Maritime Transport Research 2 (2021) 100038

Contents lists available at ScienceDirect

Maritime Transport Research

journal homepage: www.sciencedirect.com/journal/maritime-transport-research

An updatable and comprehensive global cargo maritime network and strategic seaborne cargo routing model for global containerized and bulk vessel flow estimation

Wenjie Li [a], Ralph Pundt [b], Elise Miller-Hooks [a,*]

[a] Sid & Reva Department of Civil, Environmental and Infrastructure Engineering, George Mason University, United States
[b] Maine Maritime Academy, United States

ARTICLE INFO

Keywords:
Global cargo shipping
Maritime networks
Trade
Liner shipping
Bulk shipping
Trade routes

ABSTRACT

The global maritime system provides the backbone of logistics operations for global supply chains and international trade. This paper aims to develop a unifying global network representation and strategic, system-wide decision model, the Strategic Cargo Routing Model, incorporating both liner and bulk shipping markets to estimate real-world traffic flows and study traffic patterns at the global scale. Specifically, taking a shipper's perspective, containerized and bulk movements are jointly modelled within a mixed-integer linear program that includes inbound, outbound, and transshipment cargo flows at ports. An iterative approach that combines heuristic Gradient Descent and Relax-and-Fix Decomposition methods is proposed for the calibration and solution of the Strategic Cargo Routing Model over a proposed joint liner and bulk services Global Cargo Shipping Network representation. The Global Cargo Shipping Network contains 161 seaports covering 52 countries. It is created from updatable, publicly available, data sources, and all data needed for the network representation are made available. Sufficient network details, as well as data sources and methods for extracting needed inputs, are given to allow others to use and update the network. Using the developed maritime network, mathematical model and calibration-solution methodology, 2018 global maritime traffic flow patterns were estimated. The estimates were found to achieve a 91% fit overall to real-world average annual port throughputs. This strategic model provides support to evaluate future, real-world, worldwide changes, such as increased seaborne trade demand, new routes, shipping infrastructure expansion, and transport policies.

1. Introduction

Past decades of increased globalization have led to increases in international trade and the importance of maritime shipping. As recent as early 2020 (just prior to the onset of the COVID-19 pandemic), the international shipping industry carried approximately 90% of world trade by weight (International Chamber of Shipping, 2021). In 2018, world seaborne trade reached nearly 11 billion tons (UNCTAD, 2019). Data from UNCTAD (2015) and new updates as plotted in Fig. 1 show that world trade, including world seaborne trade, outpaced two key industrial indicators, including the world gross domestic product (GDP).

The global maritime system provides the backbone of logistics operations for global supply chains and international trade. The

* Corresponding author.
E-mail address: miller@gmu.edu (E. Miller-Hooks).

https://doi.org/10.1016/j.martra.2021.100038
Received 3 June 2021; Received in revised form 31 August 2021; Accepted 1 September 2021
Available online 13 September 2021

system is complex, involving stakeholders from port authorities, ocean carriers and shippers. Its structure and the flows it carries are impacted by changing economic strategies and trade policies between nations, growth and contraction in demand, and supply-side advances aimed at meeting potentially high demand levels.

After the financial crisis in 2008, but prior to the COVID-19 pandemic of 2020, there was a positive, steady recovery in world seaborne trade (UNCTAD, 2019). To satisfy the increased demand in international shipping, in 2016 the Panama Canal was expanded. Additionally, the width and depth of vessel traffic lanes and locks were increased, thereby allowing the New Panamax vessels with capacities each of 8000 Twenty-foot Equivalent Units (TEUs) to pass through, ultimately, doubling the Canal's capacity (Maroday, 2013). Additionally, technology to support the construction of mega-ships with capacities greater than 18,000 TEUs was developed. However, due to ship construction times for such large ships (on the order of 2 years for a single ship), an imbalance between total ship capacity and demand in seaborne transport continued over recent years (UNCTAD, 2018). In response, freight rates increased. In 2019, the United States and China (accounting for around 30% of total world trade at the time) entered into a trade war, changing global seaborne trade patterns. Examples of the effects of the trade war include a ban on the shipment of steel to the United States from China and the cancelation of soybean contracts from the United States to China (Amadeo, 2020). Thus, flows between these two nations diminished. This was soon followed by unprecedented changes in maritime traffic in 2020 as the world entered a period of global pandemic. In response to the COVID-19 pandemic, the largest ocean carrier, Maersk, announced their decision to reduce their capacity of Transpacific services during this time period (Maersk, 2020).

Given these complexities and political, economic and technological dynamics, a tool for forecasting worldwide maritime shipping patterns and flows to aid the various stakeholders in strategic decision making is needed, yet very few works have sought to provide such a tool and none supports the modeling of both container and bulk shipping markets. Yet, bulk shipping plays a dominant role in carrying world seaborne trade. For example, in 2016, tanker and dry bulk shipments accounted for nearly 80% of the world seaborne trade weight, and approximately 28% of the world's seaborne trade by value. This is described more fully in Table 1 where world seaborne trade classified by cargo type is described through weight and value. Unlike the liner services that are modeled in these two works, bulk services are not fixed, and thus replicating the global bulk shipping network and related flows requires additional consideration.

This paper aims to develop a unifying global network representation and decision model, the Strategic Cargo Routing Model, incorporating both liner and bulk (or tramp) shipping markets to estimate real-world traffic flows and study traffic patterns at the global scale. The model takes shippers' perspective rather than that of a ship operator or carrier, enabling global and strategic analyses, as well as the study of shipper response to service changes. While other works develop tools for, or investigate, tactical decisions (e.g. vessel scheduling) by carriers and operators whose decisions support their own objectives, few works (specifically, Tavasszy et al., 2011; Lin and Huang, 2017; Shibasaki et al., 2017; Wang et al., 2018) take a system-perspective or provide tools to support analysis of the whole maritime system. A strategic, system-wide model as proposed herein can support, for example, port or canal capacity expansion or other development investment studies.

In the Strategic Cargo Routing Model, containerized and bulk movements are jointly modelled within a mixed-integer linear program (MILP). Port capacity, which depends on not only berth, but navigation channel and outer anchorage capacities, is shared between container and bulk vessels at locations where both operations are supported. Shipper route choice is characterized through a disutility function that is minimized in the objective function of the model to determine global containerized and bulk flows. An iterative approach that combines heuristic Gradient Descent and Relax-and-Fix Decomposition methods is proposed for the calibration and solution of the Strategic Cargo Routing Model over a global liner and bulk cargo network, a representation for which is equally

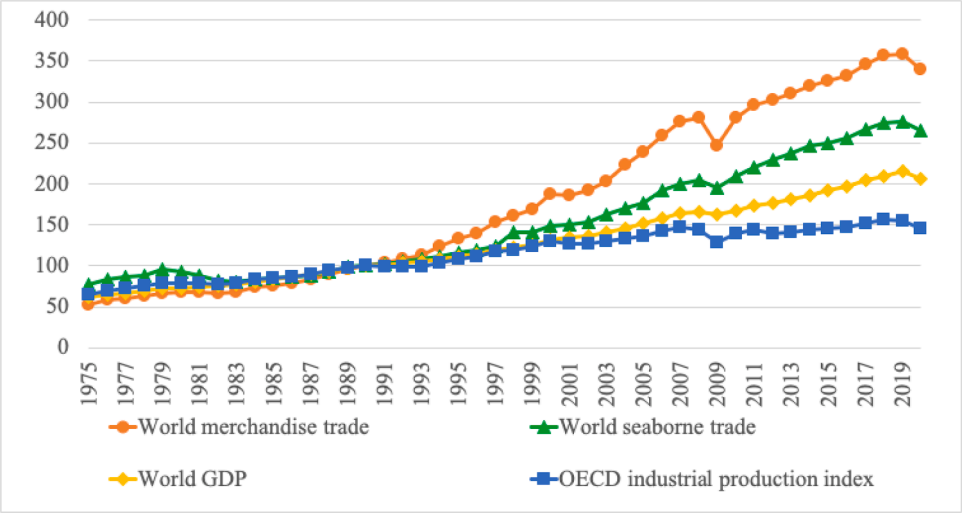


Fig. 1. Economic Cooperation and Development (OECD) industrial production index and indices for world GDP, seaborne trade, and merchandise trade, data obtained from UNCTAD through 2014 (UNCTAD, 2015) and UNCTADSTAT from 2015 through 2020.

Table 1
World seaborne trade by cargo types in 2016.

Cargo type	Trade weight (billion tons)	Percentage share by weight (%)	Percentage share by $ value (%)
Tanker shipments	3.1	29.7	22
Dry bulk shipments	4.9	47.8	6
Containerization	1.7 (139 million TEUs)	16.9	52
General cargo	0.6	5.7	20
Total	10.3		

* Percentage share by weight calculated from 2016 trade volumes reported in (UNCTAD, 2019); percentage share by value obtained from Lloyd's Maritime Intelligence Unit.

important and is developed herein.

The proposed Global Cargo Shipping Network representation contains 161 seaports covering 52 countries created from updatable, publicly available data sources. Methodologies used in the identification of key ports, shipping routes, vessel classes and numbers, bulk and container origin-destination (O-D) demand, and other needed parameters in constructing the network are presented. Sufficient network details, as well as data sources and methods for extracting needed inputs, are given to allow others to use and update the network. Using the developed maritime network, mathematical model and calibration-solution methodology, global maritime traffic flow patterns were estimated and the estimates were evaluated.

A review of the literature related to global maritime network representation and seaborne route-flow estimation is provided in the next section. In addition to justifying choices made in the calibration and solution processes presented and employed herein for route-flow estimation, the review revealed that no prior global maritime network representation with sufficient breadth and fidelity to support a worldwide analysis with realism exists in the literature. Neither does a prior model exist that combines liner and bulk cargo shipping operations, despite that they share capacity at key network elements. This paper seeks to overcome these gaps. Specifically, the methodological contributions of this work are: (1) a unifying strategic cargo routing network model that estimates worldwide containerized and bulk cargo flows within a single framework; (2) a calibration-solution methodology with applicability to detailed, world-wide, large-scale network; and (3) the creation of a high-fidelity, global shipping network with all details needed for model reconstruction and ready updates as data evolves.

The literature review is followed by presentation of the Strategic Cargo Routing Model with details of required inputs and assumptions in Section 3. Section 4 presents techniques for developing the topology of the proposed Global Cargo Shipping Network, along with its attributes, all of which is provided for wider use through Github. Sufficient details of the processes behind the network's development are provided to enable changes to the network with supply-side enhancements, e.g. introduction of a new vessel or route, and demand-side changes, e.g. annually updated seaborn, country-to-country trade volumes. Section 5 describes the optimization-calibration solution algorithm for solving and calibrating the Strategic Cargo Routing Model. In Section 6, the calibrated model is applied to the Global Maritime Network to estimate 2018 global seaborne trade flows. An analysis of the results is also given. Finally, future considerations are described in Section 7.

2. Literature review

Maritime systems have been studied and modelled from a wide variety of perspectives, including, for example, liner-shipping network design (Agarwal and Ergun, 2008, 2010; Brouer et al., 2014), vessel routing and scheduling (Bilgen and Ozkarahan, 2007; Álvarez, 2009; Lin and Liu, 2011; Meng et al., 2014; Christiansen et al., 2017; Yang and Wang, 2017), and cargo routing (Fan et al., 2010; Tavasszy et al., 2011; Bell et al., 2013; Lin and Huang, 2017; Achurra-Gonzalez et al., 2019), among other areas. Only a few of these works applied their ideas at a global scale. In line with the focus herein, works in the literature are reviewed that develop global cargo shipping network representations or support seaborne cargo transport analyses (SubSection 2.1) and seaborne trade flows estimation (SubSection 2.2) at a global scale.

2.1. Representing the global maritime network in the literature

Models of the world's maritime shipping systems have been presented in relation to specific applications. These models are: (1) detailed and global, but with data that is proprietary and/or for a subset of flows, e.g. pertaining to a single industry or ocean carrier or carrier collaboration; (2) highly aggregated, even with a single node representing an entire continent; or (3) detailed but focused on a specific trade route or small region.

Li et al. (2015) and Xu et al. (2015) divided the world into regions and studied changing global container shipping through concepts of traffic development, centrality, dominance and vulnerability. These works relied on data from Containerization International (CI-Online). CI-Online provides datasets associated with liner shipping. Their data is not freely accessible since the company's acquisition by Lloyd's List in 2012. Tavasszy et al. (2011) created a global transport network using country nodes (origin/destination centroids), as well as port and maritime nodes, for a total of over 400 nodes. Transit between nodes, though, is presumed to be direct (as the crow flies), ignoring obstructions from land masses, and thus, the network is overly simplistic. Lin and Huang (2017) included 31 ports from the former G6 strategic alliance in their global network representation and observed sea route information from the International Liner Shipping Database from the Institute of Transportation of Taiwan. This database is not open-source and specific

details are not published. Generally, most maritime data is commercially owned and is either protected or for sale.

