and unpaved storage) are obtained from Braham et al. (2017). These costs include labor and materials as well as general overhead. Braham et al. (2017) selected these costs from a material, construction and equipment cost database from (RS Means, 2014) and (RS Means, 2017), and validated through interviews with industry representatives. All costs for our study are calculated based on the 2020-dollar value. The infrastructure cost used in our study are shown in **Table 6** and **Table 7** in Appendix A2.

3.2.1.5 Scenario definitions

To capture the impacts of stochasticity in commodity demand to infrastructure decisions, we generate 10 different demand scenarios. Scenarios 1 to 8 are based on historical commodity throughput data gathered from the LPMS between 2009 and 2016. We assign a probability of 6.5% to the scenario developed using historical data from 2009 and increase this probability by 0.5% per year for subsequent years. This is done with the assumption that the demand scenarios developed from recent years have more probability of occurrence compared to scenarios developed from previous years. The probability of scenarios developed from historical data sum to 66%. Two additional scenarios: scenario 9 with probability of 20% and scenario 10 with probability of 14%, are developed based on the 15 and 25 year commodity demand projection from BTS (2019), respectively.

3.2.2 Case Study Results

We evaluate the impacts of varying investment for infrastructure investments on system cost and volume of commodities moved via waterways. The total system cost, which includes transportation cost and investment cost, decreases from $1.245 billion to $1.225 billion ($20 million decrease) as the investment increases from $2 to $6 million ($4 million increase) (**Figure 4**). The rate of decrease in total system cost reduces after the investment reaches $6 million. The total system cost drops to $1.216 billion at an investment of $10 million. We see similar results for unit supply chain cost where the unit cost decreases from $21.15/ton to $20.74/ton ($0.41/ton decrease) as the investment increases from $2 million to $6 million ($4 million decrease) and eventually to $20.55/ton for an investment of $10 million (**Figure 5**). This decrease in unit cost can be attributed to the increased volume of commodity shipped via waterways which have the lowest transportation cost compared to truck and rail, as shown in **Figure 7**.

We discuss the frequency and the range of investments at individual ports for five different investment scenarios, $2M, $4M, $6M, $8M, and $10M (**Figure 6**). In most of these scenarios, the model determines that investments should be made on ports located near Little Rock (central Arkansas), Delaware (mid-western section of the Arkansas river) and Fort Smith (western most port on the Arkansas river). In every scenario considered, the model determines that investments in capacity expansion equipment should be made at the port located near Little Rock (Port 15 in **Figure 6**). Across all scenarios, several ports repeatedly receive no investment e.g., Menifee, Morrilton, and Russellville.

