Transportation Research Part E 47 (2011) 593–608

Contents lists available at ScienceDirect

Transportation Research Part E

journal homepage: www.elsevier.com/locate/tre

Long- and Short-Run supply-chain optimization models for the allocation and congestion management of containerized imports from Asia to the United States

Payman Jula [a,*], Robert C. Leachman [b]

[a] *Technology and Operations Management, Faculty of Business Administration, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada V5A 1S6*
[b] *Department of Industrial Engineering and Operations Research, University of California at Berkeley, Berkeley, CA 94720-1777, USA*

ARTICLE INFO

Article history:
Received 4 August 2010
Received in revised form 11 August 2010
Accepted 29 January 2011

Keywords:
Supply-chain optimization
Intermodal freight transport
Transportation planning
Containerized imports
Mixed integer non-linear programming
Queuing

ABSTRACT

Analytical models are introduced for predicting the allocation to ports and transportation channels of containerized goods imported from Asia to the USA. Assuming fixed distributions for container flow-times, the Long-Run Model heuristically solves a mixed integer non-linear program to determine the least-cost supply-chain strategies for importers. The Short-Run Model uses estimates of the flow times as a function of traffic volumes on fixed infrastructure to iteratively develop the best near-term strategies. Minimum volume requirements and capacities for ports and landside channels are considered. The models are applied to predict the effects of container fees at the San Pedro Bay ports.

1. Introduction

This article proposes two analytical models for optimization of the supply chains of importers of waterborne containerized goods from Asia to USA. The basic aim of these models is to predict the distribution of Asia – United States waterborne import volumes by port and transportation modes (channels) for a given distribution of total import volume by declared value, pre-specified potential port fees and minimum required port volumes, assumed inventory holding cost rates, assumed geographical distribution of import destinations and given transportation and handling rates.

The first model, termed the Long-Run Model, is a large mixed integer non-linear programming model, and a set of heuristics to solve that. The objective is to minimize the total costs for transportation and handling, pipeline inventory, and safety-stock inventories. In the Long-Run Model, the mean and standard deviation of container flow times by channel are fixed, reflecting an assumption that over the long term the various ports and transportation carriers would make investments to maintain existing service quality and thereby protect market share.

In contrast, the second model, termed the Short-Run Model, assumes the infrastructure of the entire transportation network is fixed. Container flow times are endogenous in this model, responding to congestion (or lack thereof) in various ports and channels, thereby capturing the effects of capacity constraints. The Short-Run Model is thus useful for projecting more near-term responses of importers to changes in fees, rates or infrastructure. The Short-Run Model uses the Long-Run model as a component and integrates it with a set of analytical queuing models which estimate the import container flow times through port terminals, rail intermodal terminals and rail line-haul channels as a function of traffic volumes, infrastructure

---

* Corresponding author. Tel.: +1 778 782 3564; fax: +1 778782 4920.
E-mail address: pjula@sfu.ca (P. Jula).

1366-5545/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tre.2011.02.006

and staffing hours. (Hereafter, we refer to the set of analytical queuing models as the Queuing Model.) Calculations of the Short-Run Model involve iterative runs of the Long-Run Model and the Queuing Model.

A typical large US importer/retailer operates Regional Distribution Centers (RDCs) that restock its retail outlets or retail customers. In this article, we consider the origins for import shipments to be factories in China and elsewhere in Asia, and the destinations are RDCs spread across the Continental USA. Marine containers from Asian origins are shipped on vessels to ports of entry (POE) to the USA. The containers may be directly shipped inland to the RDCs, called "direct shipment," or they may be unloaded at trans-load or import warehouse facilities and the contents sorted and re-shipped in domestic vehicles to multiple RDCs, under a strategy termed "consolidation–de-consolidation" shipment. In the consolidation–de-consolidation case, marine containers carrying goods destined to multiple RDCs are channeled through a common port and routed to a de-consolidation center (trans-load warehouse). The goods are unloaded from the marine boxes, sorted and re-loaded into domestic containers or trailers for final landside movement to the RDCs. Both direct and consolidation–de-consolidation shipments may use different land-side transportation modes (channels) to RDCs; i.e., train, truck, and local drayage (dray). Depending on the selected port of entry and the landside mode of transportation, the importer will face different transportation costs. Another source of cost is the opportunity cost of working capital tied up in the inventory throughout the supply chain. This cost is customarily expressed as an interest rate times the amount of capital invested per unit of inventory times the average inventory level.

There are three types of inventories in a chain: cycle inventory, in-transit inventory, and safety stock. Cycle inventory is a function of the replenishment frequency (e.g., weekly) and is otherwise independent of the selection of the supply-chain strategy and channel, and therefore is not considered in this article. In-transit inventory is the amount of inventory in the pipeline, and it is a function of the transportation time. Safety stock is the extra inventory kept by retailers to satisfy the customer demand on time. It is a function of the customer service level, the uncertainties in the shipment lead time, and the demand forecast error.

Both proposed models observe pre-specified minimum volumes that must be channeled through various ports, reflecting the requirements of prevailing contracts. Steamship lines enter into long-term (10–30 year) contracts with ports. Many of these contracts involve fixed payments and/or volume incentives. Some offer incentives for rail intermodal movement of the marine containers (as opposed to placement of containers on truck chasses). These contracts limit or delay the flexibility of steamship lines in restructuring their vessel strings or their strategies for which port to off-load cargoes destined to inland points.

These models extend previous literature for predicting import flows by incorporating formal optimization of import flows from the point of view of the importers, and by providing both Short- and Long-Run predictions of import flows. The underlying optimization model extends the existing literature in location/allocation problems by considering risk pooling economics during the selection of transportation mode and routing, and by integrating analysis of stochastic demand and random transportation times with location decisions in a multi-echelon supply chain to achieve desired customer satisfaction levels. It also provides a strategic-level and tactical-level tool for service design of intermodal freight transportation systems. Analysis using the proposed models can be valuable to many stakeholders of the supply chain including importers, public policy makers, port authorities, and transportation companies (steamship lines, trucking companies, railroads, dray companies, etc.). These stakeholders need to consider the response of all importers to changes in services, rates, fees or infrastructure, and so models that can rapidly predict import flows in what/if scenarios are of great interest.

The rest of the paper is organized as follows: in Section 2, the data and the framework used by the models are introduced. In Section 3, the methodology and the structure of models are discussed. In Section 4, an application of the models for predicting the effect of imposing container fees at the San Pedro Bay Ports (i.e., the Ports of Los Angeles and Long Beach) is discussed. Finally, conclusions, recommendations and directions for future research are presented in the last section.

2. Background

Melo et al. (2009) provide a review of recent literature on facility location models in the context of supply chain management (SCM) and report that the majority of the literature deals with deterministic environments and ignores the randomness involved in location decisions in SCM. Their survey further shows that the facility location decision is frequently combined with inventory decisions. In contrast, routing and the choice of transportation modes (alone or integrated with other types of decisions) have not received much attention. Shen and Qi (2007), Ambrosino and Scutellà (2005), Ma and Davidrajuh (2005), and Liu and Lin (2005), have considered routing decision-making simultaneous with inventory management. Wilhelm et al. (2005) have considered choice of transportation mode along with inventory management. Manzini and Bindi (2009) consider transportation mode selection along with routing and inventory management.

Inventory control policies may be included in a facility location problem to recognize risk pooling benefits due to stochastic demands or randomness in supply. This combination of tactical and strategic decisions has been addressed by a number of authors – see Snyder et al. (2007), Shu et al. (2005), Miranda and Garrido (2004), Shen et al. (2003), Daskin et al. (2002), and Erlebacher and Meller (2000). However, within the context of the location/allocation problem, there is a lack of publications which consider risk pooling simultaneously with choice of transportation mode and routing. Melo et al. (2009) report that the existing literature on location/allocation problems falls short of fully integrating the many aspects relevant to SCM. The literature integrating uncertainty in SCM with location decisions is still scarce. Many relevant tactical/operational decisions in SCM, as it is the case with routing and the choice of transportation modes, are far from being integrated with location decisions. In this article, we target these gaps by addressing a location–allocation problem involving stochastic demands and random transportation times, while considering risk pooling, routing, and choice of transportation modes to achieve desired customer service levels.

In terms of solution methodology, optimization models, especially non-linear optimization models, have been widely used to solve the minimization of the total cost that includes location costs and inventory costs at the facilities, and distribution costs in the supply chain. Shen (2007) provides a survey of recent developments in this research area. The author indicates that most of these problems are NP-hard, and therefore difficult to solve. Thus, researchers have used techniques such as Lagrangian-relaxation and genetic algorithms to provide solutions. For industrial applications, researchers mostly have focused on exploiting the properties of the problems and developing special solution techniques to reduce the computational effort.