IHS Markit provides country-to-country seaborne trade data by cargo type, i.e., containerized, dry, liquid, and general cargo. IHS Markit sells this data at a significant rate, making it inaccessible to the majority of the shipping research community. For this reason, Brouer et al. (2014) pointed out that limited access to data is a key barrier in liner shipping network design studies. To overcome this data access difficulty, Brouer et al. (2014) created benchmark data instances with various scales from a variety of sources, including historical data from Maersk, the top liner ocean carrier by capacity. Their benchmark problems are specific to liner shipping and are limited to one ocean carrier's routes. Similarly, Fremont (2007) used a global maritime network from Maersk to test a hub-and-spoke alternative to port-to-port maritime services. While Maersk is one of the largest international shipping companies, its operations contribute only a fraction of world flows.

With a focus on cargo-routing under disruption, Bell et al. (2013) and Achurra-Gonzalez et al. (2019) used networks with six nodes and direct links to represent trade lanes between the Far East and Europe. Their networks are highly aggregated and greatly abstracted from reality, with a single centroid node or a specific port representing a point of entry for a whole continent.

There also exist detailed maritime networks for specific regions. To test their liner shipping network design strategy, Agarwal and Ergun (2008, 2010) created maritime networks with ports distributed in Asian and North American regions. They considered service networks of two shipping companies, OOCL (Orient Overseas Container Line) and APL (American President Lines). Fan et al. (2010) focused on container flows coming into the U.S. from Asia and Europe, both of which are represented as centroids. In a Northern Sea Route planning problem for the Arctic, Lin and Chang (2018) collected data from an existing, primary international, Asia-Europe shipping service. Their network includes 20 ports and 53 arcs specific to the study area.

More than half of the world's maritime cargo by weight is in the form of dry and liquid bulk. Prior works mentioned thus far have focused entirely on liner (container-based) shipping. Different from liner services, bulk services are not fixed, but oftentimes operate along specific passages to fulfill short or long-term contracts (Global Security, 2021). Alizadeh and Talley (2011) discussed global shipping route characteristics for categorized bulk cargo vessels based on shipping routes as summarized by The Baltic Exchange. Since Capesize and large-sized Panamax ships mostly transport dry bulk commodities between export and import nations, their shipping routes are distinct. Small bulk ships serve a more fragmented market and have scattered trade routes that may use any combination of ports. Liquid bulk cargo network configurations were designed in the literature based on pre-defined routes of bulk ships (Lin and Liu, 2011; Christiansen et al., 2017). The routes are specific to vessel type (Bilgen and Ozkarahan, 2007; Yang and Wang, 2017) used in specific regions or for specific commodities. These works suggest the use of historical routes for bulk carriage as a reasonable approach to forming the bulk shipping network. This approach is employed herein in forming the bulk shipping subnetwork of the combined global liner and bulk maritime network.

Thus, there is no comprehensive, global maritime network representation in the literature that can facilitate analysis for real-world understanding and no simple means to obtain such a network even with sufficient funds. Only highly aggregated networks or disaggregate representations with proprietary data or data representing a specific industry, region or ocean carrier were found. Further, there is no clear and systematic method for representing the cargo shipping network or its attributes. Nor are the attributes based on open-source data. Moreover, bulk and liner shipping markets are studied entirely separately. This paper seeks to overcome these omissions by creating a global cargo shipping network of both liner and bulk services, along with methodologies for generating the needed data objects from publicly available, and updatable data sources.

2.2. *Global route modeling in the literature*

A few works have developed container assignment models that aim to replicate real-world, annual global container traffic flows (Tavasszy et al., 2011; Lin and Huang, 2017; Shibasaki et al., 2017; Sun et al., 2020). To this end, Tavasszy et al. (2011) formulated a generalized linear function of costs and time spent at ports and over links. Routes are selected through a simple logit route choice model on the cost function. Lin and Huang (2017) proposed a generalized nonlinear cost function of port and transport link costs and times. Routes are chosen to achieve a user equilibrium, thus accounting for traffic congestion along routes or at the ports. Cost parameters in both models are calibrated to minimize differences in calculated versus real-world annual port throughput data; although, only the latter provides details of the methodology and outcomes. With container, port and route data from 2010 from CI-Online, which prior to 2012 was available to the public, Shibasaki et al. (2017) developed a global shipping network and calibrate a user equilibrium model for estimating global container cargo flows taking a shipper perspective. These works rely on data from 2010 that is no longer available and cannot be readily retrieved or updated. Also in an effort to model global container cargo flows, a cargo flow assignment model with a concave cost function is used by Sun et al. (2020) that determines flows along routes that satisfy multiregional seaborne trade demand. These works are simplistic in their modeling of vessels and ports as they presume a single, uniform vessel type that serves only containerized cargo and, further, neglect the role of ports and their limited capacities.

The most relevant works consider various aspects in estimating global seaborne cargo flows. Fan et al. (2010) modelled and optimized container flows imported to the United States from China and Europe only. They noted that the ship size, route, port and interior shipping corridor play a key role in maritime systems, influencing optimal shipping strategies. Bell et al. (2013) modelled liner shipping markets using a cost-based container assignment model characterized through a linear program. Their model includes port and route capacity limitations, but no vessel type, loading capacity, vessel draft or cargo type.

In the application of dry bulk shipping, Bilgen and Ozkarahan (2007) formulated the grain routing and inventory management problem as a MILP in which total vessel capacity of each vessel type on a given route is used to constrain transport flows. They test their model on a small, hypothetical network specific to the grain market. Lin and Liu (2011) recognized the importance of vessel characteristics in bulk shipping. They study a combined ship allocation, freight assignment, and ship routing problem that accounts for

vessel type with loading capacity. These prior works on bulk cargo routing have, thus, only considered a single-product market or specific regions. There appears to be no prior global bulk routing models in the literature. Moreover, no prior works develop routing models for containerized and bulk cargo jointly. This paper aims to fill these gaps through the construction of a comprehensive model constructed on concepts from classical containerized and bulk cargo route modeling.

Building on the advances made in these prior works, the Strategic Cargo Routing Model proposed in this paper seeks optimal global containerized and bulk flows within a single model to minimize total transportation costs, including transit costs over sea and handling costs at ports. It takes concepts from both markets related to service capacities and accounts for differing vessel types. To understand and study global maritime traffic patterns, the proposed model is calibrated and representative global traffic patterns are generated and studied.

3. Problem conceptualization and formulation

3.1. Model input overview

Global containerized and bulk maritime flows are affected by transport capacities and efficiencies of shippers, ocean carriers and ports. Operations are specific to liner and bulk shipping markets, and containerized and bulk cargo are carried by different types of vessels and handled by specialized equipment at different terminals for ports with both container and bulk operations. Liner shipping companies provide scheduled shipping loops, while bulk shipping companies serve their markets fluidly. For example, a liquid cargo vessel will make stops according to customer requests in response to changing inventories. Their differences create a challenge for their simultaneous modeling; thus, they have been treated separately in the literature. Yet, they share system capacities and need to be considered simultaneously. The problem of estimating global maritime annual traffic flows is cast herein as a MILP i.e., the Strategic Cargo Routing Model. The model assigns and routes both types of cargo over a shared container and bulk cargo shipping network.

The Global Cargo Shipping Network is composed of four subnetworks, three for bulk (by vessel draft - below 12 m, 12 to 18 m and 18 to 28 m drafts) and one for containerized passages, as displayed on world maps in the Appendix A. In bulk shipping, vessel draft and route and port bathymetry affect a vessel's route options. For example, to complete the carriage of coal from Port Dampier in Australia to Port Hamburg in Germany, a Suezmax Bulker would go through the Suez Canal to Europe while a Very Large Bulk Carrier limited by depth at the Suez Canal would need to detour, perhaps taking a long distance to navigate through the Cape of Good Hope. In reality, as ships operate over a plane, there are infinite possible subnetworks. Here, the possibilities are reduced to only three for bulk cargo ships. Shallower routes can be chosen by vessels with smaller drafts. While the smaller vessels can use any of the routes in the three subnetworks, it is presumed that they will choose routes within the category that just meets their requirements as these paths tend to be shorter to navigate.

The subnetworks are interconnected at ports, their outer anchorage, and other geographic chokepoints (e.g. the Panama and Suez Canals). Vessels flow through each subnetwork and each unit of flow, a vessel belonging to an ocean carrier, is assigned to a route between its O-D ports. Ports are classified as container ports, bulk ports, and those that handle both. They do not all serve the same businesses. For example, some ports may serve vessels carrying liquid natural gas (LNG), where others will not serve this market at all.

Route options depend on the type of cargo to be shipped. For containerized shipments, ocean carriers ship according to scheduled routes formed through a sequence of port calls with a possible change in vessels. These routes generally form a loop. Transshipment operations are typically handled at large hub ports; thus, it is common for a container to change from one vessel to another and for its route to include a call to one or more transshipment ports. Equipment at the ports may only be able to handle smaller vessel sizes, thus limiting options for larger vessels. Only a restricted number of major ports around the world, such as in Shanghai, Singapore, Rotterdam, Antwerp, Long Beach, for example, can support the newest, largest mega-ships with capacities larger than 18,000 TEUs. To characterize container traffic flows over routes, Bell et al. (2013) introduced a concept of task-based legs and links. In their cost-based container assignment model, inbound and outbound leg flows are restricted by route carrying capacities and port handling capacities. Their leg-based modeling technique is adopted herein.

Bulk shipping is demand-driven, and routes flex with this demand. Route carrying capacities are a function of the capacities of the vessels that operate along them, and the ability of a route to support vessel passage depends on the bathymetry of the route and vessel draft, as well as the channel depth at the ports.

3.2. Input details and assumptions

Restricted by port handling and route transport capacities and bathymetry, containerized and bulk cargo O-D flows are generated over the set of interconnected subnetworks with the aim of minimizing a generalized cost function. Resulting flows provide an estimate of global maritime traffic patterns. The Strategic Cargo Routing Model that seeks these patterns is described in terms of vessels,ports, O-D pairs, routes, and costs. Before proceeding to give the formulation, these details, along with assumptions, are first presented. How these elements serve as data objects in creating an updatable global cargo shipping network representation using publicly available data is described in detail in Section 4.

3.2.1. Vessels

Vessels considered include: containerships, tankers, and bulkers that serve either containerized, liquid- or dry-bulk cargo, respectively. Vessels are grouped by the cargo type they can hold and carrying capacity. A set of 10 such vessel types were constructed, where each set is delineated by its range on its carrying capacity (Table 5). The required inputs, capacity and draft for each vessel type

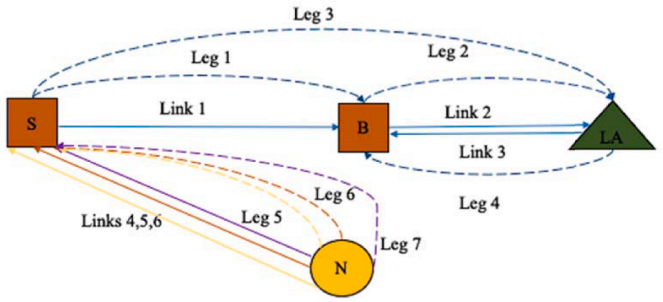
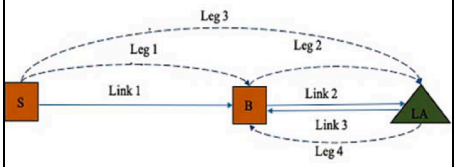
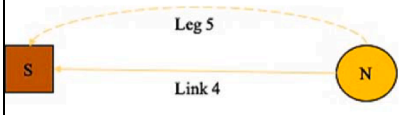
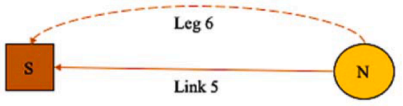
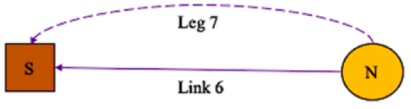
are set as the maximum capacity and draft for each vessel type.

3.2.2. Ports

Each seaport is categorized as a container, bulk (serving both liquid and dry bulk), or both container and bulk port based on its business focus. Within the model, port depths were employed in determining the largest vessel size that a port can serve based on the drafts of vessels when filled to capacity. The size of the largest vessel that can call at a terminal is restricted by infrastructure limitations (both air draft and channel depth). As not all ports have or provide air draft restrictions, only the port's channel depth is considered in determining the largest vessel size a port can serve.

