A Two Stage Stochastic Optimization Model for Port Infrastructure Planning

Sanjeev Bhurtyal (sbhurtya@uark.edu)
University of Arkansas at Fayetteville

Sarah Hernandez
University of Arkansas at Fayetteville

Sandra Eksioglu
University of Arkansas at Fayetteville

Manzi Yves
University of Arkansas at Fayetteville

Research Article

Keywords: Inland waterway port, Benders decomposition algorithm, Stochastic programming, Port infrastructure investment

Posted Date: January 30th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2511701/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.
A Two Stage Stochastic Optimization Model for Port Infrastructure Planning

Corresponding Author:

Sanjeev Bhurtyal
Department of Civil Engineering
University of Arkansas
4190 Bell Engineering Center, Fayetteville, Arkansas, 72701
Email: sbhurtya@uark.edu
Orcid: 0000-0002-6289-1625

Sarah Hernandez
Associate Professor
Department of Civil Engineering
University of Arkansas
4190 Bell Engineering Center, Fayetteville, Arkansas, 72701
Email: sarahvh@uark.edu
Orcid: 0000-0002-4243-1461

Sandra Eksioglu
Professor
Department of Industrial Engineering
University of Arkansas
4207 Bell Engineering Center, Fayetteville, Arkansas 72701
Email: sandra@uark.edu
Orcid: 0000-0002-6674-2133

Manzi Yves
University of Arkansas
Department of Civil Engineering
4190 Bell Engineering Center, Fayetteville, Arkansas, 72701
Email: Ymanzi@garverusa.com

Acknowledgements
This work is supported by a research grant (Grant number 69A3551747130) from the Maritime Transportation Research and Education Center (MarTREC).

Word Count: 8023 words

Submitted to Transportation on January 24, 2023.
ABSTRACT
This paper investigates inland port infrastructure investment planning under uncertain commodity demand conditions. A two-stage stochastic optimization is developed to model the impact of demand uncertainty on infrastructure planning and transportation decisions. The two-stage stochastic model minimizes the total expected costs, including the capacity expansion investment costs associated with handling equipment and storage, and the expected transportation costs. To solve the problem, an accelerated Benders decomposition algorithm is implemented. The Arkansas section of the McClellan - Kerr Arkansas River Navigation System (MKARNS) is used as a testing ground for the model. Results show 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 is able to identify a cluster of ports that should receive investment in port capacity under any investment scenario. The use of a stochastic approach is justified by calculating the value of the stochastic solution (VSS).

Keywords: Inland waterway port, Benders decomposition algorithm, Stochastic programming, Port infrastructure investment
1 INTRODUCTION

Barge shipments via inland waterways are one of the most efficient modes of freight transportation. A standard 15-barge tow can move about 22,500 tons of commodities, which is equivalent to 225 rail cars or 870 tractor-trailer trucks. Similarly, it takes a gallon of fuel to ship one ton of cargo 59 miles by truck, 202 miles by rail, and 514 miles by barge (USACE 2000). Hence, freight shipped by waterways tend to be more fuel efficient and have less impacts on the environment compared to other modes of transportation (USACE 2000). Leveraging inland waterways for freight shipments could mitigate the negative impacts associated with increasing commodity flow in the US by alleviating landside bottlenecks.

Inland waterway ports are a critical part of the national freight network, since they carry near half of the domestic waterborne freight in the US (USACE 2021a). 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 challenging to understand ports’ commodity throughput and capacity. Such data are valuable for infrastructure investment planning, such as adding capacity to port servicing roads, waterway dredging schedules, lock and dam maintenance programming. Furthermore, port level data enables port agencies and authorities to identify ports needing 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 (Hu et al. 2021). However, statistics are limited to the top 25 ports, and only few inland waterway ports are included. Additional data include the US Port Data (US Census Bureau 2021), Commodity Flow Survey (CFS) (Bureau of Transportation Statistics 2021) and Waterborne Commerce Statistics (USACE 2021c). More recently, the US Army Corps of Engineers (USACE) has made available monthly commodity flow data at lock locations through the Lock Performance Monitoring System (LPMS) (USACE 2021b). However, data from above mentioned sources are constrained in their spatial and/or temporal disaggregation making it challenging to collect port specific data for inland waterways. This study addresses this limitation with the formulation of an optimization model that relies on public data sources to derive investment decisions at port level.

While collecting data at locks and dams may address inland waterway performance related questions, it alone does not answer port specific investment questions. 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. Therefore, it would be difficult to strategically guide capacity expansion investments without additional 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 such as a wide investment across many ports.

Another facet of strategic decision-making in this context is the potential for change in port usage as a function of overall freight shipment demand. Thus, decisions about investments in port infrastructure, as for any transportation investment, should be evaluated for different scenarios of freight demand that reflect the unknown nature of economically driven trends seen for freight transport. This calls for an investment model that considers multiple scenarios and provides optimal inland waterway port infrastructure investment solutions when uncertainty is present.

As such, the objective of this research is to develop a modeling tool to guide strategic decisions about investments in waterway transportation infrastructure. To this end, 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 use of 2-SOP model, rather than a deterministic model, avoids misestimation of supply chain costs when stochastic parameters exist. In addition, 2-SOP problem...
mimics reality in which decisions regarding infrastructure investments are made first and without full
information about future freight flows (uncertainty).

The key contributions of this paper are (i) a multi-modal, multi-period, and multi-commodity
optimization model for port-specific investment decision under stochastic commodity demand; (ii) an
application of the model using real data from an inland waterway; and (iii) managerial insights on
investment decisions for port authorities and stakeholders.

The remainder of this paper is structured as follows. A review of prior research related to inland
waterway port infrastructure is presented followed by the formulation of a two-stage stochastic
optimization model. Next the paper discusses an application of the proposed model to the McClellan-Kerr
Arkansas River Navigation System (MKARNS). The final section provides the main contributions,
conclusions, and limitations of the paper while highlighting future work.

2 LITERATURE REVIEW

This section reviews literature pertaining to inland waterway infrastructure planning and two-stage
stochastic optimization models for transportation planning. This section also discusses investment models
developed for inland waterway infrastructure, highlighting research gaps.

