

DESIGN OF ROBUST DISTRIBUTION NETWORKS RUN BY THIRD PARTY LOGISTICS SERVICE PROVIDERS

M. P. M. HENDRIKS*, D. ARMBRUSTER^{*,†,¶}, M. LAUMANN^{‡,||},
 E. LEFEVER^{*,§} and J. T. UDDING*

**Department of Mechanical Engineering,
 Manufacturing Networks Group,
 Technische Universiteit Eindhoven,
 P. O. Box 513, 5600 MB Eindhoven, The Netherlands*

*†School of Mathematical and Statistical Sciences,
 Arizona State University,
 Tempe, 85287, USA*

*‡Institute for Operations Research,
 ETH Zurich, 8092 Zurich, Switzerland*

§A.A.J.Lefeber@tue.nl

¶Armbruster@asu.edu

||Laumanns@ifor.math.ethz.ch

Received 11 August 2011

Revised 15 November 2011

Published 26 June 2012

We consider a third party logistics service provider (LSP), who faces the problem of distributing different products from suppliers to consumers having no control on supply and demand. In a third party set-up, the operations of transport and storage are run as a black box for a fixed price. Thus the incentive for an LSP is to reduce its operational costs. The objective of this paper is to find an efficient network topology on a tactical level, which still satisfies the service level agreements on the operational level. We develop an optimization method, which constructs a tactical network topology based on the operational decisions resulting from a given model predictive control (MPC) policy. Experiments suggest that such a topology typically requires only a small fraction of all possible links. As expected, the found topology is sensitive to changes in supply and demand averages. Interestingly, the found topology appears to be robust to changes in second order moments of supply and demand distributions.

Keywords: Network topology design; logistics service provider; bi-level optimization; robustness.

1. Introduction

The logistics networks considered in this paper consist of production facilities, warehouses, and consumers, which are geographically connected by links, e.g. roads, railways, waterways. In long-distance transportation networks, the distribution of products is often performed by a logistics service provider (LSP). As opposed to the

common supply chain studies, the problem addressed here is not to control the amount of products in the supply chain. Instead, this study addresses one of the typical services an LSP provides: Supply and demand cannot be influenced by the LSP, but are simply a (stochastic) reality, that is revealed only a few days in advance. A supplier pushes its products from several production facilities to the logistics provider. By the same token, consumers pull products from the network. The logistics provider then has to decide whether to store the products in a warehouse or to immediately match them with a consumer demand.

The decisions regarding design and operation of distribution networks are typically classified into three levels: the strategic level, the tactical level, and the operational level. The *strategic level* deals with decisions regarding the number, location and capacities of warehouses. These decisions have a long-lasting effect on the system's performance. The goal on the *tactical level* is to construct a network topology with a small number of fixed line hauls (transportation links between two facilities). Decisions on the network topology are made on the medium term time scale. To preserve continuity for employees and to restrict the complexity of organizational tasks, such a network topology should not change too frequently. According to the economies of scale principle, the efficiency of a line haul increases if more products flow through it. The average line haul utilization can simply be increased by decreasing the number of line hauls in the network topology. Therefore, an LSP strives for a network topology with a small number of links. The *operational level* refers to short term decisions such as scheduling and routing of the daily shipments given the tactical topology.

In this paper, we focus on the interplay between the operational level and the tactical level, i.e. we are interested in determining the topology of the network for a given operational strategy and given operational parameters (see [7]). We specifically strive for a robust topology, which can be established for a relatively long period of time (months or years) and is still cost-effective when operational parameters (supply and demand distributions) change.

The operation of an LSP is run as a black box for a fixed price. Within certain boundaries with respect to variability in supply and demand, the LSP is contractually bound to supply the desired quantities at the desired day. In this case, there are high incentives for an LSP to reduce its operational costs, within the limits set by the service contract.

The LSP can compensate for the stochasticity in supplies and demands by temporarily storing products in warehouses rather than shipping them directly from supplier to consumer. Furthermore, shipping via a warehouse enables product mixing so as to leverage on the economies of scale. On the other hand shipping through a warehouse introduces additional delay due to the handling activities, e.g. (un)loading and consolidation. In settings like the automotive industry, actual orders are accurately released to the LSP a couple of days in advance. The LSP can use this information to decide which transportation links to use in the long run and how much of which product(s) to ship through them each day, such that total

costs, including transportation and storage costs and penalties for early and late deliveries, are minimized.