Port capacities were calculated based on historical port throughput data. According to the U.S. Department of Transportation (USDOT 2017), there are many factors that influence annual throughput capacities of ports, including channel depth, berths and berth lengths, loading and unloading equipment, storage space for cargo, containers, and chassis, modal connections, and port operating factors. They note that capacities are also affected by events, such as weather, ice, and other disruption events. Based on interviews with port authorities, Lin and Chang (2018) found that a port's capacity is generally maintained at 20% above demand. Thus, herein 120% of the port's average throughput over the past five years is utilized as its annual throughput capacity. Note, too, that most large ports capable of handling containerized and bulk cargo will do so separately at specific terminals. However, the terminals share human resource and power capacities. They are also limited by shared outer anchorage capacity. Thus, containerized and bulk flows are bundled in considering adherence to port capacity constraints. Likewise, inbound and outbound traffic are combined in requiring adherence to total port capacity limits.

Table 2
Network structure with subnetworks, ports, routes, links and legs of the model.

<table>
<tr><td>Network</td><td colspan="5"> 
</td></tr>
<tr><td>Subnetwork</td><td>Subnetwork structure</td><td>Port</td><td>Route</td><td>Link</td><td>Leg</td></tr>
<tr><td>Liner-shipping subnetwork</td><td> 
</td><td rowspan="4">Container port: Los Angeles (LA) Bulk port: Newcastle (N) Both container and bulk ports: Shanghai (S) and Busan (B)</td><td>Shanghai →Busan→Los Angeles→Busan</td><td>1, 2, 3</td><td>1, 2, 3, 4</td></tr>
<tr><td>Bulk-shipping subnetwork (vessel draft ≤ 12 meters)</td><td> 
</td><td>Newcastle → Shanghai</td><td>4</td><td>5</td></tr>
<tr><td>Bulk-shipping subnetwork (12 <vessel draft ≤18 meters)</td><td> 
</td><td></td><td>5</td><td>6</td></tr>
<tr><td>Bulk-shipping subnetwork (18 < vessel draft ≤ 28 meters)</td><td> 
</td><td></td><td>6</td><td>7</td></tr>
</table>

3.2.3. O-D pairs

The model builds on annualized cargo-based, port-to-port shipping demand. This demand data includes demand by containerized and both dry and liquid bulk cargo. O-D demand is assumed to be static over the study period. Domestic trade is not considered and, thus, traffic flows between ports of the same country are excluded.

The cargo-based port-to-port O-D pairs data include not only the origin and destination ports, but also associated cargo types and weights (tons). For unifying the data, containers were converted to weight at a rate of 12 tons per TEU assuming containers are filled to capacity. This weight conversion standard is adopted from the European Sustainable Shipping Forum Sub-group (European Sustainable Shipping Forum, 2021).

3.2.4. Routes

To model traffic flows over the developed cargo shipping network, routes along with constituent links and legs are introduced to generate the task-based shipping network representation as first introduced by Bell et al. (2013) for liner shipping. By definition, in liner shipping, routes are sequences of port calls operated by ships. Adjacent pairs of port calls served in the route are defined as links, while legs are based on tasks (the loading and unloading of cargo at two ports, for example) executed over the route. Such tasks may require transshipments through intermediate ports. Bulk shipping vessels are assigned between ports, and while a bulk vessel may serve multiple ports one after another, each route consists of a single O-D pair, link and leg (containing only one link). There are four types of legs in the liner-shipping subnetwork, i.e., O-D, origin-transshipment, transshipment-destination, and transshipment-transshipment, while there are only origin-destination legs in the bulk shipping subnetworks. A bulk route between an O-D pair will consist of one link; however, the specific voyage undertaken by a vessel depends on several factors, including its draft.

The three bulk shipping subnetworks contain routes over port pairs based on the three vessel draft ranges: (1) under 12 m, (2) between 12 and 18 m, and (3) between 18 and 28 m. Boundaries of these three ranges were chosen to be consistent with vessel draft limits at globally geographical chokepoints (Panama and Suez canals). For example, 12 and 18 m are vessel drafts for Panamax and Suezmax, respectively. The vessel draft of 28 m is set for very large crude tankers or bulkers. The three bulk vessel classes by draft creates the need for three navigable routes for each O-D pair. A vessel is presumed to use the link with the shallowest depth that permits its passage.

Table 2 illustrates the shipping network with its four subnetworks, ports, routes, links and legs. The network contains four ports, including a container port (Los Angeles, LA), a bulk port (Newcastle, N), and two ports that handle both container and bulk cargo (Shanghai, S, and Busan, B). For example, the liner route loops through ports in Shanghai, Busan and Los Angeles and then back to Busan. There are three links for adjacent port calls and four legs based on loading-and-unloading tasks. Leg 1 represents the task of loading containers at the port in Shanghai and unloading the containers at the port in Busan. Containers are transported through leg 1 to leg 2 to their final destination port in Los Angeles. Thus, a transshipment task is executed at the port in Busan. Alternatively, containers can be transported along leg 3 without any loading/unloading actions in Busan.

Also illustrated in Table 2 is a bulk route involving port calls between Newcastle and Shanghai. Since bulk routes are created with consideration of vessel draft, there are three potential links and, thus, three legs for this route, each associated with one of the three bulk-shipping subnetworks. Multiple legs, typically with differing links, can exist for a given pair of ports, whether in liner or bulk applications.

3.2.5. Costs

Shipping costs account for the costs associated with the voyage over sea, including depreciation, rental, insurance and maintenance costs associated with the vessel, as well as fuel consumption, tariffs at canals, and crew costs. Shipping costs depend heavily on the vessel type and size. They also account for handling costs at ports, including fees and time charged for completing loading, unloading and transshipment tasks.

This paper adopts a vessel-based generalized shipping cost approach, including transit costs over sea and handling costs at ports. A linear relationship between weight and transit costs over the routes and handling costs at ports is assumed. Generalized transit costs c_{ϕ}^{v} (\$) per transit by vessel are defined with respect to route ϕ and vessel type v. A generalized port handling cost (h_p, \$/ton) was obtained from De et al. (2019). Seedah et al. (2013) proposed a port operating cost model, including navigational service costs, berth service costs and cargo operation costs. De et al. (2019) used this cost model to calculate port costs in the U.S. Gulf Coast, providing estimates of cargo handling costs at ports. For lack of better cost estimates, based on their calculation, all ports considered herein are assumed to initially use the same generalized port handling cost, h_p, of \$20 /ton, regardless of cargo type or the loading, unloading or transshipment services. Generalized port handling costs, h_p, vary across ports and impact model solutions. These costs are not public, and thus, are obtained through calibration on published annual port throughputs as is described in Section 5. They are presumed to represent not only handling charges, but also handling times. Such handling times depend on infrastructure and space availability (e.g. yard and terminal spaces) as well as equipment and labor efficiencies. Thus, a smaller, less efficient port with relatively low handling charges may have higher generalized port handling costs due to the longer handling times.

3.3. Mathematical formulation

The Strategic Cargo Routing Model formulation is given in Table 3. Containerized and bulk cargo are routed and assigned over the Global Cargo Shipping Network. The assignment of cargo to routes is strategic and aims to minimize total shipping cost through objective function (1), where route options and service capacities are flow-independent, and therefore fixed. Thus, shipping costs do

not change as flows increase. Shipping costs include transit costs by sea and handling costs at ports. Key decision variables determine the network flows, specifically the leg flows and route transits, and are restricted by network capacities. Constraints (2) define the largest cargo ship size that can be handled at each port. Constraints (3) - (5) define route capacity limits as a function of the vessel types that serve the routes within each subnetwork, as well as number of vessel transits. Constraints (6) are cargo flow balance constraints at each port given by cargo type (containerized, liquid and dry bulk). Total inbound and outbound leg flows at ports are restricted by port throughput capacity limitations in constraints (7). Non-negativity of decision variables is enforced through constraints (8).

Objective function Minimize

$$\sum_{v\in V}\sum_{\phi\in\Phi} c_{\phi}^{v} * f_{\phi}^{v} + \sum_{\mathrm{p}\in\mathrm{P}} h_{p} * \sum_{v\in V}\sum_{\phi\in\Phi}\sum_{(\mathrm{i,p}),(\mathrm{p,j})\in\mathrm{L}(\Phi)}\sum_{d\in D}\left(y_{\phi}^{v}(\mathrm{i,p,d}) + y_{\phi}^{v}(\mathrm{p,j,d})\right) \tag{1}$$

Subject to

$$\sum_{d\in D} y_{\phi}^{v}(\mathrm{i,p,d}) \ \leq \mathrm{capv_p} * f_{\phi}^{v}, \forall\ v\in V,\ \phi\in\Phi,\ (\mathrm{i,p})\in\{\mathrm{L}(\phi)\} \tag{2}$$

$$\sum_{l\in\mathrm{L}(\phi)}\zeta_{kl\phi}\sum_{d\in D} y_{\phi}^{v}(\mathrm{i,p,d}) \leq \mathrm{cap}_{v} * f_{\phi}^{v},\ \forall\ v\in V,\ \phi\in\Phi,\ k\in K(\phi) \tag{3}$$

$$f_{\phi}^{v} \leq\ m_{\phi}^{v},\ \forall\ v\in V,\ \phi\in\Phi \tag{4}$$

$$\sum_{\phi\in\Phi_a} f_{\phi}^{v} \leq\ n_{a}^{v},\ \forall\ v\in V,\ a\in A \tag{5}$$

$$\sum_{v\in\mathrm{V}_g}\sum_{\phi\in\Phi}\sum_{(\mathrm{i,p})\in\ \mathrm{L}(\Phi)} y_{\phi}^{v}(\mathrm{i,p,d}) - \sum_{v\in\mathrm{V}_g}\sum_{\phi\in\Phi}\sum_{(\mathrm{p,j})\in\ \mathrm{L}(\Phi)} y_{\phi}^{v}(\mathrm{p,j,d}) = \begin{cases} \sum_{\mathrm{o}\in\mathrm{O}} TD_{odg}, & \text{for p} = \mathrm{d}\in\mathrm{D} \\ -TD_{odg}, & \text{for p} = \mathrm{o}\in\mathrm{O} \\ 0, & \text{otherwise} \end{cases} \forall d\in D, g\in G \tag{6}$$

$$\sum_{v\in V}\sum_{\phi\in\Phi}\sum_{(\mathrm{i,p}),\ (\mathrm{p,j})\in\ \mathrm{L}(\Phi)}\sum_{d\in D}\left(y_{\phi}^{v}(\mathrm{i,p,d}) + y_{\phi}^{v}(\mathrm{p,j,d})\right) \leq \mathrm{cap_p}, \forall \mathrm{p}\in\mathrm{P} \tag{7}$$

Table 3
Notation and mathematical formulation.

Sets	
P	Set of ports
O	Set of origin ports
D	Set of destination ports
G	Set of cargo types, containerized, dry and liquid bulk cargo
V	Set of vessel types
V_g	Set of vessel types carrying cargo type g, $g\in G$
A	Set of subnetworks
Φ_a	Set of routes operated by subnetwork a, $a\in A$
Φ	Set of routes in the global shipping network
$K(\phi)$	Set of links $k(i,j)$ within route $\phi\in\Phi$, $i,j\in P$
$L(\phi)$	Set of legs $l(i,j)$ within route $\phi\in\Phi$, $i,j\in P$
Parameters	
c_{ϕ}^{v}	Generalized transit cost per ton of a vessel of type v deployed by route ϕ, $\phi\in\Phi$, $v\in V$
D_{ϕ}	Distance (nautical miles) over route ϕ, $\phi\in\Phi$
ω_v	Transit cost (US $) per nautical mile per ton for a vessel of type v, $v\in V$
h_p	Handling cost per ton at port p, $p\in P$
m_{ϕ}^{v}	Available number of annual transits for vessel type v by route ϕ, $\phi\in\Phi$, $v\in V$
n_{a}^{v}	Available number of annual transits for vessel type v by subnetwork a, $v\in V$, $a\in A$
cap_v	Capacity of vessel type v, $v\in V$
cap_p	Annual throughput capacity of port p, $p\in P$
$capv_p$	Largest vessel size that can be handled at a port p, $p\in P$
TD_{odg}	Cargo type g to be transported from origin port o to destination port d, $g\in G$, $o\in D$, $d\in D$
$\zeta_{kl\phi}$	1 if leg l uses link k on route ϕ, 0 otherwise
Decision variables	
$y_{\phi}^{v}(i,\ j,\ d)$	Units of leg flows $(i,j)\in\ L(\Phi)$ using vessel type v en route to destination d, $v\in V$, $d\in D$
f_{ϕ}^{v}	Number of transits on route $\phi\in\ \Phi$ using vessel type v, $v\in V$

$$y_{\phi}^{v}(i,\ j,d) \in \mathbb{R}_{+},\ f_{\phi}^{v} \in \mathbb{Z}_{+} \tag{8}$$

4. Building the global cargo shipping network topology and its node and link attributes

The Global Cargo Shipping Network is constructed based on concepts of four data objects: O-D pairs, ports, routes and vessels (Fig. 2). Methods for generating the contents of the data objects and potential, publicly available, and updatable data sources, along with data samples specific to the year 2018, are described for each data object in the following subsections. All data are shared on GitHub.