2.1 Inland Waterway Infrastructure

Research on investment planning of inland waterways primarily focuses on the allocation and scheduling
of inland waterway maintenance. Mahmoudzadeh et al. (2021) proposed a two-stage deterministic model
that allocates funds for channel dredging and lock maintenance by minimizing commodity transportation
cost required to serve the demand. In a similar context, Mitchell, Wang, and Khodakarami (2013)
formulated integer programming models to strategically fund dredging projects considering the
interdependencies of the project. The model’s solution was shown to increase commodity throughput.
Ratick and Morehouse Garriga (1996) developed a risk-based spatial decision support system to address
trade-offs in cost and reliability. The model schedules dredging such that channel reliability is maximized
in any period. These studies focus on channel and lock maintenance to increase throughput; however, the
increased throughput must be processed at ports. If infrastructure within a port (i.e., cranes, lifts,
elevators, berths, etc.) is not adequate to handle the increased throughput, then delays may be introduced,
which in turn increases delays at locks. Therefore, to compliment and improve the existing suite of inland
waterway planning models, there is a need for decision models to determine optimal port infrastructure
investment.

Along the line of port-specific infrastructure investment, the work by Lagoudis, Rice, and
Salminen (2014) proposes an investment decision-making process for port infrastructure investment
considering uncertainty in commodity throughput. The model identifies the port’s most profitable
infrastructure and ranks investment strategies. This study focuses on profitability rather than
serviceability. Considering higher profit does not necessarily mean better performance of the
transportation system, there is a need to focus on optimizing system performance. Koh (2001) developed
a heuristic algorithm that determines optimal locations, size, and timing of inland container port
developments. The infrastructure needed by ports that handle containers (cranes) differs from that of ports
that handle dry bulk (grain elevators). Therefore, the models developed for inland container ports may not
apply to ports along inland waterways. Whitman, Baroud, and Barker (2019) integrated a dynamic risk-
based interdependency model with weighted multi-criteria decision analysis techniques to rank
investment allocation strategies based on economic loss due to disruptive events. The model determines
which commodity to allocate resources towards but does not get to the level of detail that recommends the
type of infrastructure investment that would be needed to handle that commodity. Further, since
commodity movements can share infrastructure, investment strategies based solely on commodity
movements will likely differ when considering how the infrastructure is shared across commodities. In
addition, to our knowledge, there is a gap in addressing the needs of multi-modal, multi-period, and
multi-commodity inland waterways port infrastructure investment decisions.
Two-stage stochastic optimization models for transportation planning

Two-stage stochastic optimization models are used in a wide array of transportation planning problems, including disaster response (Barbarosoğlu and Arda 2004), supply chain management (Marufuzzaman, Eksioglu, and Huang 2014), and pavement maintenance and rehabilitation (Ameri et al. 2019). Capacity, supply, demand, technology advancement, and budget are widely used stochastic parameters. Similarly, two-stage stochastic models are used to solve waterway transportation-related problems related to empty container management (Mittal et al. 2013), storage capacity optimization (Liu et al. 2020), port management (Aghalari, Nur, and Marufuzzaman 2020; 2021), and biomass supply chain (Marufuzzaman and Ekşioğlu 2017). We focus our discussion on studies related to inland waterway ports.

Aghalari, Nur, and Marufuzzaman (2021) provided a two-stage mixed-integer linear programming model to optimize the assignment of towboats and barges within a inland waterway system. This model assumes stochastic supply of commodities and fluctuation of water levels. The model minimizes the sum of the expected operational cost of towboats and barges, and the cost of inventory, transportation, processing, and shortage of commodity. Similarly, Aghalari, Nur, and Marufuzzaman (2021) proposed a two-stage stochastic programming model for inland waterway barge and towboat management considering perishable commodities with the aim of minimizing the operational cost of towboats and barges and the transportation cost. Mittal et al. (2013) proposed a two-stage stochastic model to determine an optimal number of inland depots that should be opened within a 10-year time frame while minimizing system cost of repositioning empty containers while satisfying supply and demand requirements. Marufuzzaman and Ekşioğlu (2017) developed a dynamic multimodal transportation network design that copes with fluctuating biomass supply. The studies by Aghalari, Nur, and Marufuzzaman (2020; 2021) focused on port management, while the studies by Marufuzzaman and Ekşioğlu (2017) and Mittal et al. (2013) focused on the identification of biorefineries and inland container depot location, respectively. In synthesizing the above-mentioned studies, there is a research gap regarding application of two-stage stochastic models to inland port-specific infrastructure investment planning. The two stage stochastic model compliments applications in inland waterway infrastructure planning where infrastructure investment decisions are made prior to the realization of uncertainty. Hence, such models are deemed appropriate for port infrastructure planning. Likewise, since uncertainty in volume of commodity is a deciding factor in port capacity expansion decisions, it is vital to consider it within investment decision models.

Research gap

Our work makes a distinct contribution to the studies mentioned above. First, this study considers the multi-modal, multi-period, and multi-commodity nature of inland waterways simultaneously to develop a stochastic port-specific investment decision model. Second, this study captures the uncertainty in the volume of commodities moved via the waterways, which is critical for effective port infrastructure investment planning. While studies attempted to develop a stochastic model for port management and scheduling, no prior attempt has been found for multi-modal, multi-period, and multi-commodity stochastic modelling applied to port-specific infrastructure decisions.

PROBLEM DESCRIPTION AND MATHEMATICAL MODEL FORMULATION

We consider port capacity to include physical space such as land for inventory storage and operational equipment such as cranes and grain elevators. The model determines port infrastructure investments and commodity flow such that the total of the investment cost and the expected transportation cost are minimized within a specified investment budget.
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. The set $C = \{1,2,3, \ldots, |C|\}$ represents the set of commodity groups transported along network $W$ over a set time period $P = \{1,2, \ldots, |P|\}$.

As an example, consider a supply chain network consisting of three counties that supply commodities via rail, truck and waterways (Figure 1). Two ports receive the commodities from the counties and ship them via barges to two destination ports. Finally at the destination port, the commodity is shipped to the destination county by truck or rail where there is corresponding demand for that commodity. Any unmet demand within a state boundary is satisfied from other states.