Intermodal transportation of containerized imports to the United States has been a subject of several recent articles. For example, Fan et al. (2010) develop a linear-programming based optimization model that integrates international and North America inland transport networks. The authors consider direct shipment of marine containers (implicitly assuming consolidation–de-consolidation volumes are fixed) and make an arbitrary allocation of railroad line capacities for direct shipments of marine containers. Levine et al. (2009) formulated a linear program with an integrated gravity model to estimate the origin–destination table for US containerized imports. In this work, Transportation Analysis Zones composed of one or more Bureau of Economic Analysis economic areas are specified as US destination regions. Imai et al. (2009) compare multi-port calling served by smaller vessels with hub-and spoke operation using mega-containerships for the Asia–North America and Asia–Europe trades. They consider empty container repositioning in their research. These studies assume deterministic environments, i.e., the uncertainties in demand and transportation times haven't been addressed. Safety-stock inventories and the effect of risk pooling on container flows are ignored.

As discussed in the introduction section, the models developed in this article can influence the practices of many stakeholders of the supply chain. In the domain of intermodal freight transport systems, Caris et al. (2008) provide an overview of the planning decisions and solution methods proposed in the scientific literature and find a lack of research concerning the strategic- and tactical-level issues facing intermodal operators. Intermodal decision-makers can benefit from the proposed methodology in this article to make better decisions at strategic, tactical and operational levels.

Except for Leachman (2008, 2010), and Jula and Leachman (2011), the problem, as outlined in the preceding section of this article, has not been addressed by researchers. Leachman (2008) assume a single homogenous supply-chain strategy practiced by each importer. The author assumes a predetermined assignment of regional distribution centers (RDCs) to ports of entry within each supply-chain strategy and investigates the effect of increasing container fees at the San Pedro Bay Ports in terms of diversion of cargoes to other ports. Jula and Leachman (2011) introduce a mixed integer non-linear programming model for a single importer and develop solutions for the case of non-identical-volume RDCs. They provide details of a heuristic approach to solve the problem to near-optimality over a wide range of parameters, and provide sensitivity analysis and general recommendations for an importer to choose supply-chain strategies most suitable for their businesses. These studies did not incorporate constraints reflecting contractual agreements for minimum volumes by port in their analysis. These models implicitly assumed that infrastructure at ports and channels would be expanded as necessary to maintain current container flow times in the face of increased shares of imports routed through ports and landside channels, i.e., the capacity constraints are not addressed in their work.

In this article, first we extend the mixed integer programming model introduced by Jula and Leachman (2011) to encompass all importers. We further enhance the model to address the complexities arising from minimum contractual requirements for port volumes. Then, we integrate the proposed model with a queuing model to estimate container flow times. This integrated model is suitable for short-term analysis in which port and landside infrastructure and operating schedules are fixed, rather than fixing statistics on container flow times. In a Short-Run analysis, container flow times by port and channel are calculated as a function of traffic levels, thereby capturing the effects of capacity constraints.

In this article we use the sets of queuing models proposed by Leachman and Jula (submitted for publication-a,b) to estimate the containers flow times at ports and railroads based on traffic levels. Interested readers are encouraged to refer to that article for a survey on existing literature on applications of queuing models in estimating the container flow times.

For this research, we secured US customs data for year 2006 as summarized in the PIERS (http://www.piers.com/) commercial data subscription. These data specify for each US port, each importer, and each of 99 commodity codes the total volumes of imports from Asian origins (measured in 20-foot equivalent units, or TEUs). We also secured the customs data for year 2006 as summarized in the World Trade Atlas commercial data subscription, which summarizes total volumes of imports to the Continental USA from Asian origins by total declared value for each of the 99 commodity codes. These data enabled the authors to make estimates for volumes and declared values per cubic foot by commodity type. Here we assume a particular distribution of imports among 83 large, nation-wide importers and 19 sets of generic importers acting as proxies for small and regional importers. We have classified these importers based on their US customs declared value for imported goods, which is the value per cubic foot of container capacity and not the value per cubic foot of the actual cargo within the container. These values reflect the manufactured or purchased cost of the goods in Asia plus waterborne transportation rather than their full retail values in the United States.

The 11 major North American ports of entry considered in this study are as follows: Prince Rupert, BC; Vancouver, BC; Seattle-Tacoma, WA; Oakland, CA; Los Angeles–Long Beach, CA; Lazaro Cardenas, Mexico; Houston, TX; Savannah, GA;

Charleston, SC; Norfolk, VA; and Port of New York–New Jersey, NY. There are other ports handling Asian imports to North America, but in much smaller volumes than handled by the above ports.

In our study, the continental United States is divided into 21 regions, with the entire import demand for each region concentrated at a single location (the assumed site of the RDCs serving that region). The list of regions and RDCs are as follows (the location of RDCs are listed in parenthesis following each region): Atlanta Region (RDC in Duluth, GA); Baltimore Region (Frederick, MD); Boston Region (Milford, MA); Charleston Region (Summerville, SC); Charlotte Region (Salisbury, SC); Chicago Region (Joliet, IL); Cleveland Region (Chagrin Falls, PA); Columbus Region (Springfield, OH); Dallas Region (Midlothian, TX); Harrisburg Region (Allentown, PA); Houston Region (Baytown, TX); Kansas City Region (Lenexa, KS); Los Angeles Region (Ontario, CA); Memphis Region (Millington, TN); Minneapolis Region (Rosemount, MN); New York Region (RDCs assumed to be in East Brunswick, NJ and in Allentown, PA); Norfolk Region (Suffolk, VA); Oakland Region (Tracy, CA); Pittsburgh Region (Beaver Falls, PA); Savannah Region (Garden City, GA); Seattle Region (Fife, WA). The definition of the regions and the particular locations within each region are specified in Leachman (2008).

The geographical distribution of import volumes by destination is assumed to be the same for all importers. In this study, this distribution was set to be proportional to purchasing power in the regions (based on statistics available on the US Department of Commerce website), but other distributions could be input to the models. Year 2007 rate quotations to various importers from steamship lines, non-vessel-operating common carriers, intermodal marketing companies, trans-loading warehouse operators, railroad carriers and trucking companies were obtained to each RDC via each channel for which service was offered and quotes were secured. Considerable variation in rates from carrier to carrier and customer to customer was encountered. Average rates were developed from a basket of rates for each channel. There were major increases in rail rates charged to certain steamship lines beginning with the 2007 season, partially passed on as higher inland-point intermodal (IPI) rates for the 2007 season and more fully reflected in the 2008 rates. Such changes tend to promote market share for all-water steamship line service via the Panama Canal and diminish shares of West Coast ports. This affects the validation of the model, as will be discussed in the results section.

Transportation costs to importers for routing imports from Asia via eleven alternative North American ports of entry to the 21 RDC destinations were developed. For each port of entry and each destination, rates were developed for two alternative supply-chain channels: (1) shipping marine containers direct from Asia to RDC destinations, and (2) shipping marine containers to trans-loading warehouses in the hinterlands of the ports of entry, thence re-loading the imports in domestic rail containers or truck trailers for re-shipping from trans-loading warehouses to regional destinations. We have observed in practice that typically each RDC is supplied using only one channel. Volume is concentrated on a channel in order to negotiate a favorable rate as well as to simplify information management. We have therefore assumed in the models below that each RDC must be replenished using a single port and a single landside channel.

3. Methodology

In this section, first we provide a high-level overview of the proposed models and introduce their components and how they interact. Then, we focus on each of the components of models and provide explanations.

3.1. Structure and overview of models

The Long-Run Model is an extension of the model introduced by Jula and Leachman (2011). We provide a Mixed Integer Non-Linear Programming formulation of the supply-chain optimization from the importers' point of view, and provide heuristics to solve the model efficiently and effectively. The model is initialized with assumed or actual container flow time statistics by port and channel and with an assumed distribution of imports by declared value, importer and region. Calculations in the Long-Run are made to identify the best supply-chain strategy for each type of importer, identifying among strategies suitable for the importer the particular strategy that minimizes its total inventory and transportation costs. The resulting import volumes are then tallied by port and landside channel while considering optional minimum volume requirements for each of the various ports.