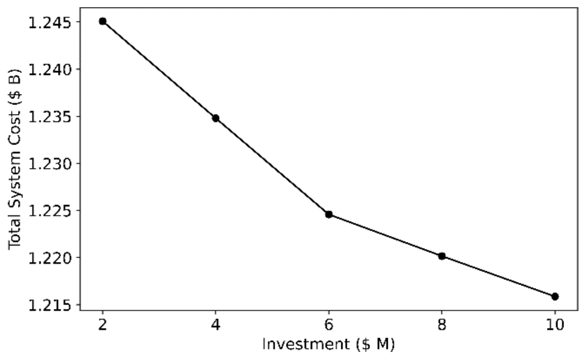

Figure 4 Total System Cost

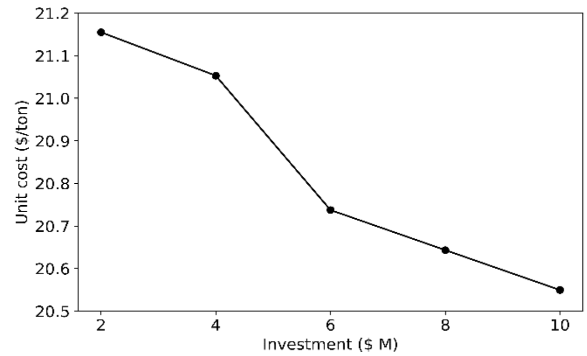

Figure 5 Unit Supply Chain Cost

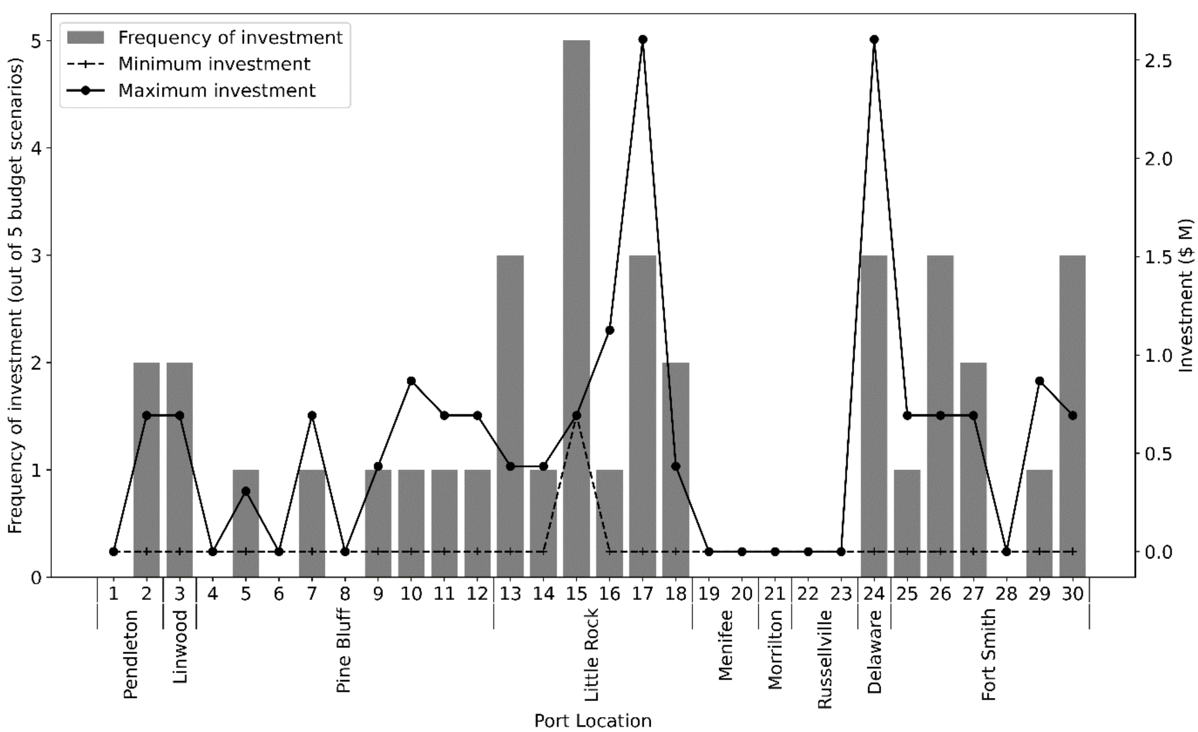

Figure 6 Port Investment

The volume of commodity shipped via waterways increases as the investment in port capacity increases (**Figure 7**). A total of 2.74 million tons of commodity is shipped via waterways (6% of total shipments by all three modes) when the total investment is $2 million. This volume increases to 3.47 million (8% of total shipments by all three modes) when the amount invested in infrastructure is $10 million. We also report the change in ton-miles shipped by waterways for varying investment (**Figure 8**). Ton-mile reflects both the volume (tons) and distance (miles) shipped and is one of the most well used measures of the physical volume of freight transportation services (United States Department of Transportation BTS, 2012). The waterway serves nearly 188 million ton-miles of freight (5% of all three modes) at an investment of $2 million and increases to 231 million ton-miles (6% of all three modes) at an investment of $10 million.