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

### 3.1 Sets:
- $I', I''$ 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, $s \in S$

### 3.2 Problem parameters:
- $\omega_s$ is probability of occurrence of scenario $s$
- $\kappa_e$ is the cost of equipment $e$
- $\eta_f$ is the cost of storage facility $f$
- $b$ is the available budget
\( \alpha_{ij} \) and \( \alpha_{ij} \) are the unit cost of transporting commodity via truck and rail, respectively 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 \( c \)

\( \mu \) is the unit penalty cost for unmet demand (in $/ton)

\( l_{jm}^t \) and \( l_{jm}^r \) are the unit cost for transporting commodity via truck and rail, respectively, between county \( j \in J' \) and \( m \in J'' \) (in $/ton)

\( q_{jcs} \) and \( d_{jcs} \) are the supply and the demand of commodity \( c \) in county \( j \in J' \) in month \( p \) in scenario \( s \) (in tons)

\( \iota_{fc} \) is the normalized tonnage of commodity \( c \) for inventory in storage facility \( f \)

\( \Lambda_{ec} \) 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 \( i \in I \)

\( \iota_{icpfs} \) is the tonnage of inventory of commodity \( c \) in origin port \( i \in I' \) in storage facility \( f \) in month \( p \) in scenario \( s \)

\( \iota_{icpkes} \) is the tonnage of inventory of commodity \( c \) in origin port \( i \in I' \) in storage facility \( f \) in month \( p \) in scenario \( s \)

\( \iota_{icpkes} \) 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 \( p \)

\( \iota_{icpfs} \) is the tonnage of inventory of commodity \( c \) in destination port \( i \in I'' \) in storage facility \( f \) in month \( p \) in scenario \( s \)

\( \iota_{ijcpes} \) and \( \iota_{ijcpes} \) are the tonnage of commodity \( c \) processed using equipment \( e \) and transported from port \( i \in I'' \) to county \( j \in J'' \) via truck and rail, respectively, in month \( p \) in scenario \( s \)

\( Q_{jcs} \) is the tonnage of commodity \( c \) shortage in county \( j \in J'' \) in month \( p \) in scenario \( s \)

3.3 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 \( i \in I \)

\( X_{ijcpe}^t \) and \( X_{ijcpe}^r \) are the tonnage of commodity \( c \) shipped via truck and rail, respectively, 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 \( s \)

\( W_{icpkes} \) 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 \( s \)

\( R_{ijcpes}^t \) and \( R_{ijcpes}^r \) are the tonnage of commodity \( c \) processed using equipment \( e \) and transported from port \( i \in I'' \) to county \( j \in J'' \) via truck and rail, respectively, in month \( p \) in scenario \( s \)

\( O_{ijmcps} \) and \( O_{ijmcps} \) are the tonnage of commodity \( c \) shipped from county \( j \in J' \) to county \( m \in J'' \) via truck and rail, respectively, in month \( p \) in scenario \( s \)

\( Q_{jcs} \) is the tonnage of commodity \( c \) shortage in county \( j \in J'' \) in month \( p \) in scenario \( s \)

Our proposed mathematical model is defined as follows,

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

subject to:

\[
\sum_{i \in I} \sum_{e \in E} \kappa_e Z_{ie} + \sum_{i \in I} \sum_{f \in F} \iota_f Y_{if} \leq b
\]

\( Z_{ie} \in Z^+, \ \forall i \in I, e \in E \)

\( Y_{if} \in Z^+, \ \forall i \in I, f \in F \)
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 to invest in storage facilities and operational equipment, respectively. Let $Y = \{Z_{ie}, Y_{if} | i \in I, e \in E, f \in F\}$ represent solutions to problem (1) - (4). For a given value $\bar{Y} \in Y$, and a realization $d$ of random demand $\bar{d}$, the following is the formulation of second stage problem $H(\bar{Y}, d)$.

$$
H(\bar{Y}, d) = \min \sum_{i \in I'} \sum_{f \in F} \left( \sum_{e \in E} \left( \sum_{j \in J'} \left( \alpha_{ij}^t X_{jicpe}^t + \alpha_{ij}^r X_{jicpe}^r \right) + \sum_{i \in I'} \sum_{f \in F} h_c U_{icpe} \right) + \sum_{i \in I'} \sum_{k \in I''} \left( \sum_{e \in E} \left( \alpha_{ik} W_{ikcpe} + \sum_{i \in I'} \sum_{f \in F} h_c V_{icpe} \right) + \sum_{i \in I'} \sum_{j \in J'} \sum_{e \in E} \left( \alpha_{ij}^t R_{jicpe}^t + \alpha_{ij}^r R_{jicpe}^r \right) \right) \right) \\
+ \sum_{j \in J'} \sum_{m \in J''} \left( \sum_{l \in D_{jm}} O_{jm} f_{mcpe} + \sum_{l \in D_{jm}} O_{jm} R_{mcpe} \right) + \sum_{i \in I'} \sum_{f \in F} \mu \star Q_{icpe} 
$$

subject to:

$$
\sum_{m \in J'} \left( O_{jm} f_{mcpe} + O_{jm} R_{mcpe} \right) + \sum_{i \in I'} \sum_{f \in F} \left( X_{jicpe}^t + X_{jicpe}^r \right) \leq q_{icpe} \forall j \in J', c \in C, p \in P \\
Q_{mcpe} + \sum_{i \in I'} \sum_{f \in F} \left( R_{icpe}^t + R_{icpe}^r \right) + \sum_{j \in J'} \left( O_{jm} f_{mcpe} + O_{jm} R_{mcpe} \right) = d_{mcpe} \forall m \in J', c \in C, p \in P \\
\sum_{j \in J'} \sum_{e \in E} \left( X_{jicpe}^t + X_{jicpe}^r \right) + \sum_{f \in F} U_{icpe-1f} = \sum_{f \in F} U_{icpe} + \sum_{k \in I''} \sum_{e \in E} W_{kcpe} \forall i \in I', c \in C, p \in P \\
\sum_{i \in I'} \sum_{f \in F} W_{icpe} + \sum_{f \in F} V_{mcpe-1f} = \sum_{f \in F} V_{mcpe} + \sum_{m \in J''} \sum_{e \in E} \left( R_{kmcpe}^t + R_{kmcpe}^r \right) \forall k \in I'', c \in C, p \in P \\
\sum_{e \in E} \left( \sum_{j \in J'} \left( R_{jicpe}^t + R_{jicpe}^r \right) + \sum_{k \in I''} \left( W_{kicpe} + W_{kicpe} \right) + \sum_{j \in J'} \left( X_{jicpe}^t + X_{jicpe}^r \right) \right) \leq m_e (n_e + Z_{ie}) \forall i \in I, p \in P, e \in E \\
\sum_{e \in E} \xi_{icpe} \left( U_{icpe} + V_{icpe} \right) \leq l_{ie} (k_{ie} + Y_{ie}) \forall i \in I, p \in P, f \in F \\
U_{icpe} = 0 \forall i \in I', c \in C, p \in \{0\}, f \in F \\
V_{icpe} = 0 \forall i \in I'', c \in C, p \in \{0\}, f \in F \\
X_{jicpe}^t, X_{jicpe}^r, U_{icpe}, W_{kicpe}, V_{icpe}, R_{jicpe}^t, R_{jicpe}^r, O_{mcpe}^t, O_{mcpe}^r, Q_{icpe} \in R^+ \forall i \in I', k \in I'', j \in J', m \in J'', c \in C, p \in P, f \in F 
$$