Minimum volume requirements may arise from two different issues. First, transportation rate quotes from steamship lines to large importers sometimes specify minimum required volumes or maximum allowed volumes at particular ports of entry. These contracts are re-negotiated annually. For an analysis of the reaction of import flows to infrastructure changes or new container fees within one year, such constraints could be important. However, our time horizon is multiple years. Large importers we spoke to indicated they have no trouble securing contracts with volume minimums and maximums appropriate for their business, as no single importer accounts for more than 10% of imports and there are many lines competing for their business. So we choose not to include such constraints in our model. Second, steamship lines enter into long-term contracts with ports that may feature minimum volume guarantees or cost penalties to the line for tendering volumes below specified thresholds. These constraints can be met by the lines by appropriately allocating so-called discretionary imports, i.e., imports moving under inland-point intermodal (IPI) rates, among the alternative ports of entry. (The lines have control of the port of entry for IPI imports, but not for imports moving in other channels.) While we do not analyze minimums line by line, we emulate the impact of the lines' allocation decisions in our model. We consider the total contractual volume at each port, and insure the total port volume is consistent with its minimum volume. If contractual commitments are not met by the aggregate volumes resulting from application of optimal strategies for the importers, then volumes of discretionary imports are adjusted to satisfy these commitments with least total increase in cost. Once all port constraints are satisfied, volumes by port and landside channel are tallied a final time and reported as output of the model. In the Long-Run Model, the mean and standard deviation of container flow times by channels are fixed, reflecting an assumption that over the long term the various ports and transportation carriers would make investments to maintain existing service quality and thereby protect market share. The details of the Long-Run Model are described in Section 3.2.

The proposed Short-Run Model incorporates two main modules: the Long-Run Model and the Queuing Model. Application of the model involves iterative calculations of the Long-Run Model and the Queuing Model, mimicking the actual behavior of importers and carriers. Fig. 1 shows the structure of the Short-Run Model and the interaction between its modules.

The Queuing Model is initialized with data concerning port terminals, rail intermodal terminals, and rail line-haul characteristics. Port terminal input data includes available acreage and crew shifts operated. Rail terminal input data includes terminal acreage, shifts operated per day, and port shares. Port shares are specified separately for direct inland movement of marine containers and for imports trans-loaded to domestic containers for the cases where ports are served by more than one rail terminal. The volumes calculated by the Long-Run Model may be fed into the Queuing Model, or user-defined volumes may be tendered to it. The Queuing Model calculates updates to the container flow times, considering congestion or lack thereof at the various ports or in the various channels. The results of the Queuing Model are summarized in a format suitable for input to the Long-Run Model in another iteration of that model.

The Long-Run Model is re-applied after application of the Queuing Model to re-optimize supply-chain strategies for each type of importer. Again, volumes are tallied by port and channel, and compared to contractual minimums, again adjusting discretionary volumes as required. If there is significant change in volumes by port or channel, the adjusted volumes are fed back to the Queuing Model for re-calculation of flow times. This iteration continues until either (1) flow times and volumes by port and channel stabilize, or (2) a pre-specified maximum number of iterations is performed. Case 2 occurs if the model cycles between two different allocations of imports to ports and channels. As discussed below, the model has been engineered so that cycling, if it occurs, is between two import distributions with small differences. These differences are at the “noise level” of the model, reflecting alternative, equally-efficient import strategies for the importers.

The iteration of the Long-Run Model and Queuing Model reflects the following basic phenomenon: Changes in flow times result in changes in inventory costs for importers. These changes induce changes in supply-chain strategies. This mimics real-life, iterative behavior of importers, transportation providers, and ports. For example, in the summer of 2004, the San Pedro Bay ports experienced severe congestion. Container flow times through the ports increased sharply. Because of contractual commitments and other operational impediments, reactions of importers and steamship lines were impeded and delayed. For the 2005 season, several steamship lines changed certain of their Asia – North America vessel strings that called at San Pedro Bay ports to call at Puget Sound ports instead. Thus during the 2005 season, import share through the Puget Sound ports was up sharply, while it was down at the San Pedro Bay ports. During the 2005 season, the “PierPass” program of congestion pricing and night-time operational shifts at port terminals was successfully implemented at San Pedro Bay, and container flow times experienced at the San Pedro Bay ports in the 2005 season were as short as, or shorter than, they were in 2003. But, on the other hand, container flow times through the Puget Sound ports and over railroads serving those ports increased during 2005. So for the 2006 season, the vessel strings that moved up to Puget Sound for the 2005 season were moved back to San Pedro Bay. Note how the lines and importers over-reacted to the 2004 congestion. It was not until the start of the 2006 season that a new equilibrium was reached. Without intervention, the same sort of over-reaction can happen in the iterations between the Long-Run Model and Queuing Model. Without some sort of feedback control, the iterative model may not reach a steady-state and could cycle indefinitely between two different allocations.

For this reason, a proportional controller is incorporated into the model. Once new container flow times are calculated by the Queuing Model, instead of jumping to the new flow times, a *proportional* correction is made to the old flow times to

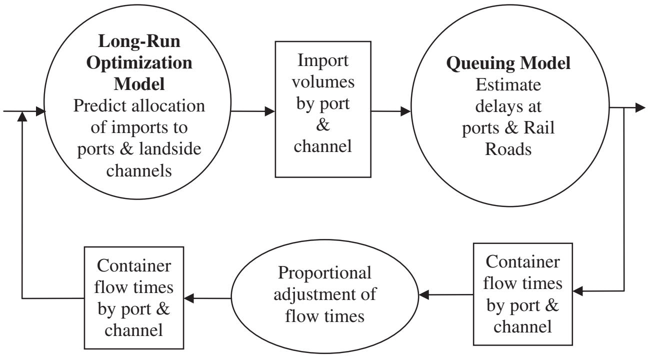

Fig. 1. The structure of the Short-Run Model and its components.

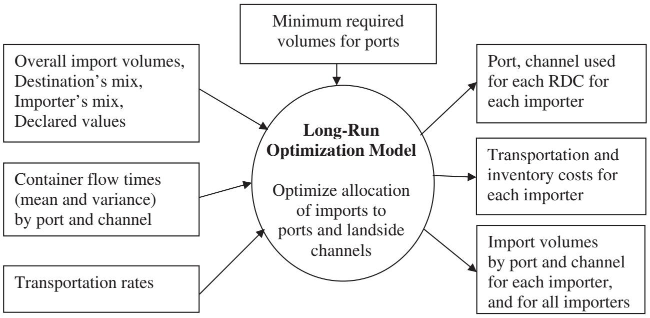

Fig. 2. Inputs & outputs of the Long-Run Model.

define input for the next run of the Long-Run Model. See Fig. 1. This correction takes the form of a weighted average of new and old flow times. This proportional correction is set so as to induce a gradual and conservative change in overall import volumes in response to perceived opportunities to avoid congestion and reduce transit time (and hence inventory expense). In subsequent iterations, if opportunity still exists, further import volume is shifted. This cautious approach enables convergence of the model to a stable solution after a reasonable number of iterations and avoids cycling instabilities.

3.2. Long-Run model

Here we propose a mixed integer non-linear programming model which optimizes the distribution of import volumes by port and landside channel for a given port and transportation infrastructure network. This model helps importers to decide what is their best allocation of Asian imports by port and channel so as to minimize their total costs for transportation and handling, pipeline inventory, and safety-stock inventory at RDCs. Here we assume all demands at each RDC must be served only by one port using one mode of transportation. Fig. 2 displays a schematic of the Long-Run optimization model and the required inputs and generated outputs.

3.2.1. Mixed integer non-linear programming (MINLP) model for all importers

Notation for parameters

g	index of importers;
n	index of set of RDCs ($n = 1, 2, \dots, N$);
m	index of set of ports of entry (POEs);
i	index of set of land transportation modes (channels);
D_g	nation-wide average sales volume for the importer g per week (expressed in physical units);
$MAPE$	mean absolute percentage error (expressed as a fraction of one) in one-week-ahead forecasts of nation-wide sales for the importer (assumed to be the same for all importers);
σ_g^D	standard deviation of errors in one-week-ahead forecasts of nation-wide sales. A standard assumption is $\sigma_g^D = (1.25)(MAPE)(D_g)$ (see, e.g., Silver and Peterson, 1985);
$D_{g,n}$, $\sigma_{g,n}$	mean (D), and the standard deviation (σ) of sales for importer g distributed from RDC n. It is assumed that $\sum_n D_{g,n} = D_g$, and the proportion of nation-wide sales handled by each RDC is fixed;
R	time between replenishment orders (from Asian suppliers). R is assumed to be 1 week for all importers;
L^A	mean value of the lead time (expressed in weeks) from when a nation-wide replenishment order is placed until an allocation of the order among USA ports of entry is fixed and vessel passages are booked; (in the case of direct shipping, the RDC destination also is chosen at the end of this lead time, whereas in the case of consolidation–de-consolidation, only the ports of entry are selected);
L_m^M, σ_m^M	mean value (L), and the standard deviation (σ) of the lead time (expressed in weeks) for shipments from point of origin to port of entry POE m, measured from when vessel passage is booked until land transport to RDC from POE m begins (direct shipping case), or until land transportation to destination RDC from POE m is booked (consolidation-de-consolidation case);
L_m^S, C_m^S	mean value (L) of the lead time (expressed in weeks), and transportation cost (C) per unit of load, from departure from point of origin until land transport from POE m to RDC begins (direct shipping), or until land transport from POE m (consolidation–deconsolidation) to destination RDC is booked;