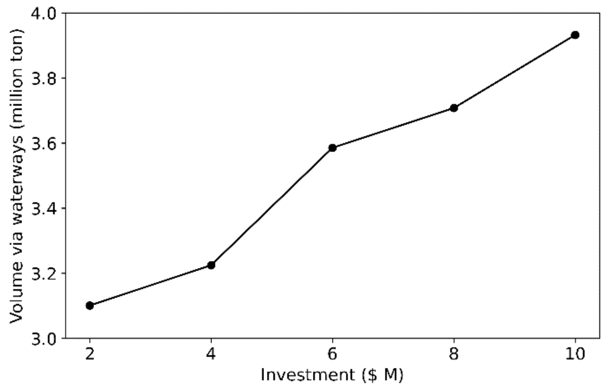

Figure 7 Total Volume

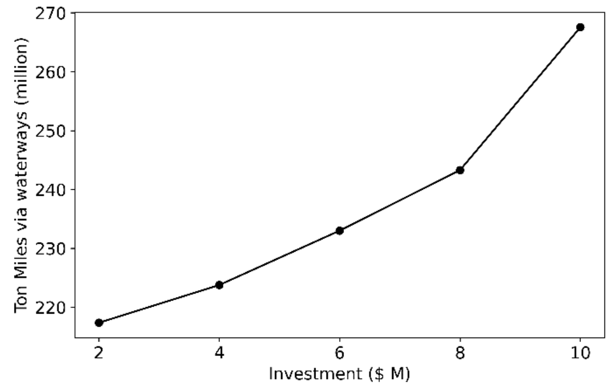

Figure 8 Total Ton-miles

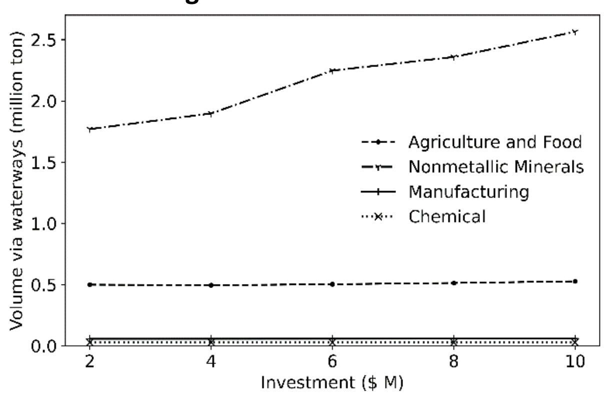

Figure 9 Commodity Volume

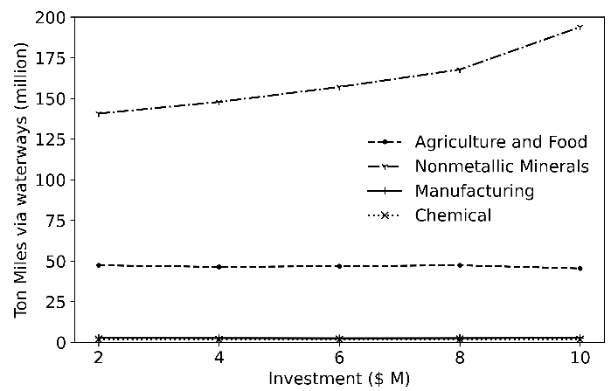

Figure 10 Commodity Ton-miles

The 2-SOP model represents 11 commodity groups and models the shipment, supply, and demand for each commodity group uniquely while considering how port infrastructure can be shared among commodity groups, e.g., a grain elevator can be used for aggregates and gain. To demonstrate this modeling contribution, we present the results for four commodity groups that dominate the Arkansas section of the MKARNS (Asborno et al., 2020): nonmetallic minerals, agriculture and food, manufacturing, and chemicals. The volume of nonmetallic minerals shipped via waterways increases from 1.5 million (56% of all commodities) to 2.2 million tons (64% of all commodities) (an increase of 0.8 million tons) as the investment increases from $2 million to $10 million (**Figure 9**). This represents a reduction of 1.4% of non-metallic minerals shipped via truck and rail. Similarly, the nonmetallic mineral freight generated via waterways increases from 121 million ton-miles (64% of all commodities) to 167 million ton-miles (72% of all commodities) (an increase of 45 million ton-miles) when investments increase from $2 million to $10 million (**Figure 10**). No significant impact of port infrastructure investment is seen for the other three commodity groups.

3.3 Evaluation of Stochastic Solutions

To demonstrate the benefit of developing the model with stochastic elements, as opposed to a deterministic model, we calculate the Value of Stochastic Solution (VSS). VSS is the difference between the objective function value of the stochastic solution and the expected value solution. We also calculate the Expected Value of Perfect Information (EVPI) which captures the value of knowing the future with certainty. EVPI is the difference between the objective function value of the stochastic solution and the wait-and-see solutions (WSS).

Table 4 Comparison of Stochastic and Deterministic Solutions

Strategies	Value ($M)
Stochastic programming	1,225
Wait and see solutions	
Scenario-1	1,433
Scenario-2	626
Scenario-3	680
Scenario-4	538
Scenario-5	1,367
Scenario-6	2,336
Scenario-7	976
Scenario-8	286
Scenario-9	1,455
Scenario-10	1,798
Expected value solution	1,246

The expected value solution is calculated in two steps. First, we solve (WSN) assuming a single scenario, represented by the expected values of commodity demand. Next, we fix the values of the first stage decisions using this solution, and resolve (WSN) to obtain the expected value solution of $1,246M. Hence, VSS = $1,246M - $1,225M = $21M. Therefore, the expected savings from solving the stochastic model, rather than the corresponding deterministic model, is $21M per year.