Function (5) minimizes the total supply chain costs. These costs include shipping costs via truck, rail, and barges, inventory cost, and a penalty cost for any unturned demand. Constraint (6) ensures that the amount of commodity shipped does not surpass the quantity of supply available. Constraint (7) captures the shortage of commodity in cases when demand exceeds supply. Constraints (8) and (9) are the flow balance constraints for origin and destination ports. Constraint (10) ensures that the commodity handled
at origin and destination ports does not exceed the port capacity. Constraint (11) ensures that the total inventory at a port does not exceed the holding capacity of that port. Constraints (12) and (13) assign initial inventory at the destination and origin ports, respectively, as zero. Constraint (14) includes the non-negativity constraints.

4 SOLUTION APPROACH

The computational burden from solving the problem formulated in (1) - (14) warrants use of advanced solution approaches such as the Benders Decomposition algorithm. This approach is discussed in this section along with techniques to accelerate its convergence rate.

4.1 Benders Decomposition Algorithm

The uncertainty in commodity demand represents the stochastic parameter in our formulation and requires the introduction of demand scenarios into the modeling approach. We approximate the distribution of stochastic demand (level of demand over time) via a discrete distribution. Let $S$ represent the discrete set of demand realizations 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, the Benders decomposition algorithm (Benders 1962) is employed. This algorithm is widely used to solve large-size, mixed integer linear problems. We first decompose the original problem into two subproblems: an integer master problem (MP) ($M$-WSN) and $|S|$ linear subproblems ($S$-WSN($s$)). The MP along with an auxiliary variable and optimality cut provides an approximation of the original problem (1) - (14).

Let $\bar{Y}$ be the solution of MP ($M$-WSN). For the given $\bar{Y}$, $|S|$ subproblems are solved, one for each realization $d_s$ for $s \in S$ of the stochastic demand $\bar{d}$. 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 $\bar{Y}$ 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 objective value of ($M$-WSN) provides a lower bound, and the solution of ($M$ - WSN) and ($S$ - WSN($s$)) 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.

$$(M - WSN): \min \sum_{i \in I} \kappa_e Z_{ie} + \sum_{i \in I, f \in F} \gamma_f Y_{if} + \sum_{s \in S} \omega_s \theta_s$$

Subject to: (2-4)

$$\theta_s^n \geq \sum_{j \in J} \sum_{c \in C} \sum_{p \in P} (\gamma_{jcps} q_{jcp} + \xi_{jcps} d_{jcps}) + \sum_{i \in I} \sum_{e \in E} \sum_{p \in P} \pi_{ipes} m_e (n_{ie} + Z_{if})$$

$$+ \sum_{i \in I} \sum_{f \in F} \sum_{p \in P} \sigma_{ipfs} I_f (k_{if} + Y_{if}) \forall s \in S, n = \{1, \ldots, N'\}$$

where $N'$ is the current number of iterations.

The Benders decomposition solves ($M$ - WSN) iteratively and in each iteration $n$, $\bar{Y}^n$ represents the corresponding solution. $\Phi^n$ is the objective function value of model (15) to (16) obtained at the $n^{th}$ iteration according to the formulation in equation (17). $\Phi^n$ represents the corresponding cost of infrastructure investment.

$$\Phi^n = \sum_{i \in I, e \in E} \kappa_e Z_{ie}^n + \sum_{i \in I, f \in F} \gamma_f Y_{if}^n$$
Given \( \overline{y^n} \), the following scenario-based subproblem \((S − WSN(s))\) is solved for each demand scenario \( s ∈ S \).

\[
(S − WSN(s)) : \min \sum_{e ∈ E} \sum_{p ∈ P} \left( \sum_{i ∈ I'} \sum_{j ∈ J'} \sum_{e ∈ E} (\alpha_{ij}^t X_{j_{ips}}^e + \alpha_{ij}^r X_{j_{ips}}^r) + \sum_{i ∈ I'} h_c U_{icpfs} \right) \\
+ \sum_{i ∈ I''} \sum_{j ∈ J} \sum_{e ∈ E} (\alpha_{ij}^l W_{l_{ips}} + h_c V_{icpfs}) \\
+ \sum_{i ∈ I''} \sum_{j ∈ J'} \sum_{e ∈ E} (\alpha_{ij}^l R_{i_{ips}} + \alpha_{ij}^r R_{i_{ips}}) + \sum_{j ∈ J'} \sum_{l ∈ E} \{ (l_j^l O_{lm_{ips}} + l_j^r O_{lm_{ips}}) \} \\
+ \sum_{j ∈ J'} \mu * Q_{j_{ips}})
\]

Subject to: (6) – (14)

Constraint (16) is the scenario specific optimality cut added to \((M − WSN)\) in each iteration of the Benders decomposition algorithm. \( ν_{j_{ips}} \forall j ∈ J', \xi_{j_{ips}} \forall j ∈ J'' \), \( ν_{icpfs} \forall i ∈ I', \chi_{icpfs} \forall i ∈ I'' \), \( π_{ipeps} \), \( σ_{ipfs} \) are dual variables for constraints (6) to (10), respectively. Let \( X^n(s) \) denote this solution for scenario \( s \) at the \( n^{th} \) iteration. Constraint (7) makes the scenario-based subproblem \((S − WSN(s))\) always feasible for any values of the 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 \((S − WSN(s))\) is shown in Appendix A.

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

\[
Θ^n = \sum_{s ∈ S} ω_s Θ^s_n \]

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 \( \overline{y^n} \)

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 \( \frac{UB^n − LB^n}{UB^n} ≤ ϵ \) then

\( abort ← true \)

else

Add cut (16) to \((M − WSN)\)

\( n ← n + 1 \)

end

end

return \( UB, \overline{y^n} \) and \( X^n(s) \)
4.2 Methods to Accelerate Benders Decomposition Algorithm

The Benders decomposition algorithm is known to be extremely slow and computationally expensive. In our study we add Knapsack inequalities and Pareto-optimal cuts to improve the convergence rate of the algorithm. The details of the proposed techniques are provided in later parts of this section.