$L^N_{m,n,i}, \sigma^N_{m,n,i}, C^N_{m,n,i}$	mean value (L) and standard deviation (σ) of transportation lead times (expressed in weeks), and transportation cost (C) per unit of load, shipped using land transportation mode *i*, from departure from POE *m* until processed through RDC *n* (direct shipping); or from when land transport from POE *m* to RDC *n* is booked until processed through the RDC *n* (consolidation–deconsolidation);
z_g	safety factor for importer *g*, determining the level of safety stocks at RDCs. (Choosing $z = 2.05$ implies approximately a 98% probability of no stock-out.) For each importer, a homogenous customer satisfaction level is assumed for all RDCs.
V^S_g	the amount of capital tied up in a unit of pipeline stock from origin to POE for importer *g*;
V^N_g	the amount of capital tied up in a unit of pipeline stock from POE to RDC for importer *g*;
V^R_g	the amount of capital tied up in a unit of RDC safety stock for importer *g* (assumed to be the same for all RDCs for each importer in this article);
r_g	inventory carrying rate (inventory holding cost rate) for importer *g*;
M_m	Minimum contractual import volume that must be routed through port *m*.
Variables	
$\delta_{g,m,n,i}$	binary variable (0 or 1) indicating if land transportation mode *i* is used for transportation from departure from POE *m* to RDC *n* for importer *g*. This variable is set to zero if land transportation mode *i* can not be used for transportation from *m* to *n*;
$\delta_{g,m,n}$	binary variable (0 or 1) indicating if RDC *n* is served by port *m* for importer *g*;
$\delta_{g,m}$	binary variable (0 or 1) indicating if port *m* is used by the importer *g*;
$L^N_{g,m,n},\ \sigma^N_{g,m,n}$	mean value (L) and standard deviation (σ) of the lead times (expressed in weeks) for importer *g*, using selected land transportation from departure from POE *m* until processed through RDC *n* (direct shipping); or from when land transport from POE *m* to RDC *n* is booked until processed into RDC *n* (consolidation–deconsolidation);
$\Re_g$	set of RDCs served by direct shipment for importer *g*; $\Re_g = \{n \mid \delta_{g,m,n,i} > 0, i \in Direct\}$;
$\Re^C_g$	set of RDCs served by trans-loading for importer *g*; $\Re^C_g = \{n \mid \delta_{g,m,n,i} > 0, i \in Transloading\}$;
$\Omega_{g,m}$	set of RDCs served using port *m* for importer *g*; $\Omega_{g,m} = \{n \mid \delta_{g,m,n} > 0\}$;
Φ_g	set of ports used by the importer *g*; $\Phi_g = \{m \mid \delta_{g,m} > 0\}$;
ss_g	positive continuous variable showing the total safety stock in the chain for importer *g*;

Constraints

$$\sum_{m,i} \delta_{g,m,n,i} = 1 \quad \forall g, n \tag{1}$$

$$L^N_{g,m,n} = \sum_i \delta_{g,m,n,i} L^N_{m,n,i} \quad \forall g, m, n \tag{2}$$

$$\sigma^N_{g,m,n} = \sum_i \delta_{g,m,n,i} \sigma^N_{m,n,i} \quad \forall g, m, n \tag{3}$$

$$\delta_{g,m,n} = \sum_i \delta_{g,m,n,i} \quad \forall g, m, n \tag{4}$$

$$\sum_n \sum_i \delta_{g,m,n,i} \leqslant N \delta_{g,m} \quad \forall g, m \tag{5}$$

$$\sum_g \sum_n \delta_{g,m,n} D_{g,n} \geqslant M_m \quad \forall m \tag{6}$$

Constraint (1) guarantees that each RDC for each importer is served, and it is served only by one port and one mode of transportation. The mean and standard deviation of the lead times from ports to RDCs are calculated in (2) and (3) for each importer. Constraint (4) sets the port – RDC combinations used by each importer. Constraint (5) identifies which ports are used by each importer. Finally, constraint (6) guarantees that minimum contractual volumes are met for each port.

Objective function

Our objective is to minimize the total cost (i.e. transportation cost plus inventory holding cost). The cost of the cycle stock has been omitted because that cost is independent of the supply chain channel alternatives. Eq. (7) shows the total transportation cost.

$$\sum_g \sum_m \sum_n \sum_i (\delta_{g,m,n,i}(C^S_m + C^N_{m,n,i}) D_{g,n}) \tag{7}$$

The inventory holding cost is due to the in-transit inventory cost and the required safety-stocks, and is expressed in Eq. (8).

$$\sum_g \sum_m \sum_n \delta_{g,m,n} r_g V^S_g L^S_m D_{g,n} + \sum_g \sum_m \sum_n \delta_{g,m,n} r_g V^N_g L^N_{g,m,n} D_{g,n} + \sum_g r_g V^R_g ss_g \tag{8}$$

The first two terms of Eq. (8) show the in-transit inventory cost from Asia to POEs, and from POEs to RDCs, respectively. The last term shows the cost of the safety stock.

Jula and Leachman (2011) developed an expression for safety stock for a single importer assuming a proportional fractile allocation policy with independent and identically distributed normal variables for demands and lead-times, and no correlation among variables. Ports do not hold inventories, and all inventories are at RDCs. Assuming non-collaborative importers, in this article we extend the authors' safety stock formula to encompass the safety stocks of all importers (i.e. the added index g represents the safety stock for importer g). Eq. (9) shows the formula for the mixed strategies of direct shipment to some RDCs and consolidation–de-consolidation shipments to others. This equation also accommodates unequal channel lead times and unequal RDC demands.

$$
\begin{aligned}
ss_g = {} & z_g \sum_{n \in \Re_g} \sum_m \delta_{g,m,n} \left[L^A \sigma_{g,n}^2 \frac{\sum_v \sigma_{g,v}^2}{\left(\sum_v \sigma_{g,v}\right)^2} + \sigma_{g,n}^2 \left(L_m^M + L_{g,m,n}^N + R \right) + D_{g,n}^2 \left(\left(\sigma_m^M\right)^2 + \left(\sigma_{g,m,n}^N\right)^2 \right) \right]^{1/2} \\
& + z_g \sum_{n \in \Re_g^C} \sum_m \delta_{g,m,n} \left[L^A \sigma_{g,n}^2 \frac{\sum_v \sigma_{g,v}^2}{\left(\sum_v \sigma_{g,v}\right)^2} + \sigma_{g,n}^2 L_m^M \frac{\sum_{v \in \Re_g^C} \delta_{g,m,v} \sigma_{g,v}^2}{\left(\sum_{v \in \Re_g^C} \delta_{g,m,v} \sigma_{g,v}\right)^2} + \sigma_{g,n}^2 \left(L_{g,m,n}^N + R \right) \right. \\
& \left. + D_{g,n}^2 \left(\frac{\sum_{v \in \Re_g^C} \delta_{g,m,v} D_{g,v}^2}{\left(\sum_{v \in \Re_g^C} \delta_{g,m,v} D_{g,v}\right)^2} \left(\sigma_m^M\right)^2 + \left(\sigma_{g,m,n}^N\right)^2 \right) \right]^{1/2}
\end{aligned} \tag{9}
$$

where v is an additional index notation for RDCs. The first term of Eq. (9) shows the safety stock in RDCs served by direct shipment for each importer. The second term shows the safety stock in consolidation–de-consolidation cases.

3.2.2. Proposed solution

In an oversimplified form, the structure of the problem under study can be viewed as a p-median problem, in which p facilities are to be selected to minimize the total (weighted) distances or costs for supplying customer demands. In addition, the problem includes more complexities such as the inventory costs, which are non-linear in the assignment variables, and the selection of transportation modes in a multi-echelon setting. Thus the problem we are studying is more difficult than the standard p-median problem, which is already a notorious NP-hard problem (e.g. see Revelle et al., 2008).

Since the individual importers are not generally collaborative with each other (e.g. their goods are not shared to satisfy each others demands), our main heuristic decomposes the problem first to individual importers. The heuristic then selects among alternative location-level strategies (i.e., direct shipping vs. use of trans-load warehouses) for each importer, and optimizes the choice of modes and allocation of volumes to routes for each location-level strategy, and finally identifies the best of the alternatives considered for each of the importers. As reported by Jula and Leachman (2011), this approach will result in optimal or close to optimal solution for each importer in this environment. However, when tallying the loads going through the ports, there is no guarantee that the minimum port constraints (6) are satisfied. Therefore, here we introduce a complementary heuristic that adjusts the volumes of discretionary imports to satisfy these constraints with least total increase in cost.