The following is the approach we use to calculate EVPI. Let us assume that we know exactly what scenario is realized in the future. Then, we can solve (WSN) for this particular scenario. This is the "wait-and-see solution" (WSS). **Table 4** summarizes the WSS of each scenario. We calculate the expected value of WSS to be $1,221M ($1,433M*0.065 + $626M*0.07 + $680M*0.075 + $538M*0.08 + $1,367M*0.085 + $2,336M*0.09 + $976M*0.095 + $286M*0.1 + $1,455M*0.20 + $1,798M*0.14). Recall the multipliers in the above calculation correspond to the likelihood of each scenario occurring. Therefore, EVPI is $4M ($1,225M - $1,221M). Hence, the price we should be willing to pay for correctly predicting future realizations of commodity demand, should be no more than $4M.

4 Conclusion

In this research, we develop a model to guide strategic investments in inland waterway port infrastructure investments given uncertainty in commodity demand. This study proposes 2-SOP model that seeks to minimize the total of port infrastructure investment costs and the expected transportation costs. We implement a Benders decomposition algorithm to solve the model and accelerate this algorithm using Knapsack inequalities and Pareto-optimal cuts. The computational analysis reveals that Benders with Knapsack inequalities and Pareto-optimal cuts outperforms Gurobi and the traditional Benders algorithm for large-sized problems.

We apply the two-stage stochastic optimization model to the Arkansas section of the McCllelan-Kerr Arkansas River Navigation System (MKARNS). The model results show that while the total system cost (transportation plus investment costs) decreases with increasing investment, the rate of decrease in system cost is convex in nature, i.e., the rate of change decreases with the dollar amount invested in port capacity expansion. Our model shows that commodity volume and, as expected, the percent of that volume that moves via waterways (in ton-miles) increases with increasing investment in port infrastructure. The model captures individual commodity movements in terms of tons and ton-miles shipped by transportation mode. Results show that among all commodities, nonmetallic minerals experience the largest fluctuation in the tonnage and ton-miles shipped as a consequence of changing investment amounts. Furthermore, since the model estimates investments and commodity throughput at individual ports, we can identify a cluster of ports (Little Rock, Fort Smith) that should receive investment in port capacity under any investment scenario.