4.1. O-D pair demand data construction and sources

4.1.1. Country- or regional-based demand data preparation

Global shipping demand data are unified under a cargo-based, port-to-port demand data object. More refined values at a port-to-port level can be employed, but publicly available data at a global level of this nature is nonexistent. Even country-to-country seaborne trade data are not publicly accessible. Such a dataset can be obtained from IHS Markit, but at a substantial cost. Country-to-country trade data do exist and is readily available (e.g. UNCTADstat and United Nations (UN) Comtrade Database). This data is given in monetary units rather than in units of weight or volume. Thus, conversion to vessel flows or TEUs can be difficult.

To build this data object, country- or regional-based demand data and historical port throughput data are utilized. There are several sources for these inputs. To obtain country- or regional-based demand data, two key sources of data were employed: resourcetrade.earth (https://resourcetrade.earth/about#top) and UNCTAD (2019).

Country-based demand data for bulk cargo were collected from resourcetrade.earth, which provides a comprehensive list of bulk commodities associated with natural resources. Bulk cargo is divided into dry and liquid bulk. Dry bulk commodities include major (e.g., iron ore, coal and grain) and minor bulk commodities (e.g., agricultural products, fertilizers, forestry products, coal, metals and minerals, cement etc.), while tankers carry liquid fuels, such as oil, oil products, natural gas, LNG, liquefied petroleum gas (LPG), and chemicals.

Due to the availability of alternative modes of transport between adjacent countries, bulk trade estimates obtained from resourcetrade.earth cannot be used directly when identifying country-based seaborne-bulk trade pairs. Based on the TransBorder Freight Dataset from the USDOT (USDOT, 2021), approximately 60% of trade between the United States and Canada or Mexico is transported by seaborne vessel. For lack of better information, it was assumed, therefore, that 60% of bulk trade by weight is seaborne for adjacent countries globally and 100% otherwise.

As there is no open-source, country-based demand data for containerized cargo, aggregated trade flows are used. UNCTAD publishes annual container trade volumes over highly aggregated trade routes. Table 4 lists traffic flows over these routes for 2018 obtained from UNCTAD (2019). Containerized trade on routes within three regions (Asia, Europe, and North America) are included in constructing the global cargo shipping network. UNCTAD (2019)) notes that total container port throughput from these three regions accounted for 88% of global maritime trade in 2018. Also, from Table 4, based on the container volumes shipped over these routes, the total market share for the routes within these three regions is approximately 80% of the world's seaborne shipments. Consequently, smaller container ports in the areas of Africa, Latin America, and Oceania are excluded in constructing the global maritime network. They may be added in future network extensions if desired.

To begin the process of generating regional-based demand data, a top-down approach is applied. The first step is to ascertain trade flows between the three regions (Asia, Europe and North America). These country-based bulk and regional-based container demands serve as inputs to estimate port-to-port shipping demand as discussed in the next subsection.

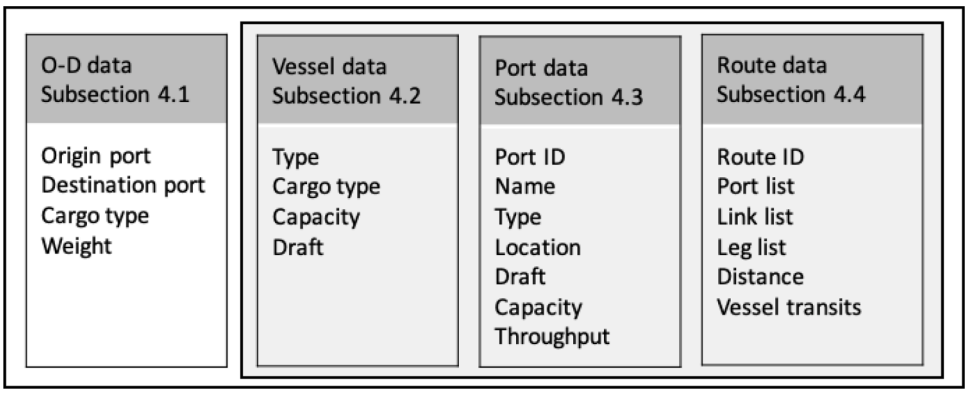


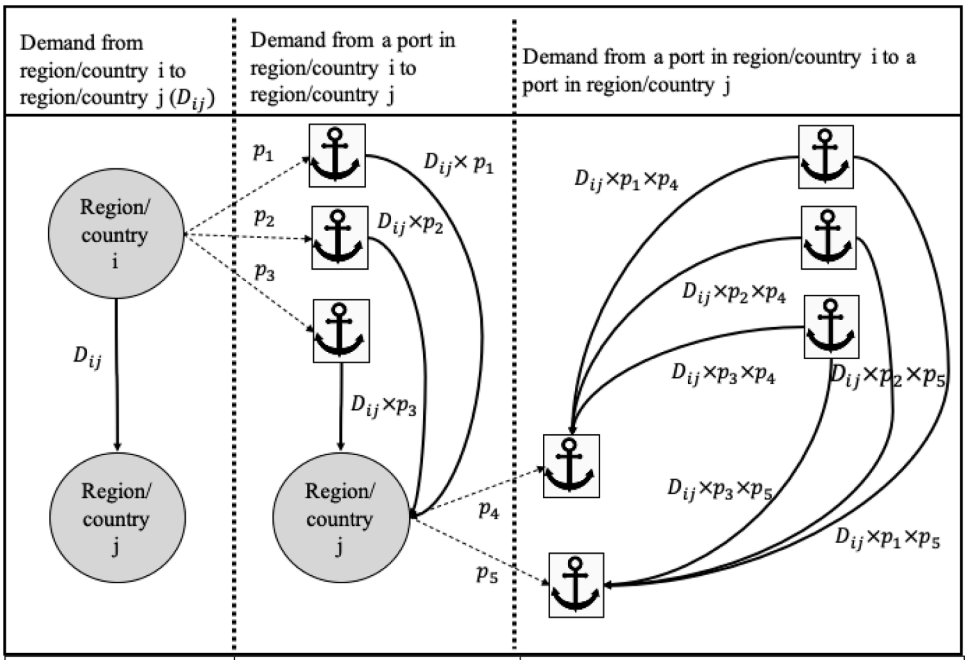
Fig. 2. Data objects underlying the Global Cargo Shipping Network.

Table 4
Containerized cargo flows along major trade routes in 2018, raw data taken from Table 1.9 in (UNCTAD, 2019).

Trade route	Shipped volumes (million TEUs)
East Asia – North America	20.9
North America – East Asia	7.4
Northern Europe and Mediterranean to East Asia	7.0
East Asia to Northern Europe and Mediterranean	17.4
North America to Northern Europe and Mediterranean	3.1
Northern Europe and Mediterranean to North America	4.9
Total	**60.7**

4.1.2. Port-to-Port shipping demand estimates through portion-based allocation

To estimate port-to-port shipping demand, country-based bulk demand estimates and region-based container cargo demand estimates are allocated proportionally across ports based on each port's market share in terms of total inbound and outbound throughput following a similar portion-based allocation method in (Lin and Huang, 2017). Region-based data is used herein, as the country-to-country data they used in their work is not publicly available. This approach ensures supply and demand balance is maintained at the ports. The portion-based allocation process is depicted in Fig. 3. The figure illustrates the three key steps as applied to



Obtain demand from region/country i to region/country j.	Estimate demand from a port in region/country i to region/country j.	Compute demand from a port in region/country i to a port in region/country j.
Region-to-region containerized demand; country-to-country dry/liquid bulk demand.	A port's market share is calculated by the percentage of the port's throughput of total throughput of all included ports within a region/country. Outbound flow volumes from region/country i are distributed among ports by multiplying the total outbound volume by the port's percentage of market share, denoted as p_k (k is the port ID number).	Port-to-port shipping demands are computed from the product of demand in flow volumes from a port in region/country i to region/country j and the percentage market share of each port in region/country j.

Fig. 3. Port-to-port shipping demands distributed from regional/country-based cargo demand.

generate port-to-port O-D pairs from regional/country-based cargo demand pairs.

4.2. Vessel

Given the listing of vessel types provided by UNCTAD (2019), and the focus herein on international cargo shipping, ten vessel types, with two types of containerships, four types of tankers and four types of bulkers as listed in Table 5, are included.

4.3. Ports

The world-wide seaports together with shipping passageways create the backbone of the maritime network, yet there is no consistent, comprehensive list of ports that should be included in a global study. In existing works related to global container shipping networks, Brouer et al. (2014) and Fremont (2007) considered container ports only from the ocean carrier Maersk. Moreover, Álvarez (2009) included the top 120 container ports by throughput in TEUs, but ignored ports that are not contained in this list but are important to key, loop-based routes of liner services. Álvarez assumed containers are shipped directly between ports and ignore loop-based routing strategies that are undertaken by the large ocean carriers. Álvarez further left out many ports that participate in these loops, including, for example of the Ports of Balboa and Colon located in Panama. These ports seldomly appear in a top port list, yet they are key to routes that pass through the Panama Canal.

Ports included in the Global Cargo Shipping Network were identified from two sources. First, ports involved in the container shipping loops provided in Table 6 are included. After adjusting for minor variations in naming across routes, 114 container seaports were identified for inclusion. Over 80 ports in the top 100 container seaports by volume listed by the American Association of Port Authorities (AAPA) are included in this list. The 20 excluded container ports in the top 100 list are located in Oceania and South America, which are not served by the main West - East trade routes.

The network also includes 88 bulk ports, which were selected from the top 100 seaports by weight from the AAPA. 12 of these ports handle mostly containerized cargo and were already included as container ports. Although no container ports are included from Oceania and South America, several large bulk ports from these areas are included in the bulk port list.

The 41 ports arising in both container and bulk port lists were labeled as serving both container and bulk cargo. In total, 161 ports, including 73 container ports, 47 bulk ports and 41 both container and bulk ports, comprise the Global Cargo Shipping Network. These nodes are depicted in Fig. 4. The vast majority of included ports are located in Asia, Europe, and North America with 73, 39 and 33 seaports, respectively. While the ports in Asia and Europe are a mix of container, bulk and both container and bulk seaports, two thirds of the seaports in North America are container ports. Only large, bulk ports located in South Africa, South America, and Australia are included. The attributes associated with these ports include: geographical information, draft (channel depth) and 2014 to 2018 annual port throughput used for port capacity estimation and in the calibration. This data was collected from publicly accessible sources, specifically: individual port websites, World Port Source, ports.com, Lloyd's List Maritime intelligence, and AAPA.

4.4. Routes

4.4.1. Creating the liner shipping routes with legs and links

Earlier works (Li et al., 2015; Xu et al., 2015) relied on CI-Online data for route-related input to their liner shipping models. This data was last available in 2012 and, thus, an alternative approach to obtaining the route data is needed. Here, published service loops were collected for the three largest global ocean carrier alliances (2M, Ocean Alliance and THE Alliance) together hosting 100 routes. Global container shipping market share for these three global ocean carrier alliances is approximately 80% and is nearly 95% along the East-West trade lanes (International Transport Forum, 2018). Published service loops from these ocean carrier alliances are summarized in Table 6 and plotted as a liner-shipping subnetwork in the Appendix A. These routes overlap and may even be identical in terms of ports of call, but are considered as independent routes. Each route has its own available number of vessel transits and distance information (nautical mile). Route distances are assumed to be fixed and were collected from port.com. By using the routes along with their frequencies from actual alliance member operations, carrier route design decisions are implicitly modeled.

Table 5
Vessel types in the shipping market of 2018.

Vessel type	Cargo type	Capacity (1000 tons)	Draft (meter)
Large containership	Container	200	18
Neo-Panamax	Container	120	15
Ultra large tanker	Liquid bulk	400	28
Very large tanker	Liquid bulk	300	22
Suezmax tanker	Liquid bulk	200	18
Neo-Panamax tanker	Liquid bulk	120	15
Very large bulker	Dry bulk	400	28
Capesize bulker	Dry bulk	300	22
Suezmax bulker	Dry bulk	200	18
Neo-Panamax bulker	Dry bulk	120	15

Table 6
. Global ocean carrier alliances and number of service loops over trade routes.