4.2.1 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). Let \(LB^n\) denote the lower bound obtained in the \(n^\text{th}\) iteration of the algorithm. Therefore, the following knapsack inequalities are added to \((M-WSN)\) in the \((n+1)^\text{th}\) iteration to accelerate the convergence rate of the Benders decomposition algorithm:

\[
LB^n \leq \sum_{i \in I, e \in E} \kappa_e \cdot Z_{ie} + \sum_{i \in I, f \in F} \lambda_f \cdot Y_{if} + \sum_{s \in S} \omega_s \cdot \theta_s
\]

4.2.2 Pareto-optimal cuts

The subproblems \((S-WSN(s))\) are capacitated transportation problems. The transportation problem is degenerate in nature (Ahuja, Magnanti, and Orlin 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 and 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 and 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, \(\overline{Y}^n\) be the solution of the MP at the \(n^\text{th}\) iteration and \(\overline{Y}^0 = \{ Z_{ie}^0 \mid n = 0, i \in I, e \in E, f \in F \}\) be the set of initial core points. The subproblem to generate Pareto-optimal cuts is given as:

\[
(MWS - WSN(s)): \max \sum_{j \in J} \sum_{c \in C} \sum_{p \in P} \sum_{s \in S} (v_{jpcs}q_{jpcs} + \xi_{jpcs}d_{jpcs}) + \sum_{i \in I} \sum_{e \in E} \sum_{p \in P} \sum_{s \in S} \pi_{ipes} m_e(n_{ie} + Z_{ie}^n_i) + \sum_{i \in I} \sum_{f \in F} \sum_{p \in P} \sum_{s \in S} \sigma_{ipfs} l_f(k_{if} + Y_{if}^n) + \sum_{i \in I} \sum_{f \in F} \sum_{p \in P} \sum_{s \in S} \sigma_{ipfs} l_f(k_{if} + Y_{if}^n) = \Theta^n(s) \forall s \in S
\]

subject to: (A.1) - (A.15) (Appendix A)

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 in each iteration, constraint (22) could be ignored. Pareto-optimal cuts generated by this method are known as modified Magnanti-Wong pareto-optimal cuts. The core points are updated in every iteration as follows.

\[
Z_{ie}^{o,n+1} = (1 - \lambda)Z_{ie}^{o,n} + \lambda Z_{ie}^n
\]

\[
Y_{if}^{o,n+1} = (1 - \lambda)Y_{if}^{o,n} + \lambda Y_{if}^n
\]
Papadakos (2008) and Mercier, Cordeau, and Soumis (2005) empirically showed that the value of \( \lambda = 0.5 \) 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 \( \epsilon, Z^{o,n}_{ie}, Y^{o,n}_{if} \). Set \( n \leftarrow 1, LB^n \leftarrow -\infty, UB^n \leftarrow +\infty, \) abort \( \leftarrow \) false

while abort \( = \) false do

For all \( s \in S \), Solve \((MWS - WSN(s))\) to obtain \( \Theta^n_s \) and \( \overline{y^n}(s) \)

Add cuts () to \((M - WSN)\)

Solve \((M - WSN)\) to obtain \( \Phi^n, \phi^n \) and \( \overline{y^n} \)

if \( \Phi^n > LB \) then

\[ LB^n \leftarrow \Phi^n \]

end

For all \( s \in S \), Solve \((S - WSN(s))\) to obtain \( \Theta^n_s \) and \( X^n(s) \)

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 (16) to \((M - WSN)\)

Add Knapsack inequalities (20) to \((M - WSN)\)

Update core points using equations (23) and (24)

\( n \leftarrow n + 1 \)

end

end

return \( UB, \overline{y^n} \) and \( X^n(s) \)

5 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 the performance evaluation of stochastic solutions.

5.1 Performance evaluation

We solved 15 problems corresponding to different sizes to determine the quality of the proposed algorithms (Table 1) and compare four solution approaches: (1) Gurobi solver (a state-of-the-art solver for large scale mathematical programming), (2) Benders algorithm, (3) Benders with Knapsack inequalities, (4) Benders with Knapsack inequalities and Pareto-optimal cuts.
To terminate the algorithm, we consider following stopping criterion: (i) optimality gap ≤ 1%, (ii) number of iterations ≥ 500 and (iii) algorithm run time ≥ 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. The results of the experiments are summarized by the algorithm run time in seconds \( t(s) \), the optimality gap \( \varepsilon(\%) \), and the number of iterations (n) at the time when the stop criteria is met (Table 2).

Gurobi outperforms alternative solution approaches in terms of run time and optimality gap, except for large test case sizes (8)-(10) and (12)-(15) where it fails to solve within memory limits (Table 2). The Benders decomposition algorithm solves these cases but fails to converge to solutions with a small optimality gap within the given time limit for medium and large problems. Adding Knapsack inequalities and Pareto-optimal cuts to the MP improves the solution time, optimality gap, and number of iterations, and provides solutions of high quality within the proposed threshold run time.

### Table 1 Test Case Size

| Problem Nr. | | | | | Problem Nr. | | | |
|-------------|---|---|---|---|-------------|---|---|---|
| 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 | |

### Table 2 Performance evaluation of solution approaches

| Case | Gurobi | Benders | Benders + KI | Benders + KI + PO |
|------|--------|---------|-------------|-----------------|
|      | \( t(s) \) | \( \varepsilon(\%) \) | \( t(s) \) | \( \varepsilon(\%) \) | \( n \) | \( t(s) \) | \( \varepsilon(\%) \) | \( n \) | \( t(s) \) | \( \varepsilon(\%) \) | \( n \) |
| 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 | 0.92 | 13 |
| 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 |

* Out of memory
* Representative case size for a real-world problem
5.2 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 30 freight ports (Figure 2). The Arkansas segment of MKARNS has significant economic impact in Arkansas in terms of sales, Gross Domestic Product, labor income, and jobs (Nachtmann et al. 2015).

Figure 2 Arkansas Segment of MKARNS

5.2.1 Data Description

(i) 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 to create demand scenarios in the 2-SOP model.

The Arkansas Statewide Travel Demand Model (Alliance Transportation Group and Cambridge Systematics 2012) maintained by the Arkansas Department of Transportation (ARDOT) reports amounts (in tons) of commodities shipped between counties. The total amount of commodity shipped from a county is the supply for the county and considered the origin supply. Similarly, the total amount of commodities received by a county is the demand for that county and considered the destination demand. 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.
(ii) 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 US Department of Agriculture (2021) for major cities along the 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.