1.1. Contributions

We determine the structure of the tactical network topology (the link or line hauls to choose) dependent on the decisions constructed at the operational level. We specifically strive for a network with a very small number of links that still has close to minimal operational cost. Reducing the number of links is a surrogate for a more detailed optimization at the tactical level for which the associated costs (for establishing a link) are hard to quantify:

- A reduction in the number of links reduces the complexity of the network and with that the complexity of the organizational tasks of an LSP,
- Each link involves a fixed cost due to contracts and overhead,
- Reducing the number of links leads to thicker flows per link.

The overall problem can thus be characterized as a bi-level joint problem: the upper (tactical) level chooses the links and the lower (operational) level, for given choice of links, chooses the material flow. Our goal is to provide an efficient method to approximate the trade-off frontier of link cost versus operational cost for given stochastic supply and demand processes of which only the mean values are known, and in particular to determine a network topology that is (i) approximately optimal in both objectives (small number of links and minimal operational costs) and (ii) robust with respect to stochasticity (second order moments) of supply and demand.

In Sec. 2 we solve the lower-level operational problem (Sec. 2.1) of controlling the material flow for a given network topology. A stepwise cost function is introduced in the flow model to represent the costs for transportation by trucks (economies of scale). Additional cost for storage and for missed shipment dates are imposed. A model predictive control with rolling horizon (MPC) [5] is used to deal with changing demand realizations.

The upper-level refers to the tactical decision making and is discussed in Sec. 2.2. We present a branch-and-bound method that computes the operational costs as a function of the number of links for given multiple-product supply and demand time series. The method is impractical for larger networks and hence we propose a heuristic that works by iteratively dropping least used links and replaces the stepwise cost model by a linear cost model. For small networks we study the full and the approximate optimization problem and find similar results in both the costs curve and the topology structure. For the heuristic approximation, which is very efficient, we can usually provide a network with a small number of links and close to minimal operational costs. From that we can conclude that the topology is also close to optimal.

The full optimization problem is computationally intractable at various different levels: The upper-level problem involves binary decisions variables to model the

choice of links reflecting the combinatorial explosion of the number of different networks for a set of nodes. This is the reason for investigating the heuristics of dropping the least used links. For each chosen network we will have to solve the lower level problem of scheduling the flows. However, the lower-level problem is a stochastic control problem and intractable due to the short time horizon in which orders become fixed. We therefore use an MPC approach as a heuristic here (which also mimics the operational practice of the logistics provider). The MPC approach requires to solve a limited horizon optimization problem in each time step. For this problem, we consider (i) a version where we model the fixed cost for each used truck explicitly, which requires additional binary variables that make the problem again intractable (solvable in practice for only small instances) and (ii) a version that linearizes this cost function to arrive at a convex quadratic program (the quadratic term due to the quadratic backlog cost, which can be then solved efficiently). The latter one can be seen as a heuristic (or approximation) of the former one.

Section 3 compares the branch and bound method and the heuristics approach for small problems and applies the heuristics approach to the linear model for several large network configurations. The sensitivity of the resulting networks under stochastic variations of demand and supply is studied. We end with conclusions and recommendations for future work in Sec. 4.

1.2. Related work

Most related work is framed in the context of supply chain modeling which, as discussed before, has a very different decision space. There is a very limited set of papers that deal with the issues facing an LSP.

Similar to our heuristic approach [10] determines an near optimal number and locations of transshipment hubs (strategic level) in a supply network by starting from a network in which all potential hubs are present and successively dropping the hubs that generate the largest cost decrease upon deletion. The process terminates whenever deleting a hub does not significantly affect the costs. For a stochastic supply chain setting, only [6, 13] model transportation costs by a stepwise cost-function dependent on the capacity of a transportation device (economies of scale). Such an approach reflects practice more accurately than a linear model at the expense of higher computational time. As discussed we follow this practice until the networks become too large and we have to revert back to a linear model.

To our knowledge, only a restricted number of studies consider the viewpoint of an LSP where supply and demand are both stochastic and uncontrollable [2, 14]. In [2], models and algorithms for a one product, multi-stage stochastic distribution problem with recourse are developed. In [14] a company owns several production plants and has to distribute one type of product to different regional markets. In each time period, a random (uncontrollable) amount of product becomes available at each of these plants. Before the random demand becomes available, the company has to decide which proportion of the products should be shipped directly and which

proportion should be held at the production plants. It is numerically shown that a dynamic programming method yields high quality solutions of the problem. One of the results of [14] is that most improvements on the operational costs are made when a part of the consumers is served by two plants, rather than one.