Finally, to demonstrate the value of a stochastic formulation over a deterministic approach, we calculate the value of the stochastic solution, VSS. VSS shows that a failure to use stochastic model to capture variations in commodity demand, could cost up to $21 M per year. We also calculate the expected value of perfect information (EVPI). EVPI indicates the price we should be willing to pay for correctly predicting future realizations of commodity demand, should be no more than $4M.

5 Future Work

This research opens several research avenues to explore in future. The current model allows for analysis of the supply chain within a state with assumptions about out of state (external) commodity demand. In the future, the model will be expanded to incorporate multiple states to form a “freight region’ by adding railway and highway networks connecting neighboring states.

Furthermore, in our current study, there are no port disruption scenarios considered. Closure of a port due to human-induced and/or natural causes can impact port operation. This leads to temporary decrease in port capacity. The event and impact of port disruption is uncertain and therefore, future study will incorporate this uncertainty of port disruption for modeling port infrastructure planning.

While our current work used knapsack inequalities and Pareto-optimal cuts to accelerate the Benders decomposition algorithm, our future work will explore additional enhancement techniques such as maximum density cut generation and Benders-type heuristics.

6 References

Ahuja, R. K., Magnanti, T. L., & Orlin, J. B. (1988). *Network Flows*.

Alliance Transportation Group and Cambridge Systematics. (2012). *Arkansas Statewide Travel Demand Model Phase I-Freight Model*.

Arkansas Waterways Commission. (2021). *River Ports and Terminals*. https://www.waterways.arkansas.gov/ports-terminals/mississippi-river/

Asborno, M. I., Hernandez, S., & Akter, T. (2020). Multicommodity port throughput from truck GPS and lock performance data fusion. *Maritime Economics & Logistics*, *22*(2), 196–217. https://doi.org/10.1057/s41278-020-00154-7

Benders, J. F. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, *4*(1), 238–252. https://doi.org/10.1007/BF01386316

Braham, A., Hernandez, S., Kent, J., Asborno, M., & Casillas, S. (2017). *Locating Transload Facilities to Ease Highway Congestion and Safeguard the Environment*.

Bureau of Transportation Statistics. (2010). *National Transportation Statistics 2010*. https://www.bts.gov/archive/publications/national_transportation_statistics/2010/index

Côté, G., & Laughton, M. A. (1984). Large-scale mixed integer programming: Benders-type heuristics. *European Journal of Operational Research*, *16*(3), 327–333. https://doi.org/10.1016/0377-2217(84)90287-X

Hu, P. S., Schmitt, R. R., Nguyen, L., Chambers, M., & Friedman, D. (2021). *Port Performance Freight Statistics in 2019: Annual Report to Congress 2020*. https://rosap.ntl.bts.gov/view/dot/54022

Liu, A., Lau, V. K. N., & Zhao, M.-J. (2018). Online Successive Convex Approximation for Two-Stage Stochastic Nonconvex Optimization. *IEEE Transactions on Signal Processing*, *66*(22), 5941–5955. https://doi.org/10.1109/TSP.2018.2871389

Magnanti, T. L., & Wong, R. T. (1981). Accelerating Benders Decomposition: Algorithmic Enhancement and Model Selection Criteria. *Operations Research*, *29*(3), 464–484. https://doi.org/10.1287/opre.29.3.464

McDaniel, D., & Devine, M. (1977). A Modified Benders' Partitioning Algorithm for Mixed Integer Programming. *Management Science*, *24*(3), 312–319.