Alliance	Ocean carrier	Total number of loops	Number of loops over trade routes Transpacific	Asia -Europe	Asia -Mediterranean	Transatlantic (North Europe-ECNA)	Asia -Middle East	Asia -Red Sea	ECNA-Mediterranean	
			Asia -WCNA and Asia -ECNA and U.S. Gulf							
2M	Maersk MSC	32	8	6	6	6	4			2
Ocean Alliance	Cosco-OOCL CMA CGM Evergreen	38	12	7	7	5	2	4	2	
THE Alliance	Hapag-Lloyd ONE Yang Ming	30	10	5	5	3	5	2		
Total	100									

Note: West Coast North America (WCNA); East Coast North America (ECNA).
1. https://www.msc.com/nld.
2.https://www.cma-cgm.com/news/2380/ocean-alliance-day-3-product-cma-cgm-ready-for-third-phase-starting-in-april-2019.
3.https://www.one-line.com/sites/g/files/lnzjqr776/files/2018–12/EN_PR_THE_Alliance_Unveils_Enhanced_Service_Network_for_2019.pdf.

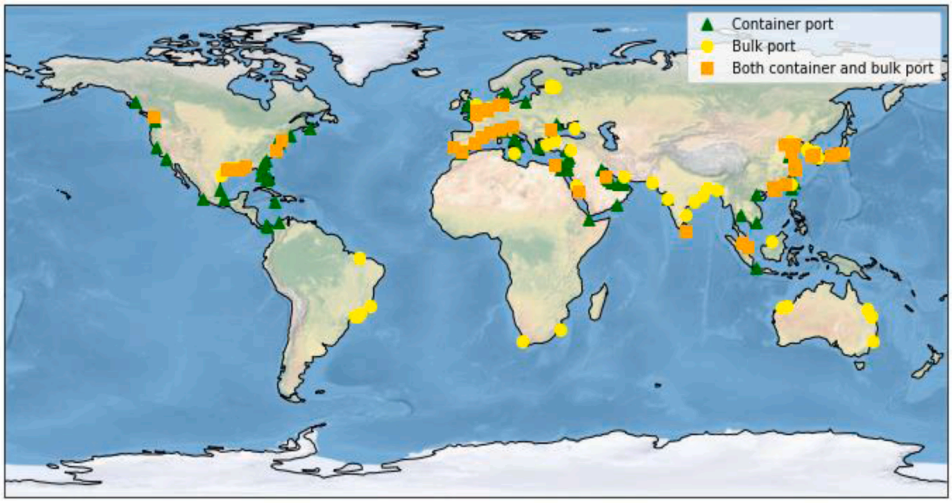


Fig. 4. Ports in the Global Cargo Shipping Network.

4.4.2. Bulk routes

For each bulk-shipping subnetwork defined with vessel draft (i.e., 12, 18, and 28 m), bulk routes were generated from pairings of ports in the bulk port list. With 88 bulk ports included, this results in 7656 (=88 × 87) port pairs per subnetwork. Intra-country bulk routes were presumed to be by land and were, thus, excluded.

Bulk route attributes, such as shipping trajectories and distances, were collected from SeaRoutes. SeaRoutes provides route navigation histories based on five-year historical AIS (Automatic Identification System) data. Given a port pair and vessel draft as inputs, SeaRoutes returns data in the structure of a data dictionary, including route existence status, route trajectories constructed from geocoordinates, and distance. To obtain the needed data for all three subnetworks, 22,968 requests were made to SeaRoutes through its API, and relevant route data were stored within the route data object. The API allows 1000 requests per day on a one-month subscription. The obtained routes associated with the three bulk-shipping subnetworks are summarized in the Appendix A.

4.4.3. Generation of links and legs along routes

The Global Cargo Shipping Network contains a liner-shipping subnetwork with 100 liner loops and three bulk-shipping subnetworks. Adjacent-port-call-based links and task-based legs were generated to serve the included routes following (Achurra-Gonzalez et al., 2019; Bell et al., 2013).

To illustrate the generation method, a liner route sequentially served by ports (Shanghai (S), Kwangyang (K), Busan (B), and Los Angeles (LA)) is taken as an example. In Fig. 5, solid directed lines represent links while dashed lines represent legs. Based on the sequence of port calls, there are 6 links connecting adjacent ports along the route. Thus, the number of links in a route is equal to the number of port calls. Legs from Port S to the other three ports, K, B, and LA, represent the loading of containers at Port S and their unloading at Ports K, B, and LA, respectively. The leg S→*K* passes through link S→*K*, while the leg S→LA traverses links S→*K*, K→*B*, and B→LA. Similarly, there are three legs from Port K to the other three ports. The total number of legs along the route is 12 (=4 × 3). Note that only three legs emanating from Port S are depicted in Fig. 5. In general, there are N ports in a route and, thus, $N \times (N-1)$ legs along the route. This differs from bulk-shipping routes that contain only one port call and two ports represented by a single link and single leg.

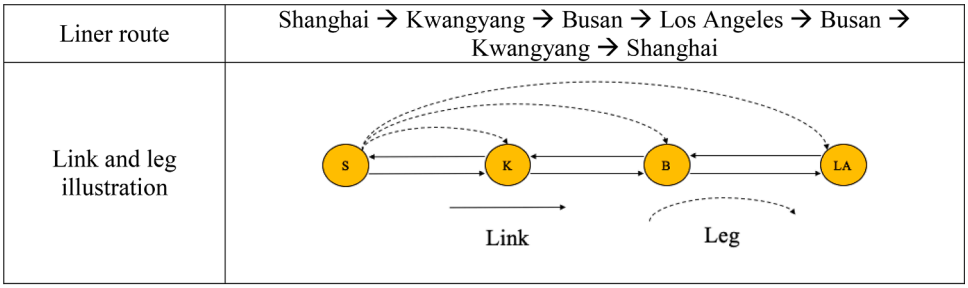


Fig. 5. Link and leg illustration along a liner route.

5. Solution algorithm

The Strategic Cargo Routing Model seeks an assignment of vessel-based flows to routes by sea to obtain the minimum total generalized transit cost. Applying the Strategic Cargo Routing Model on the global maritime network requires solution and calibration. SubSection 5.1 introduces a heuristic Relax-and-Fix Decomposition method, formally described for an unrelated application in (Noor-E-Alam and Doucette, 2015), for efficient solution of this large MILP. SubSection 5.2 presents the heuristic Gradient Descent method used in calibrating the model to obtain consistency in its outputs (measured through port throughput) with real-world observations. Once calibrated, final solution is obtained in terms of global maritime traffic patterns. Details of this optimization-calibration solution algorithm are provided in Subsection 5.3.

5.1. Model solution by Heuristic Relax-and-Fix Decomposition

Calibration requires an iterative process in which parameters are updated and the Strategic Cargo Routing Model is re-solved. As the application is large, obtaining a solution in reasonable time was difficult. Applying Gurobi, a popular off-the-shelf software product, in a high-performance computing environment to solve the MILP in a given iteration required several hours. Thus, an alternative, faster approach was needed. For this purpose, the routing problem is decomposed into three subproblems, one for each cargo type (container, dry and liquid bulk). Exact solution is obtained for each cargo type using Gurobi, and the resulting integer decision variable values are fixed within the original Strategic Cargo Routing Model. This resulting problem is called the core problem. With the integer variables preset, solution of the core problem provides optimal leg flows. This process is known as a Relax-and-Fix Decomposition method.

5.1.1. Subproblems

For each cargo type, the associated subproblem includes a subset of O-D pairs, vessels, and routes. A general version of each subproblem formulation is presented next.

Subproblem (g, $g \in G$):

Objective function

Minimize

$$\sum_{v\in V_g}\sum_{\phi\in\Phi_g} c_\phi^v * f_\phi^v + \sum_{\mathrm{p}\in\mathrm{P}} h_p * \sum_{v\in V_g}\sum_{\phi\in\Phi_g}\sum_{(\mathrm{i,p}),(\mathrm{p,j})\in\ \mathrm{L}(\Phi_g)}\sum_{d\in D_g}\left(y_\phi^v(\mathrm{i,p,d}) + y_\phi^v(\mathrm{p,j,d})\right) \tag{9}$$

Subject to

$$\sum_{d\in D_g} y_\phi^v(\mathrm{i,p,d}) \ \le \mathrm{capv_p} * f_\phi^v, \forall\ v\in V_g,\ \phi\in\Phi_g,\ (\mathrm{i,p})\in\ \mathrm{L}(\Phi_g) \tag{10}$$

$$\sum_{l\in\mathrm{L}(\phi)} \zeta_{kl\phi}\sum_{d\in D_g} y_\phi^v(\mathrm{i,p,d}) \le \mathrm{cap}_v * f_\phi^v,\ \forall\ v\in V_g,\ \phi\in\Phi_g,\ k\in K(\phi) \tag{11}$$

$$f_\phi^v \le\ m_\phi^v,\ \forall\ v\in V_g,\ \phi\in\Phi_g \tag{12}$$

$$\sum_{\phi\epsilon\Phi_a} f_\phi^v \le\ n_a^v,\ \forall\ v\in V_g,\ a\in A \tag{13}$$

$$\sum_{v\in\mathrm{V}_g}\sum_{\phi\in\Phi_g}\sum_{(\mathrm{i,p})\in\ \mathrm{L}(\Phi)} y_\phi^v(\mathrm{i,p,d}) - \sum_{v\in\mathrm{V}_g}\sum_{\phi\in\Phi_g}\sum_{(\mathrm{p,j})\in\ \mathrm{L}(\Phi)} y_\phi^v(\mathrm{p,j,d}) = \begin{cases} \sum_{\mathrm{o}\in\mathrm{O}_g} TD_{odg}, & \text{for p} = d\in D_g \\ -TD_{odg}, & \text{for p} = \mathrm{o}\in\mathrm{O}_g \\ 0, & \text{otherwise} \end{cases} \forall d\in D_g \tag{14}$$

$$\sum_{v\in V_g}\sum_{\phi\in\Phi_g}\sum_{(\mathrm{i,p}),\ (\mathrm{p,j})\in\mathrm{L}(\Phi_g)}\sum_{d\in D_g}\left(y_\phi^v(\mathrm{i,p,d}) + y_\phi^v(\mathrm{p,j,d})\right) \le cap_p^g, \forall \mathrm{p}\in\mathrm{P} \tag{15}$$

(8)

Removing transshipment activities from the formulation for dry and liquid bulk cargo types results in greater efficiency. These reduced subproblems are given as follows.

Subproblem (g, dry or liquid cargo):

Objective function

Minimize

$$\sum_{v\in V_g}\sum_{\phi\in\Phi_g} c_\phi^v * f_\phi^v + \sum_{\mathrm{p}\in\mathrm{P}} h_p * \sum_{v\in V_g}\sum_{\phi\in\Phi_g}\sum_{(\mathrm{o,p}),(\mathrm{p,d})\in\ L(\Phi_g),o\ \in O_g,d\in D_g}\left(y_\phi^v(o,\mathrm{p,p}) + y_\phi^v(\mathrm{p,d,d})\right) \tag{16}$$

Subject to

$$y_{\phi}^{v}(\mathrm{i,d,d}) \ \leq \mathrm{capv_d} * f_{\phi}^{v}, \forall\ v \in V_g,\ \phi \in \Phi_g, d \in D_g,\ (\mathrm{i,d}) \in\ \mathrm{L}\left(\Phi_g\right) \tag{17}$$

$$\sum_{l \in \mathrm{L}(\phi)} \zeta_{kl\phi} * \ y_{\phi}^{v}(\mathrm{i,d,d}) \leq \mathrm{cap}_v * f_{\phi}^{v},\ \forall\ v \in V_g,\ \phi \in \Phi_g,\ d \in D_g, k \in K(\phi) \tag{18}$$

$$\sum_{v \in \mathrm{V}_g} \sum_{\phi \in \Phi_g} \sum_{(\mathrm{o,p}) \in \mathrm{L}\left(\Phi_g\right)} y_{\phi}^{v}(\mathrm{o,p,p}) - \sum_{v \in \mathrm{V}_g} \sum_{\phi \in \Phi_g} \sum_{(p,d) \in L\left(\Phi_g\right)} y_{\phi}^{v}(\mathrm{p,d,d}) = \begin{cases} \sum_{\mathrm{o} \in \mathrm{O}_g} TD_{odg}, & \text{for p} = d \in D_g \\ -TD_{odg}, & \text{for p} = \mathrm{o} \in \mathrm{O}_g \forall d \in \\ 0, & \text{otherwise} \end{cases} \tag{19}$$

$$\sum_{v \in V_g} \sum_{\phi \in \Phi_g} \sum_{(\mathrm{o,p}),(\mathrm{p,d}) \in \mathrm{L}\left(\Phi_g\right), o \in O_g, d \in D_g} \left(y_{\phi}^{v}(o,\mathrm{p,p}) + y_{\phi}^{v}(\mathrm{p,d,d})\right) \leq cap_p^g, \forall \mathrm{p} \in \mathrm{P} \tag{20}$$

(8), (12), (13)

5.1.2. Efficiency and accuracy analysis

While the heuristic Relax-and-Fix Decomposition approach does not guarantee optimality, it was found to provide high-quality solutions with improved efficiency as compared with exact solution. To evaluate the effectiveness and efficiency of this method, the heuristic was applied on the Global Cargo Shipping Network for instances involving between 3 and 22,339 (the complete set of) randomly selected O-D pairs with associated demands.