Our success to reduce large networks to networks with a small number of links with similar costs suggest that the network with only a few links provides a large portion of the operational efficiency of the fully connected network. This confirms other studies on different kinds of two-echelon networks. Results in [9] for instance suggest that, for a small theoretical problem, limited flexibility in manufacturing processes (i.e. each plant builds only a few products) yields most of the benefits of total flexibility (i.e. each plant builds all products). However, the authors mention that for more realistic cases they have no guidelines or general approach to add flexibility in an arbitrary network. In this paper, this issue is addressed by developing a heuristic that for more realistic cases and for more than a two echelon system generates a network with a small number of links that still performs well.

2. Bi-Level Network Design Problem

We consider the distribution of K types of products by trucks through a network with S production facilities, W warehouses, and D consumers. Unless stated differently, we use the indices $i \in \{1, \dots, S\}$ for the production facilities, $w \in \mathcal{W} = \{1, \dots, W\}$ for the warehouses, $j \in \{1, \dots, D\}$ for the consumers, and $k \in \{1, \dots, K\}$, for the product types. Products can be sent either directly from supplier(s) to consumer(s) or indirectly via a warehouse, assuming that the LSP is free to choose from which supplier(s) a consumer receives a certain product. Shipments from one warehouse to another are not taken into account yet, but could easily

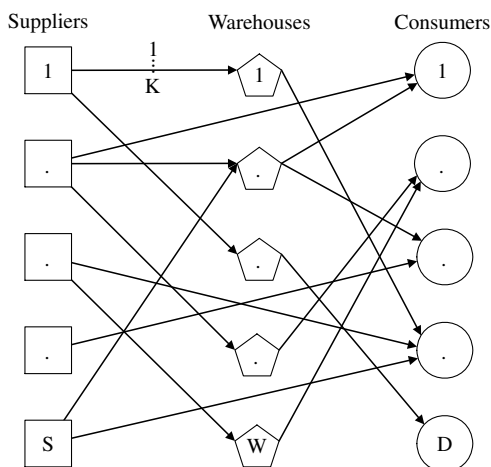


Fig. 1. Illustration of the considered logistics networks. An arrow represents a line haul (link) from one facility to another.

be incorporated into this framework as well. An example network for this type of three-echelon multi-item distribution system is depicted in Fig. 1. Even though each link in the graph has a certain length representing the geographical distance between the corresponding nodes, Fig. 1 does not show the spatial positioning of the nodes for ease of presentation.

Each production facility supplies one or multiple product types, and each consumer demands one or multiple product types, at each time step (day) $t \in \{1, \dots, T\}$. The supply and demand time series $\mathcal{S}_i^k(t)$ (supply of product type k by supplier i at day t) and $\mathcal{D}_j^k(t)$ (demand of product type k by consumer j at day t) are stochastic, and a realization is assumed to be given.

2.1. Operational decision making

Below we first present the decision and auxiliary variables and then discuss the model constraints subsequently.

2.1.1. Variables

To formalize the operational task of the LSP, we define the operational decision variables $x_{ij}^k(t) \geq 0$ as the amount of products of type k transported from node i to node j on time step t , where $(i, j) \in \mathcal{U}$, the set of links in the given network. On the supplier's side, we require that all available products $\mathcal{S}_i^k(t)$ have to be picked up at time t such that

$$\mathcal{S}_i^k(t) = \sum_{j \in P_i} x_{ij}^k(t) \quad \forall k, i, t \quad (1)$$

where the set P_i contains the indices of all the nodes to which production facility i is connected. On the consumers side, both early and late deliveries might occur and have to be accounted for. For this we introduce state variables $b_j^k(t)$ as the backlog of product type k for consumer j on day t , whose dynamics is

$$b_j^k(t) = b_j^k(t-1) + \mathcal{D}_j^k(t) - \sum_{i \in Q_j} x_{ij}^k(t) \quad \forall k, j, t \quad (2)$$

where the set Q_j contains the indices of all the nodes connected to consumer j . In the warehouses, several processes (e.g. inbound, consolidation) between arriving at and storing in a warehouse cause delay. The total delay is captured in a constant $\tau_w \in \mathbb{N}$, $\tau_w \ll T$, so that products arriving in warehouse w at time t can be shipped out τ_w time steps later at the earliest. To model this time delay, we introduce additional state variables $y_w^k(t)$ as the inventory of product k in warehouse w on day t , whose dynamics is

$$y_w^k(t) = y_w^k(t-1) + \sum_{i \in I_w} x_{iw}^k(t - \tau_w) - \sum_{j \in O_w} x_{wj}^k(t) \quad \forall k, w, t \quad (3)$$

where the set I_w contains the indices of all nodes that are sources of warehouse w , the set O_w contains the indices of all nodes that are destinations of warehouse w , and $y_w^k(0) = \mathcal{Y}_w^k$ denotes the initial inventory, which we assume to be 0.