Mercier, A., Cordeau, J.-F., & Soumis, F. (2005). A computational study of Benders decomposition for the integrated aircraft routing and crew scheduling problem. *Computers & Operations Research*, *32*(6), 1451–1476. https://doi.org/10.1016/j.cor.2003.11.013

Nachtmann, H., Boudhoum, O., Oztanriseven, F., Mack-Blackwell National Rural Transportation Study Center (U.S.), & Arkansas. State Highway & Transportation Dept. (2015). *Regional economic impact study for the McClellan Kerr Arkansas River Navigation System.* https://rosap.ntl.bts.gov/view/dot/29525

Papadakos, N. (2008). Practical enhancements to the Magnanti–Wong method. *Operations Research Letters*, *36*(4), 444–449. https://doi.org/10.1016/j.orl.2008.01.005

Poojari, C. A., & Beasley, J. E. (2009). Improving benders decomposition using a genetic algorithm. *European Journal of Operational Research*, *199*(1), 89–97. https://doi.org/10.1016/j.ejor.2008.10.033

Rei, W., Cordeau, J.-F., Gendreau, M., & Soriano, P. (2009). Accelerating Benders Decomposition by Local Branching. *INFORMS Journal on Computing*, *21*(2), 333–345. https://doi.org/10.1287/ijoc.1080.0296

RS Means. (2014). *Building Construction Costs with RSMeans Data*. Gordian.

RS Means. (2017). *Heavy Construction Costs with RSMeans Data*. Gordian.

Saharidis, G. K. D., Minoux, M., & Ierapetritou, M. G. (2010). Accelerating Benders method using covering cut bundle generation. *International Transactions in Operational Research*, *17*(2), 221–237. https://doi.org/10.1111/j.1475-3995.2009.00706.x

Santoso, T., Ahmed, S., Goetschalckx, M., & Shapiro, A. (2005). A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, *167*(1), 96–115. https://doi.org/10.1016/j.ejor.2004.01.046

Surface Transportation Board. (2003). *Public Use File (Public Use Carload Waybill Sample) text file*.

United States Department of Agriculture. (2001). *Grain Transportation Report Datasets*. https://www.ams.usda.gov/services/transportation-analysis/gtr-datasets

United States Department of Transportation BTS. (2012). *Domestic Freight Ton-Miles | Bureau of Transportation Statistics*. https://www.bts.gov/archive/publications/transportation_statistics_annual_report/2003/chapter_02/domestic_freight_ton-miles

United States Department of Transportation Bureau of Transportation Statistics. (2019). *Freight Facts and Figures*. https://www.bts.gov/product/freight-facts-and-figures

United States Department of Transportation Bureau of Transportation Statistics. (2021). *Ports*. https://data-usdot.opendata.arcgis.com/datasets/ports?geometry=3.395%2C-16.827%2C-2.582%2C72.121

University of Arkansas Division of Agriculture. (2021). *Arkansas Agriculture Profile 2021*. https://www.uaex.uada.edu/media-resources/docs/2021_AR_Ag_profile-ADA__.pdf

U.S. Army Corps of Engineers. (2017). *Final Waterborne Commerce Statistics For Calendar Year 2017*. https://usace.contentdm.oclc.org/utils/getfile/collection/p16021coll2/id/3002

U.S. Army Corps of Engineers. (2021a). *LPMS Public Web—Home*. https://corpslocks.usace.army.mil/lpwb/f?p=121:1:1

U.S. Army Corps of Engineers. (2021b). *Waterborne Commerce Statistics*. https://www.iwr.usace.army.mil/About/Technical-Centers/WCSC-Waterborne-Commerce-Statistics-Center-2/

U.S. Bureau of Transportation Statistics. (2021). *Commodity Flow Survey*. https://www.bts.gov/cfs

U.S. Census Bureau. (2021). *U.S. Port Data*. https://www.census.gov/foreign-trade/reference/products/catalog/port.html

U.S. Committee on the Marine Transportation System. (2022). *What We Do*. https://www.cmts.gov/about-us/