Solutions were compared to optimal solutions obtained from Gurobi with a specified 5% permissible optimality gap. Results of this comparison are given in Table 7 and show that the gap between the optimal and heuristic solutions was less than 3% while requiring only a small fraction of the run time (a 70% reduction for the complete network). The heuristic, thus, provides solution at the global scale with the accuracy of the exact technique.

5.2. Model calibration by Heuristic Gradient Descent

The heuristic Relax-and-Fix Decomposition approach is embedded within a calibration methodology, a heuristic Gradient Descent method. This optimization-calibration solution methodology iterates over updated values of cost parameters $h_p\ (p \in P)$ in the Strategic Cargo Routing model objective function. At each iteration, the heuristic Relax-and-Fix Decomposition approach provides the optimal or near-optimal global maritime traffic patterns across shipping routes and ports for the given cost parameters. Calibration adjusts the cost parameters such that the assignment of flows to routes (i.e. the choice of routes or global maritime traffic patterns) provides consistency with real-world port throughput measurements, specifically, publicly available annual port throughputs. That is, the calibration process minimizes the differences between estimated and observed port throughputs by iteratively modifying port-specific cost parameters, h_p, and determining the associated flows. The final set of flows obtained once these differences are sufficiently small creates the final solution.

Let $\pi = \{h_p,\ p \in P\}$, i.e., the set of port cost parameters in objective function (1). Port throughput for port p is given by $Q_p^{'}\ (p \in P)$ and is estimated in Eq. (21). In this equation, leg flows are defined as $y = \{y_{\phi}^{v}(\mathrm{i,j,d}),\ v \in V,\ \phi \in \Phi,\ (\mathrm{i,j}) \in\ \mathrm{L}(\Phi),\ d \in D\}$ and are obtained by solving the Strategic Cargo Routing Model for a given π. Observed port throughputs are given by $Q_p\ (p \in P)$.

$$Q_p^{'} = \sum_{v \in V} \sum_{\phi \in \Phi} \sum_{(\mathrm{i,p}),(\mathrm{p,j}) \in\ \mathrm{L}(\Phi)} \sum_{d \in D} \left(y_{\phi}^{v}(\mathrm{i,p,d}) + \ y_{\phi}^{v}(\mathrm{p,j,d})\right) \tag{21}$$

$$Min\ W(\pi) = \ \frac{1}{|P|} \sum_{p=0}^{P} \left(Q_p^{'}(\pi) - Q_p\right)^2 \tag{22}$$

The objective of the calibration (Eq. (22)) is to determine an optimal setting of parameter set π to minimize $W(\pi)$, the average

Table 7
Comparison of Relax-and-Fix Decomposition heuristic approach and exact method.

#O-D Pairs			Run time (s)		Objective value from solution		Optimality gap
Containerized cargo	Dry bulk cargo	Liquid bulk cargo	Gurobi	Relax-and-Fix Decomposition	Gurobi	Relax-and-Fix Decomposition	$100\% * \frac{\lvert \mathrm{obj_{gurobi}} - \mathrm{obj_{decomposition}} \rvert}{\mathrm{obj_{gurobi}}}$
1	1	1	69	37	189	189	0
5	5	5	164	65	550	542	1.34%
10	10	10	270	289	897	895	0.16%
50	50	50	2716	1015	4365	4243	2.78%
1000	1000	1000	6226	816	98,575	97,931	0.65%
11,795	5871	4673	9610	2720	639,961	657,998	2.82%

square of differences in port throughput between estimated outputs and observations. In the calibration process, the parameter set is iteratively updated using approximate gradients, improving the objective function until meeting convergence criteria. An approximation of the gradient is needed, because there is no closed form equation to represent the relationship between the cost parameters, π, and $W(\pi)$. Lin (2011) proposed an effective heuristic descent direction method for solving such an optimization problem with approximate gradients and verified the efficacy of their solution strategy. Their heuristic gradient descent method is adopted herein.

The cost parameter set $\pi = \{h_p,\ p \in P\}$ has 161 parameters corresponding to 161 ports in the Global Cargo Shipping Network. In the cost parameter set, a parameter π_p^t, the p_{th} parameter at iteration t, is updated over the iterations following Eq. (23), where $\nabla g(\pi_p^t)$ is the approximate gradient for the given parameter π_p^t, and λ^t is the step size at iteration t. The gradient $\nabla g(\pi_p^t)$ is updated using small increases in parameter settings (+1 in Eq. (24)).

$$\pi_p^{t+1} = \pi_p^t + \lambda^t \nabla g\left(\pi_p^t\right) \tag{23}$$

$$\nabla g\left(\pi_p^t\right) = \frac{\partial W\left(\pi_p^t\right)}{\partial \pi_p^t} \approx W(\pi_p + 1) - W(\pi_p) \tag{24}$$

The optimization-calibration solution methodology, including the additional steps taken to improve computational efficiency, is depicted in Fig. 6. Two specific steps are taken to improve efficiency. First, at each iteration, a subset Ω of P is identified for which the estimated throughputs are equivalent or nearly so to observed values. As tests indicated that these parameters were unlikely to change in future iterations, the cost parameters of these ports are fixed in moving to the next iteration. Second, when calculating the approximate gradient $W(\pi_p + 1)$ for each port p, only those subproblems related to cargo types that port p handles must be included in re-solving the Strategic Cargo Routing Model on parameters $(\pi_p + 1)$. Decision variables of other subproblems associated with port p are set based on solution of the prior iteration with parameters π_p.

6. Computational runs and global vessel flow estimation

The optimization-calibration solution algorithm was implemented in Python and run using on the Global Maritime Network. The run was conducted on a machine operated by macOS system with 2.6 GHz 6-Core Intel Core i7 of processor and 16 GB of memory. The run required multiple days. Practically, the run time would be expected to grow linearly with the number of ports.

6.1. Parameter calibration

The optimization-calibration solution methodology when applied on the Global Maritime Network converged in three iterations, generating maritime shipping flow estimates that fit well with the available observations. The objective value (Best W) in Eq. (22),

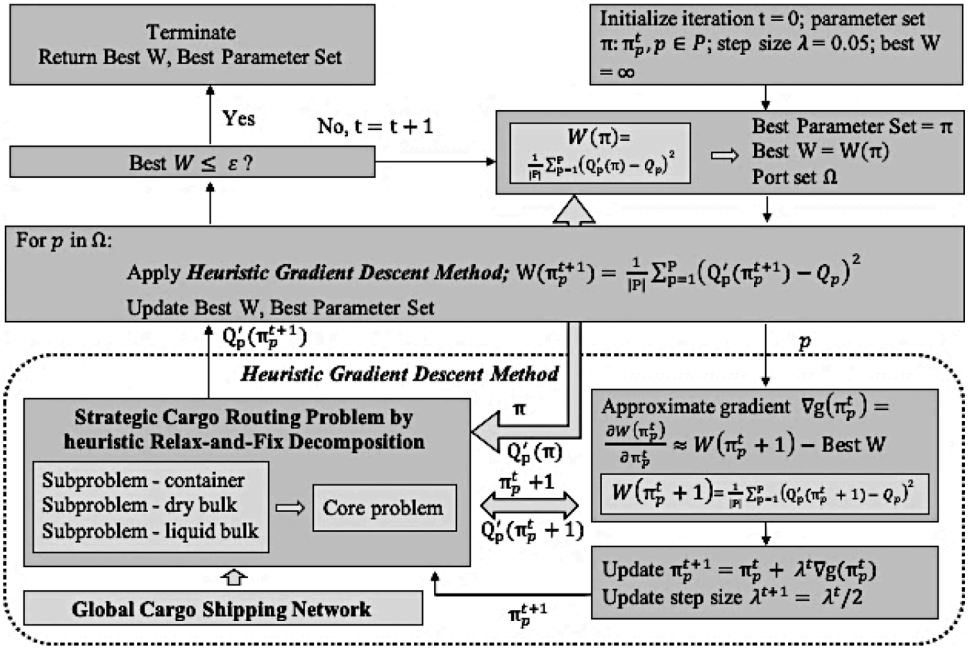


Fig. 6. Optimization-calibration solution framework.

indicating the average square of differences in port throughput values, decreased from 736 in iteration 1 to 490 in iteration 2. No notable improvement was achieved in iteration 3; thus, the procedure terminated.

At the final iteration, estimated port throughputs were calculated via Eq. (21) using the final leg flow estimates obtained from solution of the model. Appendix B lists observed and estimated port throughputs, along with their absolute differences. Tavasszy et al. (2011) and Lin and Huang (2017) similarly used the average absolute differences of port throughputs to demonstrate accuracy and precision of their models.

The average absolute difference was approximately 9 million tons (or 9% of average port throughput). Thus, the calibrated solution achieves a 91% fit overall by average annual port throughput. This finding indicates that the solution methodology provides reasonable estimates of global maritime traffic patterns.

The model was also tested with initial values of port handling costs that are lower at the smaller ports (15 $/ton) and higher at larger ports (25 $/ton). The optimal objective value (Best W) was 614, and calibration achieves a 89% fit overall by average annual port throughput. Thus, the change in initial values has some implication for the final values, and the presumed standard cost from the literature outperformed the approach using initial values that were differentiated by port size. The final, calibrated port handling cost parameters are listed in Table 8. Those parameters not included in the table remained at the initial value of 20 $/ton.

6.2. Global maritime traffic patterns

Estimated annual port throughput and shipping flows over routes for the year 2018 were obtained from the calibrated model and are visualized in Figs. 7 (estimated port throughput), 8 (vessel transits over routes) and 9 (total flows by weight over routes).

Fig. 7 suggests that the busiest ports are found in the Far East, West coast of Australia, the Netherlands, and the Gulf of Mexico. The figure also indicates differences between estimated and observed annual throughputs in Asia, where they are underestimated, and the East Coast of North America, a few countries along the Mediterranean, and areas alongside the English Channel, where they are overestimated. Similar underestimation in Asian ports was obtained by Tavasszy et al. (2011).

Overestimation at some ports is reasonable, as smaller ports are excluded from the network and their flows are likely captured at nearby ports included within the model. For example, the throughput of the Port Authority of New York and New Jersey is overestimated by approximately 21 million tons annually. This is consistent with the annual port throughput of 26 million-tons at nearby Port of Philadelphia, a port that is excluded from the Global Cargo Shipping Network. Similarly, overestimation of annual throughputs at the Ports of Singapore and Seattle is likely due to their strong network connectivity as a consequence of their geographical locations. Generally, though, the estimates are close to the observed values and the calibration is presumed to be successful.

Annual vessel transits and total flow by weight aligned with the estimated port throughputs are depicted in Figs. 8 and 9. The flows appear to be reasonable, with greater numbers of transits along shorter routes and between the busiest continents (Asia, Europe and North America).

To further validate the model, annual estimated and observed vessel transits through the Suez and Panama Canals were compared for the year 2018. Observed data were collected from the Canals' websites. Table 9 compares vessel transits by containerships, bulkers and tankers. Bulk cargo trade pairs as included in the Global Maritime Network account for 65% of world bulk cargo seaborne trade by weight. Estimated transits of bulkers and tankers were, thus, scaled proportionally to model total world bulk cargo trade. When scaled appropriately to total global cargo volumes, the total estimated number of transits matches very closely to observed transits in the two canals with one exception. Transits by bulkers and tankers through the Panama Canal are substantially underestimated. This is likely because the Global Cargo Shipping Network excludes some key shipping trade lanes served by the Panama Canal, including flows that might be completed using the Canal within and across the land masses of North and South America.

7 Conclusions and discussion

7.1 Conclusions

The vast majority of studies in which models of global seaborne trade are proposed have focused on container shipping. The networks on which these trade flows move are either highly aggregated, more detailed but regional, or are specific to a particular good

Table 8
Calibrated port handling cost parameters using constant initial generalized port handling costs, h_p, of 20 $/ton; all ports excluded from the table had final values of h_p that remained at 20 $/ton.