3.2.2.1. Main heuristic.

Following the suggestion of Jula and Leachman (2011), we allow the following strategies for each importer (we later adjust the strategies in subsequent heuristics, if needed):

(1) *TL-LA*: Consolidate–deconsolidate and trans-load using a warehouse at Los Angeles–Long Beach only,
(2) *TL-2*: Consolidate–deconsolidate and trans-load using warehouses at 2 specific ports,
(3) *TL-4*: Consolidate–deconsolidate and trans-load using warehouses at 4 specific ports,
(4) *TL-5*: Consolidate–deconsolidate and trans-load using warehouses at 5 specific ports,
(5) *Direct-WC*: Direct-ship marine box to RDCs considering use of only West Coast ports,
(6) *Direct-All*: Direct-ship marine box to RDCs considering use of all ports.

The strategies are summarized in Table 1 in terms of candidate ports and in terms of whether or not trans-loading is utilized. Here is the detail of our proposed heuristic:

```
Step 1. for each importer $g$, do
Step 2. for each strategy $s$ selected from the set of strategies specified in Table 1, do
Step 3. for each $n$ in the set of RDCs, do
Step 4. for each port $m$ in $P_s$ (set of ports in strategy $s$), do
Step 5. for each land transportation mode $i$ used in strategy$s$, do
Step 6. Calculate transportation cost using $(C_m^S D_{g,n} + C_{m,n,i}^N D_{g,n})$
```

```
Step 7. Calculate in-transit inventory cost using $(r_g V_g^S L_m^S D_{g,n} + r_g V_g^N L_{m,n,i}^N D_{g,n})$, and calculate $C_{g,m,n,i}$, the sum of transportation cost and in-transit inventory cost for importer $g$.
Step 8. end for
Step 9. end for
Step 10. Select port $m_0$ and land transport mode $i_0$ such that the sum of transportation cost and in-transit inventory cost is minimized for the selected RDC $n$, i.e. $(C_{g,m_0,n,i_0})$ is minimum.
Step 11. Set $\delta_{g,m,n,i} = 1$ for $m = m_0$, $i = i_0$; and Set $\delta_{g,m,n,i} = 0$ for all other $m$, and $i$,
Step 12. end for
Step 13. Set $\Re_g = \{\}$ for trans-loading strategies (i.e. TL-LA, TL-2, TL-4, TL-5); and set $\Re_g^C = \{\}$ for direct shipment strategies (i.e. Direct-All, Direct-WC);
Step 14. For importer $g$, derive $L^N_{g,m,n}, \sigma^N_{g,m,n}$, $\delta_{g,m,n}$, and $\delta_{g,m}$ using Eqs. (2)–(5), and derive $\Omega_{g,m} = \{n | \delta_{g,m,n} > 0\}$, and $\Phi_g = \{m | \delta_{g,m} > 0\}$;
Step 15. For importer $g$, calculate the total safety-stock ($ss_g$) using Eq. (9),
Step 16. For importer $g$, calculate:
    – transportation cost = $\sum_m \sum_n \sum_i (\delta_{g,m,n,i}(C_m^S + C_{m,n,i}^N) D_{g,n})$
    – in-transit inventory cost = $\sum_m \sum_n \delta_{g,m,n} r_g V_g^S L_m^S D_{g,n} + \sum_m \sum_n \delta_{g,m,n} r_g V_g^N L_{g,m,n}^N D_{g,n}$
    – safety stock cost = $r_g V_g^R ss_g$
    – total cost = transportation cost + in-transit inventory cost + safety stock cost
Step 17. end for
Step 18. Select the best strategy which minimizes the total cost for the importer $g$,
Step 19. For the best strategy for the importer $g$, report the total cost, and the ports and channels used for each RDC.
Step 20. end for
Step 21. Tally and report the total flows through each port and each channel considering all importers.
```

The above heuristic is designed to satisfy all the MINLP constraints except constraint (6), the minimum required contractual agreement for each port. If constraint (6) is satisfied, the heuristics stops. If not, the following complementary heuristic is applied to the output of the main heuristic.

3.2.2.2. Satisfying port minimum heuristic. Let Π be the set of *violated ports* in which $\sum_{g,n} \delta_{g,m,n} D_{g,n} < M_m$. We define $S_m = M_m - \sum_{g,n} \delta_{g,m,n} D_{g,n}$ to be the slack (shortage) for a violated port m. Π^C is the set of *non-violated* ports, in which $\sum_{g,n} \delta_{g,m,n} D_{g,n} \geqslant M_m$. For non-violated ports, we define $E_m = \sum_{g,n} \delta_{g,m,n} D_{g,n} - M_m$, to be the excess load at port m. The heuristic strives to move volume to violated ports from qualified non-violated ports with the least additional cost. The heuristic iteratively moves loads by going through an ordered set of direct-shipment importers, listed from low-value to high-values goods importers, i.e. Generic 2, Generic 6, Generic 10, etc., corresponding to the value of importers good, \$2/cu-feet, \$6/cu-feet, \$10/cu-feet, respectively. This mimics actual practice, because these importers are utilizing IPI services, for which the steamship line can control the port of entry. Steamship lines routinely balance their ports of entry for IPI traffic in order to satisfy their contractual minimum volumes at ports. By shifting the lowest-value commodities first, we minimize the increment in inventory expenses associated with shifting volume off its most preferred route. Here are the details of the heuristic:

```
Step 1. for each port $m$ in $\Pi$, do
Step 2. for each importer $g$ in the ordered set of general direct small importers {Generic 2, Generic 6, ...}, do
Step 3. for each port $u$ in $\Pi^C$, do
Step 4. for each RDC $n$ in $(\Omega_{g,u})$, do
Step 5. for each $i$ in the set of land-side transportation modes, do
```

Table 1
Alternative strategies considered by the main heuristic.

Strategy Consolidation	TL-LA YES	TL-2 YES	TL-4 YES	TL-5 YES	Direct-WC NO	Direct-All NO
Possible Ports	LA–LB	LA–LB NY–NJ	LA–LB NY–NJ Seattle-Tacoma Savannah	LA–LB NY–NJ Seattle-Tacoma Savannah Houston	LA–LB Oakland Seattle-Tacoma Vancouver Prince Rupert	LA–LB Oakland Seattle-Tacoma Vancouver Prince Rupert NY–NJ Houston Savannah Charleston Norfolk Lazaro-Cardenas

```
Step 6. Calculate $C_{g,m,n,i}$, transportation cost plus in-transit inventory cost,
        $C_{g,m,n,i} = (C^S_m D_{g,n} + C^N_{m,n,i} D_{g,n}) + (r_g V^S_g L^S_m D_{g,n} + r_g V^N_g L^N_{m,n,i} D_{g,n})$
Step 7. end for
Step 8. select transportation mode $i \doteq i_0$, such that $C_{g,m,n} = \min_i \{C_{g,m,n,i}\}$
Step 9. calculate the differential cost, $d_{g,m,u,n} = C_{g,m,n} - C_{g,u,n}$
Step 10. end for
Step 11. end for
Step 12. find the best port ($u$) and RDC ($n$), $n \in \Omega_{g,u}$, $u \in \Pi^C$, such that $d_{g,m,u,n}$ is minimized, while $E_u \geq D_{g,n}$. If there exist
         such $n = n_0,\ u = m_0$, we move the load from port $m_0$ to $m$ as follows:
    – set $\delta_{g,m_0,n_0,i} = 0, \forall i$
    – set $\delta_{g,m,n_0,i_0} = 1$
    – set $E_{m_0} = E_{m_0} - D_{g,n_0}$
    – set $S_m = S_m - D_{g,n_0}$
    – if $S_m \leq 0$, then move $m$ from set of violated ports to non-violated ports,
      (i.e.$\Pi = \Pi \setminus \{m\}, \Pi^C = \Pi^C U\{m\}$); if the new $\Pi$ is empty then exit the heuristic;
Step 13. end for
Step 14. end for
Step 15. Update the total cost, and the ports and channels used for each RDC for each importer.
Step 16. Report the total flows going through each port and each channel for all importers.
```

3.3. Queuing model

Analytical queuing formulas were developed for estimating import container flow times through port terminals, rail intermodal terminals and rail line-haul channels as a function of traffic volumes, infrastructure and staffing. The queuing-theoretic formulas express waiting time as a non-linear function of utilization and the number of parallel servers. As utilization is increased, waiting time increases exponentially. For a fixed utilization, the waiting time can be mitigated by increasing the number of parallel servers.

The queuing formulas developed for each of the three types of applications (port terminals, rail terminals, rail line hauls) were statistically fitted to 2006 industry data to provide models of container flow time as a function of parameters for traffic volume, infrastructure (e.g., terminal acreage, number of rail main tracks), staffing, and hours of operation. The formulas are applied separately to each port terminal, each rail terminal, and each segment of the rail line-haul network, from which container flow times are aggregated. The analyst may employ these formulas to calculate predictions of changes in container flow time as a function of changes in the parameters. Fig. 3 provides a schematic of the inputs and outputs of the Queuing Model.