A1. Appendix A Dual Problem

$$(DS-WSN(s)): max \sum_{j\in J}\sum_{c\in C}\sum_{p\in P}\sum_{s\in S}(v\_jcps\ q\_jcps\ +\xi\_jcpsd\_jcps) + \sum_{i\in I}\sum_{e\in E}\sum_{p\in P}\sum_{s\in S}\pi_{ipes}m_e(n_{ie}+Z_{ie}) + \sum_{i\in I}\sum_{f\in F}\sum_{p\in P}\sum_{s\in S}\sigma_{ipfs}l_f(k_{if}+Y_{if}) \quad 24$$

subject to:

$$\nu_{jcps}+\upsilon_{icps}+\Lambda_{ec}\pi_{ipes}\le\alpha_{ij}^{t}, \forall i\in I', j\in J', c\in C, e\in E, p\in P, s\in S, \beta_{ec}=1 \quad 25$$

$$\nu_{jcps}+\upsilon_{icps}+\Lambda_{ec}\,\pi_{ipes}\le\alpha_{ij}^{r}, \forall i\in I', j\in J', c\in C, e\in E, p\in P, s\in S, \beta_{ec}=1, \gamma_j=1, \delta_i=1 \quad 26$$

$$\upsilon_{icp+1s}-\upsilon_{icps}+\zeta_{fc}\sigma_{ipfs}\le h_c, \forall i\in I', c\in C, f\in F, p\in 1..P-1, s\in S, \Gamma_{fc}=1 \quad 27$$

$$-\upsilon_{icps}+\zeta_{fc}\sigma_{ipfs}\le h_c, \forall i\in I', c\in C, f\in F, p\in |P|, s\in S, \Gamma_{fc}=1 \quad 28$$

$$\chi_{kcps}-\upsilon_{icps}+\Lambda_{ec}(\pi_{ipes}+\pi_{kpes})\le a_{ik}, \forall i\in I', k\in I'', c\in C, e\in E, p\in P, s\in S, \beta_{ec}=1 \quad 29$$

$$\chi_{icp+1s}-\chi_{icps}+\zeta_{fc}\sigma_{ipfs}\le h_c, \forall i\in I'', c\in C, f\in F, p\in 1..P-1, s\in S, \Gamma_{fc}=1 \quad 30$$

$$-\chi_{icps}+\zeta_{fc}\sigma_{ipfs}\le h_c, \forall i\in I'', c\in C, f\in F, p\in |P|, s\in S, \Gamma_{fc}=1 \quad 31$$

$$\xi_{jcps}-\chi_{icps}+\Lambda_{ec}\pi_{ipes}\le\alpha_{ij}^{t}, \forall i\in I'', j\in J'', c\in C, e\in E, p\in P, s\in S, \beta_{ec}=1 \quad 32$$

$$\xi_{jcps}-\chi_{icps}+\Lambda_{ec}\pi_{ipes}\le\alpha_{ij}^{r}, \forall i\in I'', j\in J'', c\in C, e\in E, p\in P, s\in S, \beta_{ec}=1, \gamma_j=1, \delta_i=1 \quad 33$$

$$\nu_{jcps}+\xi_{jcps}\le l_{jm}^{t}, \forall j\in J', m\in J'', c\in C, p\in P, s\in S \quad 34$$

$$\nu_{jcps}+\xi_{jcps}\le l_{jm}^{r}, \forall j\in J', m\in J'', c\in C, p\in P, s\in S, \gamma_j=1, \gamma_j'=1 \quad 35$$

$$\xi_{jcps}\le\mu, \forall j\in J, c\in C, p\in P, s\in S \quad 36$$

$$\nu_{jcps}, \pi_{ipes}, \sigma_{ipfs}\in R_- \quad 37$$

$$\xi_{jcps}, \upsilon_{icps}, \chi_{icps}\in free \quad 38$$

A2. Data

Table 5 Transportation Cost for Truck, Rail and Barge

Transportation Mode	Parameters	Value
Truck	α_{ij}^{t} , l_{jm}^{t}	\$0.185/ton-mile
Rail	α_{ij}^{r} , l_{jm}^{r}	\$22.65/ton+\$0.033/ton-mile
Barge	a_{ik}^{r}	\$0.0089/ton-mile