Area	Ports	Calibrated h_p ($/ton)
Asia	Guangzhou	18.53
	Hong Kong	13.75
	Xiamen	20.38
	Nhava Sheva	18.66
	Keelung	24.22
North America	Freeport	19.46
	Vancouver	21.54
Mediterranean	Barcelona	20.37
	La Spezia	17.80

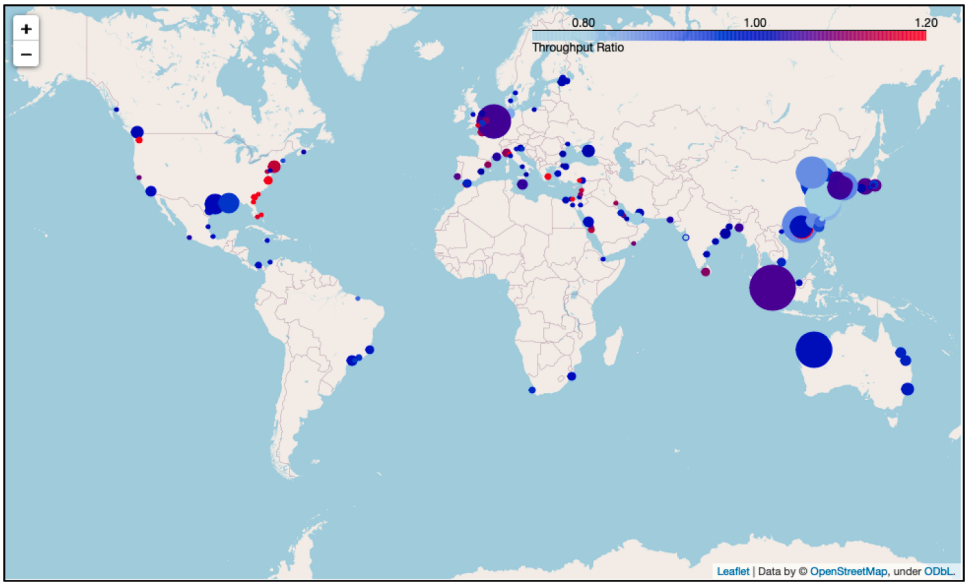


Fig. 7. Annual port throughputs: (i) estimated port throughput - radius proportional to total weight and (ii) throughput ratio – estimated to observed for 2018.

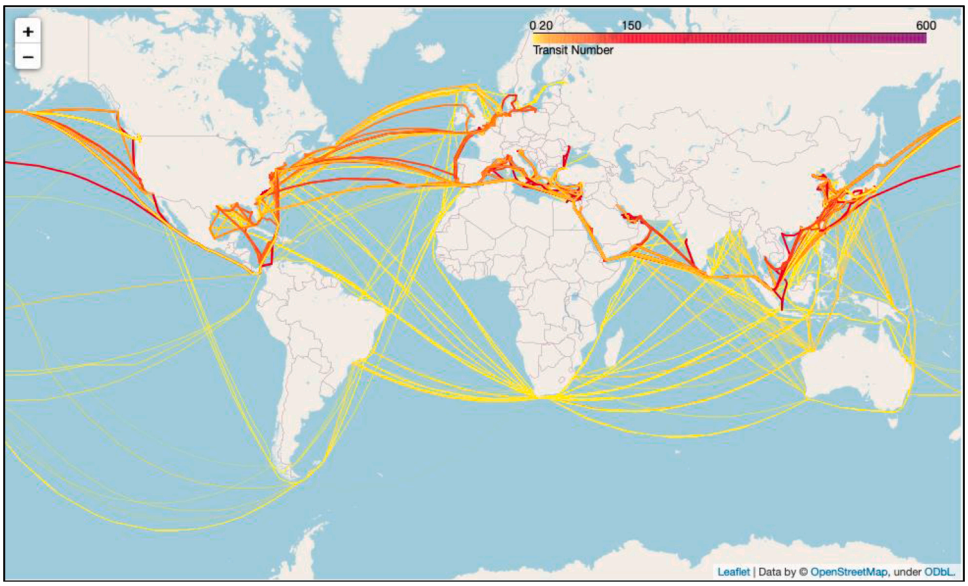


Fig. 8. Estimated number of annual transits over routes, the darker the color, the greater the number of transits.

type or shipping company. Few works have investigated global bulk shipping. Despite that seaborne container and bulk cargo carrying vessels share the shipping lanes and ports, no prior model in which both markets are treated simultaneously has been suggested. This paper addresses the problem of estimating annual world seaborne trade flows of both containerized and bulk cargo. To this end, a high-fidelity Global Cargo Shipping Network with 161 ports is constructed on publicly available, and updatable data. The Strategic Cargo Routing model was formulated for estimating annual seaborne trade flows of both container and dry and liquid bulk cargo. An iterative optimization-calibration solution approach that combines heuristic Gradient Descent and Relax-and-Fix Decomposition methods provides a solution scheme for obtaining the desired global trade flow estimates. Final network details are shared on Github to support others in conducting global maritime transport research. The calibrated solution achieves a 91% fit overall by average annual port throughputs, and thus, can be used to study the global maritime system and the effects of changes in various aspects affecting world seaborne trade into the future.

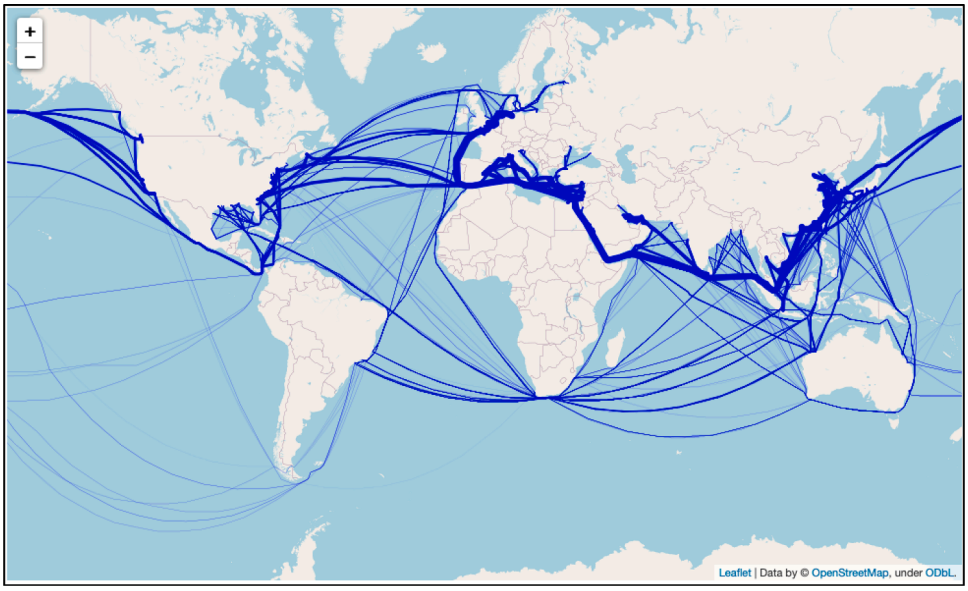


Fig. 9. Estimated tons of cargo transported across routes annually, the thicker the line the greater the number of tons transported.

Table 9
Comparison of observed vessel transits through Suez and Panama Canals with estimated transits.

	Suez Canal		Panama Canal	
	Containerships	Bulkers & tankers	Containerships	Bulkers& tankers
Observed transits	5706	9236	2604	6714
Estimated transits	5818	6236	2724	2009
Scaled to global level	5818	9594	2724	3091

7.2 Discussion

Based on findings from evaluating the proposed calibrated strategic cargo routing model with generalized port handling costs, the model can: (1) be used to evaluate and predict the performance of global maritime systems; (2) can provide port throughput estimates that capture port operational outcomes, approximating inbound, outbound and transshipment flows; and (3) can reasonably estimate global maritime traffic patterns that provide a snapshot of global shipping flows.

In future work, the Global Cargo Shipping Network could be expanded to consider more refined cargo classes, additional vessel types beyond the 10 included, greater than four shipping subnetworks, and/or more detailed port models with separated terminals. The network model could also include additional ports, route alternatives and on-land connections.

The Strategic Cargo Routing Model relies on annual demand and supply estimates, and thus, can only provide annual flow predictions. Moreover, it presumes deterministic knowledge of costs, including voyage times, port handling times, and an uncongested system. That is, it is strategic in nature and operational details, such as wait times in anchorage areas, are not explicitly modeled. For relevant applications, the model can be extended to account for operational details by incorporating additional aspects of the network, such as anchorage area and navigation channel details, terminal operations and conditions that affect the specific route navigated. Incorporating such details, however, is challenging first because obtaining the added data may be very difficult and expensive and second because the expansion of the model will likely create problem instances whose sizes make them mathematically intractable to solve. It is imperative, thus, to ensure that any model extensions maintain linearity in the formulation's objective and constraints so that the heuristic Relax-and-Fix Decomposition method may still apply.

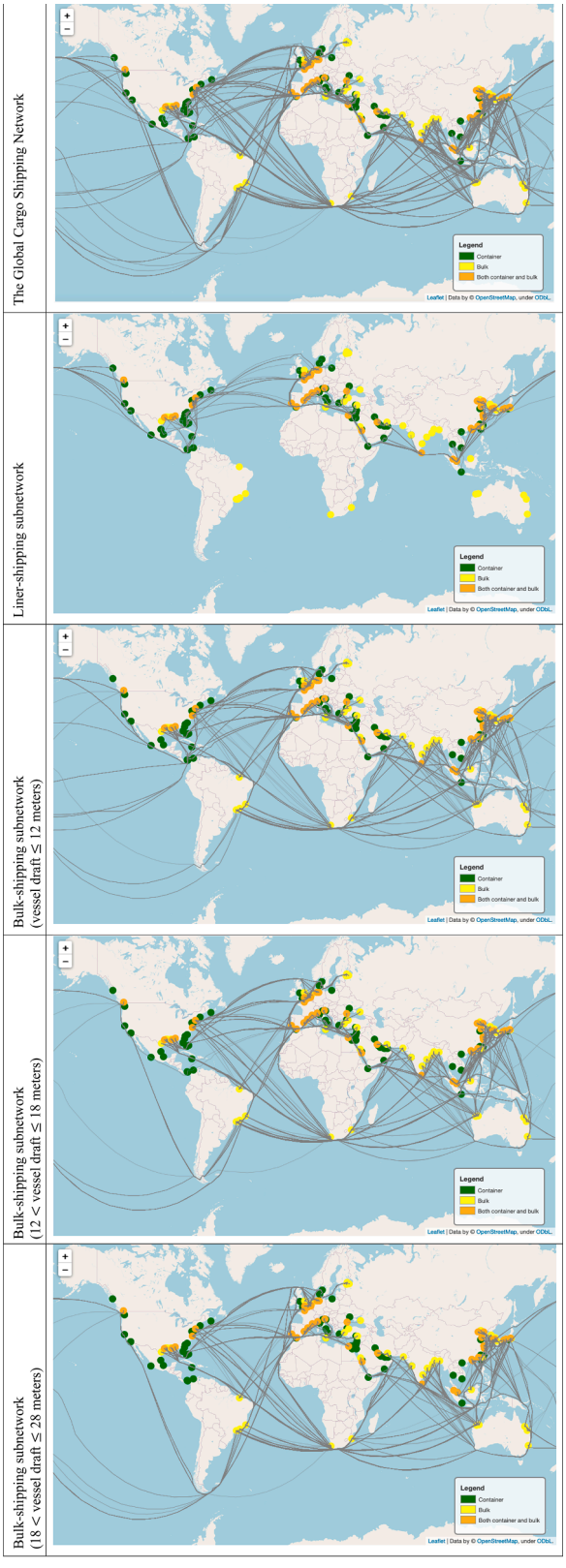
Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by **REMOVED to maintain anonymity**. This funding is gratefully acknowledged but implies no endorsement of the findings.