There are actually three sets of queuing calculations that are performed in this model: (1) queuing analysis at port terminals to determine container dwell times from vessel arrival to container departure out the truck gate or on a double-stack train, (2) queuing analysis at rail terminals to determine container dwell times from container arrival until train departure, and (3) queuing analysis of rail lines to determine transit times from origin terminal to destination terminal for rail intermodal trains handling imports. The mathematical formulas of the Queuing Model are beyond the scope of this paper. Interested readers are encouraged to refer to Leachman and Jula (submitted for publication-a,b) for the complete details of the Queuing Model. The results of the Queuing Model are summarized in a format suitable for input to the Long-Run Model in another iteration of that model.

4. Analysis and results

As discussed in previous sections, many stakeholders may benefit from analyses performed using the proposed models. For example, steamship lines and railroads could study the impact of changes in their rates. Large importers could assess the merits of investing in import warehouses and trans-loading/de-consolidation services at various ports of entry.

One application of proposed models concerns the support of public policy. Given the pollution and traffic congestion associated with high levels of imports, there is interest in assessing fees on import containers to defray costs of environmental mitigation (e.g., clean trucks) and traffic mitigation (e.g., capacity expansion of road and rail infrastructure). Here, we use the proposed models to analyze the effect of imposing hypothetical range of container fees at the Los Angeles–Long Beach ports.

As imports move through the supply chain, they accumulate more cost. One index to the amount of capital tied up can be the value declared to US customs. Based on our analysis of data and for the purposes of this study, the assumption was made that the landside pipeline inventories are valued by importers at 125% of the value declared to customs. It was further assumed that RDC inventories are valued at 150% of the value declared to customs.

The appropriate inventory carrying rate to apply depends on a number of factors. In the case of replenishment of goods with long-term demands, we use 20% per year for the inventory carrying rate. A higher rate is more appropriate if retail prices are declining with time, or if the products experience rapid obsolescence, such as is the case for technology goods,

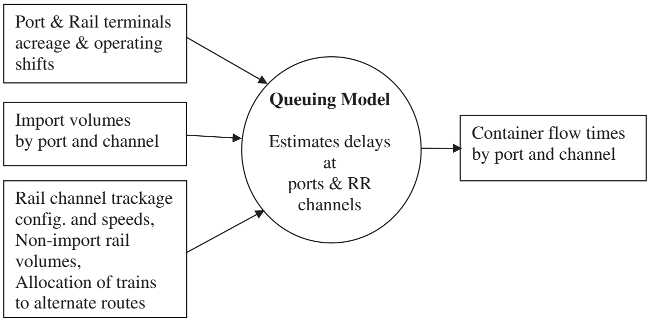

Fig. 3. Inputs and outputs of Queuing Model.

style goods and goods for special sales events. For such cases, a rate of 50% per year is considered more appropriate. The sales of most retailers are a mixture of event items and standard items. For simplicity a simple average of the two cases was assumed, i.e., an inventory carrying rate of 35% per year is assumed for the purposes of costing pipeline inventories and safety stocks. We assume a 98% target for the customer satisfaction level (i.e., fraction of demand met without stockout) for all importers. Unless otherwise specified, we use the proportional correction factor of 20% for the iteration of our Short-Run model, i.e. the weight on the newly-calculated flow times from Queue Model is set to be 20%, with 80% weight remaining on the old flow times.

We used a laptop computer with a 2-GHz CPU, and Visual Basic programming language to run our experiments. Our study shows that the proposed heuristics are very efficient in terms of CPU run-time. Long-run analysis takes around 1.5 min to generate results for a complete set of importers. The required time for Short-Run analysis depends on the number of specified iterations. For a typical 10 iterations experiment, our analysis takes around 15 min to generate results.

4.1. Models validation

Combining data from multiple sources,[1] the following break-down of 2006 containerized imports through the San Pedro Bay ports was estimated: 21% was “local” traffic, i.e., imports ultimately consumed in Southern California, Southern Nevada, Arizona, New Mexico, Southern Utah or Southern Colorado; 43% was kept in the marine box and placed on a double-stack train destined east of the Rockies (i.e. inland-point-intermodal or “IPI” volume); and the remaining 36% was either (a) unloaded from marine boxes in the local region at a warehouse or trans-loading facility, re-loaded in domestic vehicles (truck or rail) and re-shipped for consumption outside the local region, or (b) kept in a marine box that was trucked outside the above-defined “local” region. The (b) part of the 36% category is believed to be very small. Thus the amounts of traffic in IPI and trans-loading categories at San Pedro Bay are roughly equal, and each is about double the local traffic. For the West Coast as a whole, “local” traffic (i.e., imports ultimately consumed in West Coast regions or regions for which the West Coast ports are the closest ports) was about 30% in 2006; IPI traffic was about 46%; and trans-loading/long-distance trucking was about 24%.

Table 2 compares actual 2006 data to results of Short-Run Model and Long-Run Model. All statistics match fairly well. The solutions to the Models call for somewhat greater all-water share and less West Coast IPI share than was experienced in 2006, but the solutions match the actual 2007 and 2008 IPI shares fairly well. In the authors’ opinion, this is to be expected, given the assumption of 2007 season rates in the models.

Considering the extraordinary challenge posed by an attempt to precisely match reality with a simplified, nationwide model, in the authors’ opinion, the results indicate the models are satisfactory for studying shifts in port and modal shares in response to hypothetical container fees.

4.2. The effect of imposing container fees at San Pedro Bay ports

Short-Run and Long-Run calculations testing the imposition of hypothetical container fees at San Pedro Bay were made for various scenarios. Container fees in increments of $50 per 40-foot equivalent unit (FEU) were tested. The results of Short-Run analysis are summarized in Table 3. The last column of this table conveys the assumed port minimum import volumes.

The Short-Run Model run with a $0 per FEU fee stabilized to a solution after eight iterations (although ten iterations were run). In this case, about 28% of total Asia – Continental US imports fall under consolidation–de-consolidation strategies, the other 72% under direct shipping strategies. The San Pedro Bay ports account for about 51% of total waterborne imports to the continental US. Increasing the fee by $50 does not affect the relative mix of imports trans-loaded vs. direct, but it reduces the total import volume through the San Pedro Bay ports by about 4%. Most of the volume diverted was discretionary,

---

[1] Total import volume at San Pedro Bay and other West Coast ports was obtained from the Pacific Maritime Association. IPI volumes from San Pedro Bay and other West Coast ports were obtained from the Intermodal Association of North America and the ports themselves. Locally-consumed imports were estimated in proportion to purchasing power statistics obtained from the US Dept. of Commerce.

Table 2
Comparison of 2006 actual and model-predicted traffic shares.

	2006 Actual (%)	Solution to Short-Run Model (%)	Solution to Long-Run Model (%)
Port shares			
LA–LB	55.0	52.4	52.0
Other West Coast	20.0	19.0	19.0
All-water	25.0	28.5	29.0
LA-LB mix			
Regional	21.0	22.5	22.7
IPI	43.0	42.0	41.1
Trans-load	36.0	35.5	36.2
USA west coast mix			
Regional		27.9	28.1
IPI	46.0	41.9	41.2
Trans-load		30.2	30.7

inland-point intermodal (IPI) volume, mostly shifted to Seattle-Tacoma, with increases also showing up at Oakland, on the East Coast, and at Houston.

Imposing a $100/FEU will cause the total import volume through the San Pedro Bay ports declined by another 6%, or 10% compared with no fee. Again, most of this volume shifted to Seattle-Tacoma, with increases also showing up at Oakland, and, to a lesser extent, at Houston and on the East Coast. As before, the relative mix of imports trans-loaded vs. direct shipped was unchanged; the amounts of both types of imports routed via San Pedro Bay declined in the face of fees.

As indicated before, we have applied 10 iterations for the Short-Run experiments. While increasing the number of iterations will slow down the system, imposing less number of iterations may prevent the system to reach a steady-state. Based on our observations, 10 iterations with proportional factor of 20% was adequate for the system to reach to a steady-state in an acceptable time. In case of no proportional feed-back, Fig. 4 shows an example of the fluctuations of loads on west coast ports for imposing $150 container fee at San Pedro Bay ports. It appears that the model may not reach to a steady-state in this case.

A proportional correction feedback loop induces a gradual and conservative change in overall import volumes in response to perceived opportunities to avoid congestion and save transit time (and hence inventory expense). In subsequent iterations, if opportunity still exists, further import volume is shifted. This cautious approach enables convergence of the model to a stable solution after a reasonable number of iterations and avoids cycling instabilities. Fig. 5 shows the effect of 20% proportional factor in stabilizing the system.