(iii) Port capacity data

The National Transportation Atlas Database (NTAD), published by the BTS (USDOT 2021c) contains information on the location, commodity handling equipment, storage facility and road/rail 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. These data are supplemented and updated using information obtained from ports’ websites and satellite images.

(iv) Infrastructure costs

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; 2017) and validated through interviews with industry representatives. All costs for our study are calculated based on the 2020-dollar value.

(v) 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 (USDOT 2019), respectively.

5.2.2 Case study results

We evaluate the impacts of varying 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 3). 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 4). This decrease in unit cost can be attributed to the increased volume of commodity shipped via waterways (Figure 6) which have the lowest transportation cost compared to truck and rail. This finding is consistent with previous research that reported economic benefits from expanding port infrastructure (Dekker, Verhaeghe, and Pols 2003). From a managerial perspective, port authorities can use these findings as a part of benefit/cost analysis to justify funding for inland waterway ports.
We discuss the frequency and the range of investments at individual ports for five different investment scenarios, $2M, $4M, $6M, $8M, and $10M (Figure 5). In most of these scenarios, the model determines that investments should be made at ports located at Little Rock (central Arkansas), Delaware (midwestern section of the Arkansas river), and Fort Smith (westernmost 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 at Little Rock (port 15 in Figure 5). Across all scenarios, several ports repeatedly receive no investment e.g., Menifee, Morrilton, and Russellville. The findings suggest that the Little Rock port is an important hub of MKARNS, and this finding is corroborated by the prior investments of $960,000, $2 million and $3 million received by the Little Rock Port Authority in 2012 (US Economic Development Administration 2012), in 2020 (Delta Regional Authority 2020), and in 2021 (USDOT 2021a), respectively. From a managerial perspective, this result is crucial for port authorities to advocate for funds in the future. Similarly, port authorities can allocate those funds in the ports identified above for optimal movement of commodities through waterways in terms of unit supply chain cost as shown in Figure 4.
The volume of commodities shipped via the waterways increases as the investment in port capacity increases (Figure 6). A total of 2.74 million tons of commodities are shipped by waterways (6% of the total shipments by the three modes) when the total investment is $2 million. This volume increases to 3.47 million (8% of the 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 7). Ton-mile reflects both the volume (tons) and the distance (miles) shipped and is one of the most widely used measures of the physical volume of freight transportation services (USDOT 2021b). 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. The findings show that volume shipped via waterways and ton-miles increases with investment in port capacity expansions. Since barges have the least amount of carbon emission among other modes of transportation, from managerial perspective (USACE 2015), port stakeholders can use this finding to advocate for more funding for inland waterway port capacity expansion.

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 warehouse can be used for manufacturing and primary metal commodities. To demonstrate this modeling contribution, we present the results for four commodity groups that dominate along the Arkansas section of the MKARNS (Asborno, Hernandez, and Akter 2020; Nachtmann and Oztanriseven 2014): 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 8). 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 9). No significant impact of port infrastructure investment is seen for the other three commodity groups. This finding is crucial for the stakeholder, from a managerial perspective, to make decisions about their business approach. For example, since the findings show that the commodity transportation shifts from trucks and trains and toward barges, companies related to nonmetallic minerals can use this information to make decisions on their business model regarding the number of trucks to acquire and the number of truck drivers to recruit.

5.3 Evaluation of stochastic solutions

To demonstrate the benefit of developing the model with stochastic elements, rather than a deterministic model, we calculate the Value of Stochastic Solution (VSS) as the difference between the objective function value of the stochastic solution and the expected value solution. In addition, we calculate the Expected Value of Perfect Information to quantify the value of perfect information required to predict future with certainty. EVPI is the difference between the objective function value of the stochastic solution and the wait-and-see solutions (WSS).

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, \(\text{VSS} = 1,246M - 1,225M = 21M\). Therefore, $21M per year can be saved by solving the stochastic model compared to the corresponding deterministic model.

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 \textit{“wait-and-see solution” (WSS). Table 3 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). Therefore, the price we should be willing to pay to correctly predict future commodity demand realizations should not exceed $4M.

| Strategies                        | Value($M) |
|-----------------------------------|-----------|
| Stochastic programming            | 1,225     |
| \textit{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     |
6 CONCLUSIONS

In this paper, we develop a model to guide strategic investments in inland waterway port infrastructure investments given uncertainty in commodity demand. This study proposes a 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.

We apply the two-stage stochastic optimization model to the Arkansas section of the 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 each 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, VSS shows that a failure to use a stochastic model to capture variations in commodity demand, could cost up to $21M per year. EVPI indicates the price we should be willing to pay for correctly predicting future realizations of commodity demand, should be no more than $4M.

This research opens several avenues to explore in the future. The current model allows for analysis of the supply chain within a state, and hence in the future, the model will be expanded to incorporate multiple states. Furthermore, in our current study, there are no port disruption scenarios considered, and since closure of port due to human-induced or/and natural causes can impact port operation, the uncertainty of port disruption will be incorporated in a future study.
7 DECLARATIONS

7.1 Ethical Approval
This declaration is not applicable.

7.2 Competing Interests
The authors have no conflict of interest of a financial or personal nature to disclose.

7.3 Author’s contributions
Conceptualization: Sarah Hernandez, Sandra Eksioglu; Methodology: Sanjeev Bhurtyal, Sarah Hernandez, Sandra Eksioglu; Formal analysis and investigation: Sanjeev Bhurtyal, Yves Manzi; Original draft preparation, review and editing: Sanjeev Bhurtyal, Sarah Hernandez, Sandra Eksioglu; Funding acquisition: Sarah Hernandez, Sandra Eksioglu. All authors reviewed the manuscript.

7.4 Funding
This work is supported by a research grant (Grant number 69A3551747130) from Office of the Assistant Secretary for Research and Technology under Maritime Transportation Research and Education Center (MarTREC).

7.5 Availability of data and materials
The dataset can be made available upon request.
REFERENCES

Aghalari, Amin, Farjana Nur, and Mohammad Marufuzzaman. 2020. “A Bender’s Based Nested Decomposition Algorithm to Solve a Stochastic Inland Waterway Port Management Problem Considering Perishable Product.” International Journal of Production Economics 229 (November): 107863. https://doi.org/10.1016/j.ijpe.2020.107863.