Table 6 Equipment Cost

Equipment	Cost (\$)	Specification
Conveyor	18,723	36 inch
Crane	300,000	65 ton
Hopper	18,723	25 ton
Forklift	96,738	

Table 7 Storage Facility Cost

Storage facility	Cost (\$)	Specification
Grain elevator	227,866	650000 bushels
Unpaved storage	692,769	547000 ft^2
Paved storage	307,065	144000 ft^2
Warehouse	5,663,854	77000 ft^2
Chemical/Petroleum storage tank	1,109,090	125000 barrels

Table 8 Equipment Processing Capacity (ton/month)

Port	Crane, Conveyor, Hopper, Forklift	Crane, Forklift	Petroleum tank	Chemical tank
1	32,400	0	0	7,624
2	34,425	0	0	0
3	30,375	0	0	0
4	30,000	0	0	0
5	0	150,000	0	7,624
6	0	0	0	0
7	30,000	0	0	0
8	50,250	9,300	0	0
9	30,000	0	0	0
10	0	210,000	0	0
11	30,000	0	0	0
12	30,375	0	0	0
13	38,700	200,700	0	0
14	0	0	38,120	0
15	129,300	21,600	7,624	0

Table 8 Equipment Processing Capacity (ton/month)

Port	Crane, Conveyor, Hopper, Forklift	Crane, Forklift	Petroleum tank	Chemical tank
16	0	0	0	22,872
17	105,000	0	0	0
18	0	58,500	0	0
19	26,250	0	0	0
20	52,500	0	0	0
21	30,000	0	0	0
22	26,250	0	0	0
23	4,050	0	0	0
24	52,500	0	0	0
25	15,000	0	0	0
26	0	90,000	0	0
27	76,650	30,000	0	0
28	20,250	0	0	0
29	34,425	0	0	0
30	105,000	0	0	0

Table 9 Storage Facility Capacity (ton)

Port	Grain elevator	Unpaved storage	Paved storage	Warehouse	Chemical storage tank	Petroleum storage tank
1	118,800	18,687	0	4,182	0	3,600
2	15,984	0	0	0	0	0
3	61,992	0	0	0	0	0
4	11,556	0	0	0	0	0
5	0	0	0	15,410	0	0
6	0	0	0	0	0	26,250
7	11,214	0	0	0	0	0
8	324	0	0	48,956	0	0
9	0	176,380	0	0	0	0
10	0	115,352	0	3,679	0	0
11	0	0	0	4,594	0	0
12	113,400	0	0	0	0	0
13	0	143,749	191,602	5,906	0	0
14	0	0	0	0	29,700	0
15	56,700	0	5,748,048	10,731	7,950	0
16	0	0	0	0	0	27,300
17	0	261,766	0	0	0	0
18	0	9,793	45,646	0	0	0
19	0	48,243	0	0	0	0
20	0	50,614	188,185	0	0	0
21	13,500	0	0	1,254	0	0
22	0	0	316,559	0	0	0

Table 9 Storage Facility Capacity (ton)

Port	Grain elevator	Unpaved storage	Paved storage	Warehouse	Chemical storage tank	Petroleum storage tank
23	17,550	0	0	10,073	0	0
24	0	1,069,815	0	0	0	0
25	0	0	1,322,985	4,534	0	0
26	0	0	0	10,047	0	0
27	0	0	492,369	13,922	0	0
28	22,950	0	0	0	0	0
29	0	125,172	0	0	0	0
30	0	0	0	25,988	0	0