The same fee increases were tested in the Long-Run Model. According to Long-Run calculations, for a $50 per FEU fee, total volume through the San Pedro Bay ports declines by about 17%. For a $100 per FEU fee, total volume through the San Pedro Bay ports declines by 23% (compared to the base case with no fee). These results, along with results for higher fee values, are summarized in Table 4.

The Long-Run elasticity of imports via San Pedro Bay is roughly double the Short-Run elasticity. If a container fee is imposed, most of the volume leaving the San Pedro Bay ports would be diverted to the Puget Sound ports. In the Short-Run analysis, the next largest diversion of volume is to Houston and the East Coast (so-called all-water channels). A slightly smaller amount is diverted to Oakland. In the Long-Run analysis, the diversion to all-water channels is roughly half the diversion to Puget Sound, and the diversion to Oakland is in turn about half the diversion to all-water channels.

It is instructive to examine the elasticities of separate components of the overall import volume at San Pedro Bay. This is shown in Fig. 6. Imports consumed in the greater local region served by San Pedro Bay are insensitive to fees of $200 per TEU or less; trucking them from the Port of Oakland is much more expensive. High-value goods (with declared values greater than $28 per cubic foot) managed in trans-loaded supply chains also are insensitive to fees of $200 per TEU, because the savings in nationwide safety stock costs afforded by consolidating all import volume at San Pedro Bay are greater than this and, because Southern California is the largest local market, the increment in average transportation costs from performing nationwide consolidation anywhere else is greater than $200 per FEU. Moderate-value trans-loaded goods are more elastic; such goods are managed using strategies such as TL-4, and the amount assigned to each of the four corners can be unbalanced to reduce exposure to the San Pedro Bay fee. The most elastic imports are those handled in IPI supply chains. Cost differences for other ports of entry, especially the Puget Sound ports, are relatively minor, and competitive services are available via such ports. In the short run, the amount of the moderate-value trans-loaded, and IPI imports diverted are limited by the available infrastructure and potential congestion at other ports and in rail terminals and rail lines serving those ports. The Long-Run results show how much more diversion would occur if infrastructure capacities in the other channels were expanded sufficiently to maintain current container flow times at the higher volume levels.

Our analyses make clear that the discretionary inland-point intermodal (IPI) volumes are very elastic and decline rapidly with growing fee values. Trans-loaded imports of moderate declared value are somewhat less elastic, while trans-loaded imports of high declared value and imports consumed within the region are very inelastic.

Table 3
Import volumes vs. San Pedro Bay container fee, as predicted by the Short-Run Model (numbers are expressed in TEUs per day).

Fee value	$0/FEU	$50/FEU	$100/FEU	$150/FEU	$200/FEU	Minimum
LA-Long Beach	20,962	20,039	18,950	18,026	16,967	10,000
Seattle-Tacoma	4429	5162	5333	5495	6193	1000
Oakland	1760	1760	2375	2714	2756	1000
Vancouver	540	540	514	537	574	500
Prince Rupert	655	655	655	655	655	500
LC-Manzanillo	224	224	224	224	224	200
Houston	1309	1568	1782	1783	1850	500
East coast	10,091	10,023	10,139	10,536	10,751	4000

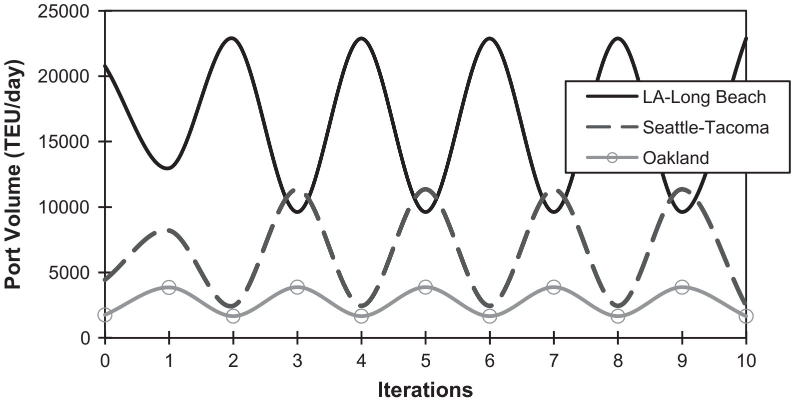

Fig. 4. The fluctuations of loads in the case of imposing $150 container fee with a simple loop.

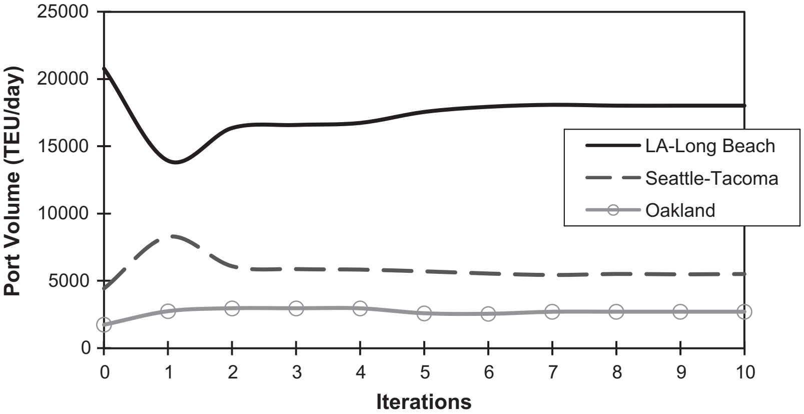

Fig. 5. The effect of 20% proportional factor feedback loop in stabilizing the system in the case of imposing $150 container fee.

Table 4
Import volumes vs. San Pedro Bay container fee, as predicted by Long-Run Model (numbers are expressed in TEUs per day).

Fee Value	$0/FEU	$50/FEU	$100/FEU	$150/FEU	$200/FEU	Minimum
LA–Long Beach	20,777	17,188	16,091	14,097	11,802	10,000
Seattle-Tacoma	4430	7346	7717	8383	9706	1000
Oakland	1755	1760	2013	2542	2954	1000
Vancouver	540	503	510	503	508	500
Prince Rupert	655	655	655	655	655	500
LC-Manzanillo	224	224	224	224	224	200
Houston	1499	1882	1960	2337	2505	500
East Coast	10,091	10,413	10,800	11,231	11,618	4000

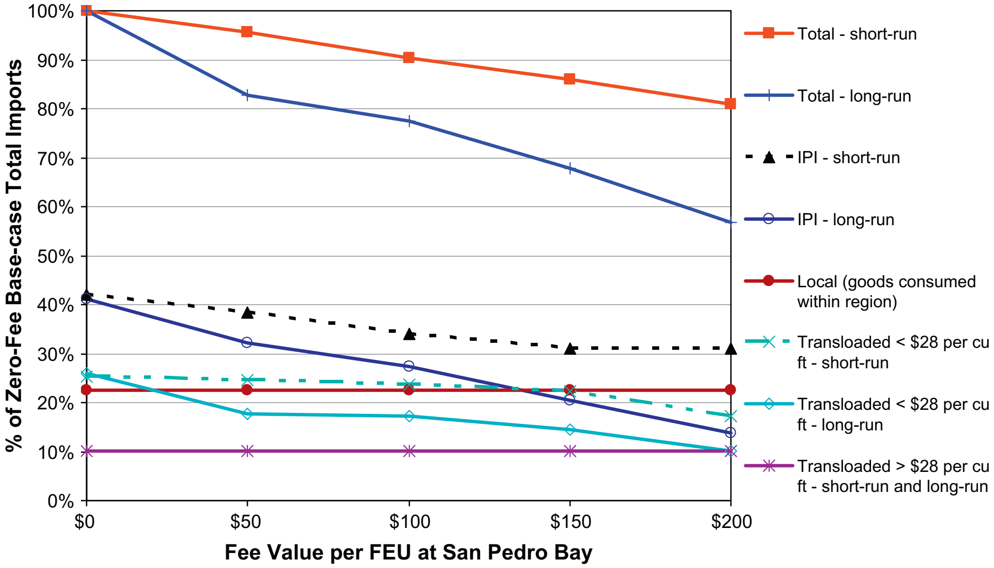

Fig. 6. Comparative Short-Run and Long-Run elasticities of direct, trans-loaded and local imports via San Pedro Bay.

Comparing Short-Run and Long-Run results, it is clear that adequate infrastructure, and staffing of that infrastructure, are not yet in place at other ports to accommodate without congestion the diversion of trans-loaded volumes away from San Pedro Bay ports. However, the economics encouraging expansion at other ports and their landside channels arise when fees greater than $150 per FEU are imposed on imports through the San Pedro Bay ports.