———. 2021. “Solving a Stochastic Inland Waterway Port Management Problem Using a Parallelized Hybrid Decomposition Algorithm.” Omega 102 (July): 102316. https://doi.org/10.1016/j.omega.2020.102316.

Ahuja, Ravindra K., Thomas L. Magnanti, and James B. Orlin. 1988. Network Flows. Massachusetts.

Alliance Transportation Group and Cambridge Systematics. 2012. “Arkansas Statewide Travel Demand Model Phase I-Freight Model.” Austin, TX.

Ameri, Mahmoud, Armin Jarrahi, Farshad Haddadi, and Mohammad Hasan Mirabimoghaddam. 2019. “A Two-Stage Stochastic Model for Maintenance and Rehabilitation Planning of Pavements.” Mathematical Problems in Engineering 2019.

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

Asborno, Magdalena I., Sarah Hernandez, and Taslima Akter. 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.

Barbaroğlu, Gülay, and Yasemin Arda. 2004. “A Two-Stage Stochastic Programming Framework for Transportation Planning in Disaster Response.” Journal of the Operational Research Society 55 (1): 43–53.

Benders, J. F. 1962. “Partitioning Procedures for Solving Mixed-Variables Programming Problems.” Numerische Mathematik 4 (1): 238–52. https://doi.org/10.1007/BF01386316.

Braham, Andrew, Sarah Hernandez, John Kent, Magdalena Asborno, and Sadie Casillas. 2017. Locating Transload Facilities to Ease Highway Congestion and Safeguard the Environment.

Bureau of Transportation Statistics. 2010. “National Transportation Statistics 2010.” 2010. https://www.bts.gov/archive/publications/national_transportation_statistics/2010/index.

———. 2021. “Commodity Flow Survey.” https://www.bts.gov/cfs.

Cheung, Raymond K, and Chuen-Yih Chen. 1998. “A Two-Stage Stochastic Network Model and Solution Methods for the Dynamic Empty Container Allocation Problem.” Transportation Science 32 (2): 142–62.

Dekker, Sander, RJ Verhaeghe, and A AJ Pols. 2003. “Economic Impacts and Public Financing of Port Capacity Investments: The Case of Rotterdam Port Expansion.” Transportation Research Record 1820 (1): 55–61.

Delta Regional Authority. 2020. “DRA Announces $2 Million Investment to Strengthen Arkansas’s Transportation Infrastructure.” https://dra.gov/newsroom/press-release/dra-announces-2-million-investment-to-strengthen-arkansass-transportation-i/.

Hu, Patricia S., Rolf R. Schmitt, Long Nguyen, Matthew Chambers, and Daniel Friedman. 2021. “Port Performance Freight Statistics in 2019: Annual Report to Congress 2020.” https://rosap.ntl.bts.gov/view/dot/54022.

Koh, Yong-Ki. 2001. “Optimal Investment Priority in Container Port Development.” Maritime Policy & Management 28 (2): 109–23. https://doi.org/10.1080/03088830117187.
Lagoudis, Ioannis N., James B. Rice Jr, and Jason B. Salminen. 2014. “Port Investment Strategies under Uncertainty: The Case of a Southeast Asian Multipurpose Port.” The Asian Journal of Shipping and Logistics, Selected Papers from the IAME (International Association of Maritime Economists, Norfolk VA, USA) Conference 2014 and the ICASL (International Conference of Asian Shipping and Logistics, Kobe, JAPAN) Conference 2013, 30 (3): 299–319. https://doi.org/10.1016/j.ajsl.2014.12.003.

Liu, Zhimin, Shaojian Qu, Mark Goh, Zhong Wu, Ripeng Huang, and Gang Ma. 2020. “Two-Stage Mean-Risk Stochastic Optimization Model for Port Cold Storage Capacity under Pelagic Fishery Yield Uncertainty.” Physica A: Statistical Mechanics and Its Applications 541 (March): 123338. https://doi.org/10.1016/j.physa.2019.123338.

Magnanti, T. L., and R. T. Wong. 1981. “Accelerating Benders Decomposition: Algorithmic Enhancement and Model Selection Criteria.” Operations Research 29 (3): 464–84. https://doi.org/10.1287/opre.29.3.464.

Mahmoudzadeh, Ahmadreza, Mohammadadel Khodakarami, Chaolun Ma, Kenneth Ned Mitchell, Xiubin Bruce Wang, and Yunlong Zhang. 2021. “Waterway Maintenance Budget Allocation in a Multimodal Network.” Transportation Research Part E: Logistics and Transportation Review 146 (February): 102215. https://doi.org/10.1016/j.tre.2020.102215.

Marufuzzaman, Mohammad, Mohammad, Mohammadadel Khodakarami, Chalo Ma, Kenneth Ned Mitchell, Xiubin Bruce Wang, and Yoonlong Zhang. 2021. “Waterway Maintenance Budget Allocation in a Multimodal Network.” Transportation Research Part E: Logistics and Transportation Review 146 (February): 102215. https://doi.org/10.1016/j.tre.2020.102215.

Marufuzzaman, Mohammad, Mohammad, Mohammadadel Khodakarami, Chalo Ma, Kenneth Ned Mitchell, Xiubin Bruce Wang, and Yoonlong Zhang. 2021. “Waterway Maintenance Budget Allocation in a Multimodal Network.” Transportation Research Part E: Logistics and Transportation Review 146 (February): 102215. https://doi.org/10.1016/j.tre.2020.102215.

Marufuzzaman, Mohammad, and Sandra D. Ekşioğlu. 2014. “Two-Stage Stochastic Programming Supply Chain Model for Biodiesel Production via Wastewater Treatment.” Computers & Operations Research 49 (September): 1–17. https://doi.org/10.1016/j.cor.2014.03.010.

Marufuzzaman, Mohammad, and Sandra Duni Ekşioğlu. 2017. “Designing a Reliable and Dynamic Multimodal Transportation Network for Biofuel Supply Chains.” Transportation Science 51 (2): 494–517.

Mercier, Anne, Jean-François Cordeau, and François Soumis. 2005. “A Computational Study of Benders Decomposition for the Integrated Aircraft Routing and Crew Scheduling Problem.” Computers & Operations Research 32 (6): 1451–76. https://doi.org/10.1016/j.cor.2003.11.013.

Mitchell, Kenneth Ned, Bruce X. Wang, and Mohammadadel Khodakarami. 2013. “Selection of Dredging Projects for Maximizing Waterway System Performance.” Transportation Research Record 2330 (1): 39–46. https://doi.org/10.3141/2330-06.