Appendix A. – Global cargo shipping network



Appendix B. Average absolute difference of port throughputs

Port	Observed port throughput (million tons)	Estimated port throughput (million tons)	Absolute difference
Dampier	175.5	172.1	3.4
Gladstone	124	120.5	3.5
Hay Point	118.3	114.8	3.5
Newcastle	165.1	161.8	3.3
Port Hedland	512.9	506.2	6.7
Freeport	12.6	15.1	2.5
Bahrain	5.2	5.1	0.1
Chittagong	89.8	94.2	4.4
Antwerp	418.8	398.3	20.5
Zeebrugge	19.5	19.4	0.1
Itaguai	56.6	54.1	2.5
Itaqui	22.4	20.1	2.3
Santos	107.1	105.6	1.5
Sao Sebastiao	44	41.3	2.7
Tubarao	103.9	102.3	1.6
Halifax	6.6	7.5	0.9
Prince Rupert	12.4	12.4	0.0
Vancouver	147.1	148.5	1.4
Dalian	467.8	375.5	92.3
Fuzhou	40.8	30.4	10.4
Guangzhou	594	506.5	87.5
Hong Kong	258.5	310.2	51.7
Lianyungang	57	42.4	14.6
Ningbo	576.5	465.3	111.2
Qingdao	542.5	503.8	38.7
Qinhuangdao	231.2	208.5	22.7
Shanghai	683.9	529.2	154.7
Shenzhen	308.9	321.8	12.9
Tianjin	507.7	436.5	71.2
Xiamen	217.2	191.7	25.5
Cartagena	34.5	34.6	0.1
Aarhus	6.5	6.4	0.1
Djibouti	10.2	10.2	0.0
Alexandria	63.4	63.9	0.5
Damietta	13.8	14.1	0.3
Port Said	37.3	44.8	7.5
Sokhna	7.9	8.0	0.1
Calais	47	47.7	0.7
Dunkirk	51.6	52.1	0.5
Le Havre	71.72	81.3	9.5
Fos	80.5	82.8	2.3
Bremerhaven	74.4	67.1	7.3
Hamburg	134	111.8	22.2
Piraeus	58.9	69.2	10.3
Bombay	60.6	61.5	0.9
Calcutta	63.8	64.9	1.1
Chennai	53.01	53.5	0.5
Nhava Sheva	61.6	41.1	20.5
Paradip	110	113.4	3.4
Visakhapatnam	65.3	66.5	1.2
Jakarta	93.6	62.4	31.2
Bandar Abbas	77.9	78.9	1.0
Umm Qasr	4.8	4.9	0.1
Ashdod	17.7	18.2	0.5
Haifa	17.6	20.0	2.4
Genoa	69.2	77.3	8.1
Gioia Tauro	27.9	28.7	0.8
La Spezia	17.8	21.4	3.6
Leghorn	9	8.9	0.1
Naples	7	7.1	0.1
Trieste	63.5	64.8	1.3
Venice	7.6	7.5	0.1
Kingston	22	22.0	0.0
Chiba	153.2	164.0	10.8
Kitakyushu	102	108.3	6.3
Kobe	95	101.0	6.0
Nagoya	197	209.6	12.6
Osaka	84.3	86.0	1.7

(*continued on next page*)

(*continued*)

Port	Observed port throughput (million tons)	Estimated port throughput (million tons)	Absolute difference
Shimizu	6.8	6.7	0.1
Tokyo	91.5	83.5	8.0
Yokohama	36.4	37.9	1.5
Aqaba	9.8	9.8	0.0
Beirut	15.7	17.7	2.0
Bintulu	48.34	49.2	0.9
Port Kelang	220.9	262.9	42.0
Tanjung Pelepas	107.5	89.7	17.8
Malta	39.8	47.8	8.0
Altamira	9.8	10.0	0.2
Manzanillo	36.9	37.9	1.0
Veracruz	10.8	10.8	0.0
Tangiers	41.7	50.0	8.3
Amsterdam	82.3	84.4	2.1
Rotterdam	469	496.9	27.9
Salalah	40.6	43.4	2.8
Sohar	16.5	16.5	0.0
Karachi	65.25	67.5	2.3
Balboa	30.3	36.4	6.1
Colon	51.9	52.3	0.4
Manila	107.6	112.4	4.8
Gdansk	23.4	23.7	0.3
Sines	47.9	50.1	2.2
Hamad	22	22.6	0.6
Constantza	61.3	62.8	1.5
Novorossiysk	154.9	157.5	2.6
Primorsk	53.5	53.7	0.2
Saint Petersburg	59.3	59.8	0.5
Ust-Luga	98.7	100.1	1.4
Dammam	18.8	20.3	1.5
Jeddah	54.8	61.0	6.2
Jubail	68.2	66.6	1.6
King Abdullah	27.6	29.5	1.9
Yanbu	123.6	124.0	0.4
Singapore	630.1	672.9	42.8
Koper	11.9	11.9	0.0
Rijeka	3.1	3.0	0.1
Richards Bay	103.4	102.2	1.2
Saldanha Bay	63.4	61.1	2.3
Busan	461.5	413.5	48.0
Daesan	92.6	96.5	3.9
Incheon	163.6	172.4	8.8
Kwangyang	301.9	316.9	15.0
Pohang	60.5	61.9	1.4
Pyeongtaek	115.1	120.7	5.6
Ulsan	202.9	214.0	11.1
Algeciras	107.2	106.6	0.6
Barcelona	66	73.7	7.7
Valencia	70.8	71.5	0.7
Colombo	84	92.1	8.1
Gothenburg	9	9.0	0.0
Kaohsiung	125.4	115.0	10.4
Keelung	17.7	13.1	4.6
Taichung	129.4	116.0	13.4
Taipei	19.9	14.8	5.1
Laem Chabang	96.8	68.4	28.4
Aliaga	53.3	53.1	0.2
Botas	60.7	60.8	0.1
Istanbul Ambarli	38.3	40.1	1.8
Izmit	72.4	73.1	0.7
Mersin	20.7	24.8	4.1
Odessa	10.2	10.2	0.0
Abu Dhabi	20.9	20.9	0.0
Dubai	179.5	137.1	42.4
Felixstowe	49.9	54.6	4.7
Immingham	55.6	57.4	1.8
Liverpool	9.8	9.7	0.1
London	53.2	50.7	2.5
Southampton	23.9	27.6	3.7
Baltimore	12.3	14.8	2.5

(*continued on next page*)

(*continued*)

Port	Observed port throughput (million tons)	Estimated port throughput (million tons)	Absolute difference
Beaumont	100.2	99.1	1.1
Boston	3.6	3.6	0.0
Charleston	27.8	33.4	5.6
Corpus Christi	93.5	92.3	1.2
Houston	268.9	275.9	7.0
Jacksonville	15.3	18.4	3.1
Lake Charles	56.9	55.2	1.7
Long Beach	97.1	98.6	1.5
Los Angeles	113	103.3	9.7
Miami	13	15.6	2.6
Mobile	58.6	56.7	1.9
New Orleans	93.3	92.9	0.4
New York	140.3	161.9	21.6
Norfolk	71.8	86.2	14.4
Oakland	30.6	36.7	6.1
Plaquemines	56.9	55.2	1.7
Savannah	52.2	62.6	10.4
Seattle	45.6	54.7	9.1
South Louisiana	275.5	260.4	15.1
Tampa	1.1	1.2	0.1
Wilmington	4.6	4.7	0.1
Cai Mep	79	75.9	3.1
Haiphong	13.9	13.8	0.1
Average absolute difference	9.04%		

References

Fremont, A., 2007. Global maritime networks: the case of Maersk. J. Transp. Geogr. 15 (6), 431–442.

Agarwal, R., Ergun, O., 2008. Ship scheduling and network design for cargo routing in liner shipping. Transp. Sci. 42 (2), 175–196.

Agarwal, R., Ergun, O., 2010. Network design and allocation mechanisms for carrier alliances in liner shipping. Oper. Res. 58 (6), 1726–1742.

Álvarez, J.F., 2009. Joint routing and deployment of a fleet of container vessels. Mar. Econ. Logist. 11 (2), 186–208.

Alizadeh, A.H., Talley, W.K., 2011. Microeconomic determinants of dry bulk shipping freight rates and contract times. Transportation (Amst) 38, 561–579.

Achurra-Gonzalez, P.E., Novati, M., Foulser-Piggott, R., Graham, D.J., Bowman, G., Bell, M.G.H., Angeloudis, P., 2019. Modelling the impact of liner shipping network perturbations on container cargo routing: southeast Asia to Europe application. Accid. Anal. Prevent. 123, 399–410.

Bilgen, B., Ozkarahan, I., 2007. A mixed-integer linear programming model for bulk grain blending and shipping. Int. J. Prod. Econ. 107 (2), 555–571.

Bell, M.G.H., Liu, X., Rioult, J., Angeloudis, P., 2013. A cost-based maritime container assignment model. Transp. Res. Part B: Methodol. 58, 58–70.

Brouer, B.D., Alvarez, J.F., Plum, C.E.M., Pisinger, D., Sigurd, M.M., 2014. A base integer programming model and benchmark suite for liner-shipping network design. Transp. Sci. 48 (2), 281–312.

Christiansen, M., Fagergolt, K., Rachaniotis, N.P., Stalhane, M., 2017. Operational planning of routes and schedules for a fleet of fuel supply vessels. Transp. Res. Part E: Logist. Transp. Rev. 105, 163–175.

De, I., Rivelino, R., Gregory, S.P., Edward, A.P., 2019. Gulf Coast port selection using multiple-objective decision analysis. Decis. Anal. 16 (2), 87–104.

Fan, L., Wilson, W.W., Tolliver, D., 2010. Optimal network flows for containerized imports to the United States. Transp. Res. Part E: Logist. Transp. Rev. 46 (5), 735–749.

GlobalSecurity, 2021. Bulk cargo carrier. https://www.globalsecurity.org/military/systems/ship/bulk.htm, Last accessed on April 25, 2021.

International Transport Forum, 2018. The impacts of alliances in container shipping. https://www.itf-oecd.org/sites/default/files/docs/impact-alliances-container-shipping.pdf, last accessed on April 25, 2021.

International Chamber of Shipping, 2021. https://www.ics-shipping.org/shipping-facts/shipping-and-world-trade, last accessed on April 25, 2021.

Amadeo, K., 2020. Trade wars and their effect on the economy and you. https://www.thebalance.com/trade-wars-definition-how-it-affects-you-4159973, last accessed on April 25, 2021.

Lin, D.Y., Liu, H.Y., 2011. Combined ship allocation, routing and freight assignment in tramp shipping. Transp. Res. Part E: Logist. Transp. Rev. 47 (4), 414–431.

Lin, D.Y., 2011. A dual variable approximation-based descent method for A Bi-level continuous dynamic network design problem. Comput. Aided Civ. Infrastruct. Eng. 26 (8), 581–594.

Li, Z., Xu, M., Shi, Y., 2015. Centrality in global shipping network basing on worldwide shipping areas. GeoJournal 80 (1), 47–60.

Lin, D.Y., Huang, K.L., 2017. An equilibrium-based network model for international container flows. Marit. Policy Manag. 44 (8), 1034–1055.

Lin, D.Y., Chang, Y.T., 2018. Ship routing and freight assignment problem for liner shipping: application to the Northern Sea Route planning problem. Transp. Res. Part E: Logist. Transp. Rev. 110, 47–70.

Meng, Q., Wang, S., Henrik, A., Kristian, T., 2014. Containership routing and scheduling in liner shipping: overview and future research directions. Transp. Sci. 48 (2), 265–280.

Maersk, 2020. Structural network changes on transpacific service due to Covid-19 impact. https://www.maersk.com/news/articles/2020/03/31/structural-network-changes-on-transpacific-services-due-to-covid-19-impact, last accessed April 25, 2021.

Noor-E-Alam, M., Doucette, J., 2015. Solving large-scale fixed cost integer linear programming models for grid-based location problems with heuristic techniques. Eng. Optim. 47 (8), 1085–1106.

Maroday, P., 2013. 9 facts about the panama canal expansion-infographic. http://www.mercatrade.com/blog/9-facts-about-the-Panama-Canal-Expansion.

Seedah, D., Harrison, R., Boske, L., Kruse, J., 2013. Container terminal and cargo-handling cost analysis toolkit. No. 0-6690-CTR-P2.

Shibasaki, R., Iijima, T., Kawakami, T., Kadono, T., Shishido, T., 2017. Network assignment model of integrating maritime and hinterland container shipping: application to Central America. Marit. Econ. Logist. 19, 234–273.

Sun, Z., Zhang, R., Gao, Y., Tian, Z., Zuo, Y., 2020. Hub ports in economic shocks of the melting Arctic. Marit. Policy Manag. 1464–5254.

Tavasszy, L., Minderhoudb, M., Perrinc, J.F., Notteboom, T., 2011. A strategic network choice model for global container flows: specification, estimation and application. J. Transp. Geogr. 19 (6), 1163–1172.

UNCTAD, 2015. Review of Maritime Transport. United Nations publication, New York and Geneva.

UNCTAD, 2019. Review of Maritime Transport. United Nations publication, New York and Geneva.
U.S. Department of Transportation (USDOT), 2017. Bureau of Transportation Statistics, Port performance freight statistics program, annual report to congress.
Xu, M., Li, Z., Shi, Y., Zhang, X., Jiang, S., 2015. Evolution of regional inequality in the global shipping network. J Transp Geogr 44, 1–12.
Yang, D., Wang, S., 2017. Analysis of the development potential of bulk shipping network on the Yangtze River. Marit. Policy Manag. 44 (4), 512–523.
Wang, H., Nozick, L., Xu, N., Gearhart, J., 2018. Modeling ocean, rail, and truck transportation flows to support policy analysis. Marit. Econ. Logist. 20, 327–357.
USDOT, 2021. TransBorder freight dataset. https://www.bts.gov/transborder, last accessed April 25, 2021.
European Sustainable Shipping Forum, Sub-group on Shipping MRV Monitoring, 2021. Draft guidance document the shipping MRV regulation – determination of cargo carried. https://ec.europa.eu/clima/sites/clima/files/docs/0108/20170517_guidance_cargo_en.pdf, last accessed April 25, 2021.