Many scenarios and fee values can be tested using the Short-Run and Long-Run Models in this article. Using the proposed models, Leachman (2010) provides elasticity curves for fee values at San Pedro Bay ranging up to $500 per FEU for scenarios involving +/− 10% changes in all-water steamship line rates relative to their rates via West Coast ports and for scenarios involving up to 10% market share gains by large, nation-wide importers able to practice consolidation–de-consolidation supply-chain strategies.

5. Summary and conclusions

In this article we proposed two analytical models for optimization of the supply chains of importers of waterborne containerized goods from Asia to USA. These models determine the least-cost supply-chain strategy for importers, in terms of ports and landside channels to be used. The costs considered include costs for transportation and handling, pipeline inventory, and safety-stock inventory at RDCs. The Long-Run Model, is a large mixed integer non-linear programming model, with a set of heuristics to solve it. In the Long-Run Model, the mean and standard deviation of container flow times by channel are fixed, reflecting an assumption that over the long term the various ports and transportation carriers would make investments to maintain existing service quality and thereby protect market share. The Short-Run Model is an integration of the Long-Run Model with the Queuing Model which estimates the flow times as a function of traffic volumes. Calculations of the Short-Run Model involve iterative runs of the Long-Run Model and the Queuing Model. The Short-Run Model assumes the infrastructure of the entire transportation network is fixed. Container flow times are endogenous in this model, responding to congestion (or lack thereof) in various ports and channels. The Short-Run Model is thus useful for projecting more near-term responses of importers to changes in fees, rates or infrastructure.

To show an application of proposed models, we used them to predict the effect of imposing container fees at the San Pedro Bay Ports (i.e. Los Angeles and Long Beach ports). We observed that the Long-Run elasticity of imports via San Pedro Bay is roughly double the Short-Run elasticity. If a container fee is imposed, most of the volume leaving the San Pedro Bay ports would be diverted to the Puget Sound ports. Our analyses make clear that the discretionary inland-point intermodal (IPI) volumes through the San Pedro Bay Ports are very elastic and decline rapidly with growing fee values. Trans-loaded imports with moderate declared values are somewhat less elastic, while trans-loaded imports with high declared values and imports consumed within the local region are very inelastic. Comparing Short-Run and Long-Run results, it is clear that adequate infrastructure, and staffing of that infrastructure, are not yet in place at other ports to accommodate without congestion the diversion of substantial volumes of trans-loaded imports away from San Pedro Bay ports. However, the economics encouraging expansion at other ports and their landside channels arises when fees greater than $150 per FEU are imposed on imports through the San Pedro Bay ports.

The contributions of this article to the literature are in several domains. First, for the challenge of predicting flows of imports by port and landside channels, our approach breaks new ground by (1) incorporating formal optimization of import flows from the point of view of the importers, and (2) providing both Short- and Long-Run predictions of import flows. We account for contractual minimum required volumes at ports, and by considering the effect of traffic volume on flow times in the Short-Run Model, we address the capacity constraints of ports and landside channels. Second, the underlying optimization model extends the existing literature in supply-chain design and management in several ways: (1) it fills a gap in the existing literature for location/allocation problems by considering risk pooling economics when making choices of transportation mode and routing; (2) it integrates analysis of stochastic demand and random transportation time uncertainties when making location decisions in a multi-echelon setting; and (3) it enriches the literature on intermodal freight transport systems by providing a strategic-level and tactical-level service design tool.

The authors believe that the proposed Short-Run and Long-Run Models show much promise for interesting policy analysis and infrastructure planning. It is exciting to be able to capture a complete view of Asia – US imports, the economics involved, and the limitations of current infrastructure and logistics services. However, in the authors' opinion, the amount of data on which the Short-Run Model was calibrated is marginally adequate; much more could be done to refine the models as well as to facilitate wider application for improved policymaking, strategic planning, capital budgeting and financing of transportation infrastructure improvements.

Several directions for future research are apparent from this study. First, in this article, we provided analysis only for hypothetical fee increases and infrastructure improvements at the San Pedro Bay Ports. Many other applications and policy-level analyses can be conducted to help decision-makers make wiser decisions on infrastructure investments and user fees for ports and landside channels. The proposed models can help transportation carriers and warehousing service providers to make more informed pricing decisions; Second, this research could be expanded to include the imports from other parts of the world to North America; Third, in addition to the imports, this research could be expanded to include the exports and suggest the best supply chains for exporters; Fourth, in addition to transportation and inventory costs, the proposed models could be further expanded to include other parameters such as environmental factors. The efficiency and effectiveness of these approaches could be further studied under different scenarios.

Acknowledgements

This research was sponsored by the Southern California Association of Governments. The authors are grateful to many port terminal operators, importers, intermodal marketing companies, third party logistics companies, steamship lines and railroads for providing data for our studies.

References

Ambrosino, D., Scutellà, M.G., 2005. Distribution network design: new problems and related models. European Journal of Operational Research 165 (3), 610–624.

Caris, A., Macharis, C., Janssens, G.K., 2008. Planning problems in intermodal freight transport: Accomplishments and prospects. Transportation Planning and Technology 31 (3), 277–302.

Daskin, M.S., Coullard, C., Shen, Z.-J., 2002. An inventory-location model: formulation, solution algorithm and computational results. Annals of Operations Research 110 (1–4), 83–106.

Erlebacher, S.J., Meller, R.D., 2000. The interaction of location and inventory in designing distribution systems. IIE Transactions 32 (2), 155–166.

Fan, L., Wilson, W.W., Tolliver, D., 2010. Optimal network flows for containerized imports to the United States. Transportation Research Part E 46, 735–749.

Imai, A., Shintani, K., Papadimitriou, S., 2009. Multi-port vs. Hub-and-Spoke port calls by containerships. Transportation Research Part E 45, 740–757.

Jula, P., Leachman, R.C., 2011. A supply-chain optimization model of the allocation of containerized imports from Asia to the United States. Transportation Research Part E 47, 609–622.

Leachman, R.C., Jula, P., submitted for publication. Congestion analysis of waterborne, containerized imports from Asia to the United States. Working Paper. Transportation Research Part E.

Leachman, R.C., Jula, P., submitted for publication. Estimating flow times for containerized imports from Asia to the United States through the western rail network. Working Paper. Transportation Research Part E.

Leachman, R.C., 2010. Final report – port and modal elasticity study, Phase II. Consulting report prepared for the Southern California Association of Governments. <http://www.scag.ca.gov/goodsmove/elasticitystudyphase2.htm>.

Leachman, R.C., 2008. Port and modal allocation of waterborne containerized imports from Asia to the United States. Transportation Research Part E 44 (2), 313–331.

Levine, C., Nozick, L., Jones, D., 2009. Estimating and origin-destination table for US imports of waterborne containerized freight. Transportation Research Part E 45, 611–626.

Liu, S.C., Lin, C.C., 2005. A heuristic method for the combined location routing and inventory problem. International Journal of Advanced Manufacturing Technology 26 (4), 372–381.

Ma, H.Z., Davidrajuh, R., 2005. An iterative approach for distribution chain design in agile virtual environment. Industrial Management and Data Systems 105 (5–6), 815–834.

Manzini, R., Bindi, F., 2009. Strategic design and operational management optimization of a multi stage physical distribution system. Transportation Research Part E 45 (6), 915–936.

Melo, M.T., Nickel, S., Saldanha-da-Gama, F., 2009. Facility location and supply chain management – a review. European Journal of Operational Research 196 (2), 401–412.

Miranda, P.A., Garrido, R.A., 2004. Incorporating inventory control decisions into a strategic distribution network design model with stochastic demand. Transportation Research Part E 40 (3), 183–207.

ReVelle, C.S., Eiselt, H.A., Daskin, M.S., 2008. A bibliography for some fundamental problem categories in discrete location science. European Journal of Operational Research 184 (3), 817–848.

Shen, Z.-J., 2007. Integrated supply chain design models: a survey and future research directions. Journal of Industrial and Management Optimization 3 (1), 1–27.

Shen, Z.-J., Qi, L., 2007. Incorporating inventory and routing costs in strategic location models. European Journal of Operational Research 179 (2), 372–389.

Shen, Z.-J., Coullard, C., Daskin, M.S., 2003. A joint location-inventory model. Transportation Science 37 (1), 40–55.

Shu, J., Teo, C.P., Shen, Z.-J., 2005. Stochastic transportation-inventory network design problem. Operations Research 53 (1), 48–60.

Silver, E.A., Peterson, R., 1985. Decision Systems for Inventory Management and Production Planning. John Wiley and Sons, New York.

Snyder, L.V., Daskin, M.S., Teo, C.P., 2007. The stochastic location model with risk pooling. European Journal of Operational Research 179 (3), 1221–1238.

Wilhelm, W., Liang, D., Rao, B., Warrier, D., Zhu, X., Bulusu, S., 2005. Design of international assembly systems and their supply chains under NAFTA. Transportation Research Part E 41 (6), 467–493.