Mittal, Neha, Maria Boile, Alok Baveja, and Sotiris Theofanis. 2013. “Determining Optimal Inland-Empty-Container Depot Locations under Stochastic Demand.” Research in Transportation Economics, Freight Transport and Sustainability, 42 (1): 50–60. https://doi.org/10.1016/j.retrec.2012.11.007.

Nachtmann, Heather, Othman Boudhoun, Furkan Oztanriseven, Mack-Blackwell National Rural Transportation Study Center (U.S.), and 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.

Nachtmann, Heather, and F Oztanriseven. 2014. “Economic Evaluation of Arkansas Inland Waterways and Potential Disruption Impacts.” Technical Report.

Papadakos, Nikolaos. 2008. “Practical Enhancements to the Magnanti–Wong Method.” Operations Research Letters 36 (4): 444–49. https://doi.org/10.1016/j.orl.2008.01.005.

Ratick, Samuel J., and Holly Morehouse Garriga. 1996. “Risk-Based Spatial Decision Support System for Maintenance Dredging of Navigation Channels.” Journal of Infrastructure Systems 2 (1): 15–22. https://doi.org/10.1061/(ASCE)1076-0342(1996)2:1(15).
RS Means. 2014. *Building Construction Costs with RSMeans Data*. Gordian.

———. 2017. *Heavy Construction Costs with RSMeans Data*. Gordian.

Santoso, Tjendera, Shabbir Ahmed, Marc Goetschalckx, and Alexander Shapiro. 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.” Washington, DC.

US Census Bureau. 2021. “United States Port Data.” https://www.census.gov/foreign-trade/reference/products/catalog/port.html.

US Department of Agriculture. 2021. “Grain Transportation Report Datasets.” https://www.ams.usda.gov/services/transportation-analysis/gtr-datasets.

US Economic Development Administration. 2012. “U.S. Department of Commerce Invests $6.9 Million in Arkansas to Build Infrastructure and Boost Economic and Job Growth.” https://www.eda.gov/archives/2016/news/press-releases/2012/09/05/ar.htm.

USACE. 2000. “Inland Waterway Navigation Value to the Nation.” https://www.mvp.usace.army.mil/Portals/57/docs/Navigation/InlandWaterways-Value.pdf.

———. 2015. “Fact Sheet 13: Comparing Navigation.” https://www.mvp.usace.army.mil/Media/Fact-Sheets/Fact-Sheet-Article-View/Article/588155/fact-sheet-13-comparing-navigation/.

———. 2021a. “Final Waterborne Commerce Statistics For Calendar Year 2017.” https://usace.contentdm.oclc.org/utils/getfile/collection/p16021coll2/id/3002.

———. 2021b. “LPMS Public Web - Home.” https://corpslocks.usace.army.mil/lpwb/f?p=121:1:1.

———. 2021c. “Waterborne Commerce Statistics.” 2021. https://www.iwr.usace.army.mil/About/Technical-Centers/WCSC-Waterborne-Commerce-Statistics-Center-2/.

USDOT. 2019. “Freight Facts and Figures.” https://www.bts.gov/product/freight-facts-and-figures.

———. 2021a. “2021 Port Infrastructure Development Program Grant Awards “textbar MARAD.” https://www.maritime.dot.gov/ports/port-infrastructure-development-program/2021-port-infrastructure-development-program-grant.

———. 2021b. “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.

———. 2021c. “Ports.” https://data-usdot.opendata.arcgis.com/datasets/ports?geometry=3.395%2C-32.16827%2C-2.582%2C72.121.

Whitman, Mackenzie, Hiba Baroud, and Kash Barker. 2019. “Multicriteria Risk Analysis of Commodity-Specific Dock Investments at an Inland Waterway Port.” *The Engineering Economist* 64 (4): 346–67. https://doi.org/10.1080/0013791X.2019.1580808.
Appendix A. Dual Problem

\begin{equation}
(\text{DS – WSN}(s)) \text{: max } \sum_{j \in J} \sum_{c \in C} \sum_{p \in P} \sum_{s \in S} (\nu_{jcp} q_{jcp} + \xi_{jcp} d_{jcp}) + \sum_{i \in I} \sum_{e \in E} \sum_{p \in P} \sum_{s \in S} \pi_{ipes} m_e (\pi_{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})
\end{equation}

subject to:

\begin{align}
\nu_{jcp} + \nu_{ipc} + \lambda_{ec} \pi_{ipes} & \leq \alpha_{i,j}^t, \forall i \in I', j \in J', c \in C, e \in E, p \in P, s \in S, \beta_{ec} = 1 & \text{A.2} \\
\nu_{jcp} + \nu_{ipc} + \lambda_{ec} \pi_{ipes} & \leq \alpha_{i,j}'^t, \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 & \text{A.3} \\
\nu_{ipc+1} - \nu_{ipc} + \zeta_{fc} \sigma_{ipfs} & \leq h_c, \forall i \in I', c \in C, f \in F, p \in 1..P - 1, s \in S, \Gamma_{fc} = 1 & \text{A.4} \\
\nu_{ipc+1} - \nu_{ipc} + \zeta_{fc} \sigma_{ipfs} & \leq h_c, \forall i \in I'', c \in C, f \in F, p \in |P|, s \in S, \Gamma_{fc} = 1 & \text{A.5} \\
\chi_{ipc+1} - \chi_{ipc} + \zeta_{fc} \sigma_{ipfs} & \leq h_c, \forall i \in I'', c \in C, f \in F, p \in 1..P - 1, s \in S, \Gamma_{fc} = 1 & \text{A.6} \\
\chi_{ipc+1} - \chi_{ipc} + \zeta_{fc} \sigma_{ipfs} & \leq h_c, \forall i \in I'', c \in C, f \in F, p \in |P|, s \in S, \Gamma_{fc} = 1 & \text{A.7} \\
\chi_{ipc+1} - \chi_{ipc} + \zeta_{fc} \sigma_{ipfs} & \leq h_c, \forall i \in I'', c \in C, f \in F, p \in 1..P - 1, s \in S, \Gamma_{fc} = 1 & \text{A.8} \\
\chi_{ipc+1} - \chi_{ipc} + \zeta_{fc} \sigma_{ipfs} & \leq h_c, \forall i \in I'', c \in C, f \in F, p \in |P|, s \in S, \Gamma_{fc} = 1 & \text{A.9} \\

\end{align}