The operational cost for time step t is the sum of transportation cost, storage cost, and backlog cost. It can be expressed as a function of the decision variables $x_{ij}^k(t)$ and state variables $b_j^k(t)$ and $y_w^k(t)$ at time t as

$$\begin{aligned} & \alpha(x_{ij}^k(t), b_j^k(t), y_w^k(t)) \\ &= \sum_{(i,j) \in \mathcal{U}} c_{ij} \phi \left(\sum_{k=1}^K x_{ij}^k(t) \right) + \sum_{k=1}^K \sum_{j=1}^D \beta^k b_j^{k^2}(t) + \sum_{w \in \mathcal{W}} h_w \sum_{k=1}^K y_w^k(t) \end{aligned} \quad (4)$$

where c_{ij} represents the geographical distance between nodes i and j , h_w represents the inventory holding costs in warehouse w , and β^k represents the “value” of product k . The reason for not using the absolute value $|b_j^k(t)|$ as a penalty for early or late delivery is that in case of shortages, the quadratic function favors equal distribution of shipments among consumers rather than shipping everything to one consumer. The function $\phi(\cdot)$ expresses quantity-dependent transportation costs. We consider two cases, (i) $\phi(x) \equiv V \lceil \frac{x}{V} \rceil$ to express stepwise-constant transportation costs due to the use of trucks with unit capacity V , and (ii) the identity $\phi(x) \equiv Vx$ to express the simplest case of linear transportation costs.

We further assume that suppliers and consumers commit to their real supplies and demands for a constant number of time steps Ω in advance, where $\Omega \ll T$. Thus, the exact supplies and demands for the current and the next $\Omega - 1$ time steps are known and can be taken into account when deciding on the transportation quantities.

The method applied by Frans Maas is to use a limited look-ahead scheme in a rolling horizon fashion, also called model predictive control (MPC). The basic idea of MPC is that optimization is not performed over an infinite horizon, but only performed over a limited horizon where more information about uncertain data is available, or even well representable by deterministic or nominal values, and to choose some reasonable approximation of the value function for the time steps outside the horizon, which typically is used as terminal costs. Here, we use a horizon of Ω steps as the exact supply and demand data is known over this time frame. Thus, we obtain an approximate control policy $\hat{\mu}$ by solving the following time-expanded network flow problem:

$$\text{minimize } \sum_{q=t}^{t+\Omega-1} \left[\sum_{(i,j) \in \mathcal{U}} c_{ij} u_{ij}(q) + \sum_{k=1}^K \sum_{j=1}^D \beta^k [b_j^k(q)]^2 + \sum_{w \in \mathcal{W}} h_w \sum_{k=1}^K y_w^k(q) \right]$$

subject to

$$\mathcal{S}_i^k(q) = \sum_{j \in P_i} x_{ij}^k(q) \quad \forall k, i, q$$

$$y_w^k(q) = y_w^k(q-1) + \sum_{i \in I_w} x_{iw}^k(q - \tau_w) - \sum_{j \in O_w} x_{wj}^k(q) \quad \forall k, w, q$$

$$\begin{aligned}
b_j^k(q) &= b_j^k(q-1) + \mathcal{D}_j^k(q) - \sum_{i \in Q_j} x_{ij}^k(q) \quad \forall k, j, q \\
V \cdot u_{ij}(q) &\geq \sum_{k=1}^K x_{ij}^k(q) \quad \forall i, j, q \\
x_{ij}^k(q) &\geq 0 \quad \forall k, i, j, q \\
y_w^k(q) &\geq 0 \quad \forall k, w, q
\end{aligned} \tag{5}$$

Minimization is performed over the variables $x_{ij}^k(q)$, $u_{ij}(q)$, $b_j^k(q)$, and $y_w^k(q)$, where $q \in \{t, \dots, t+\Omega-1\}$, initial backlog $b_j^k(0) = \mathcal{B}_j^k$ and initial inventory $y_w^k(0) = \mathcal{Y}_w^k$. The additional variables $u_{ij}(q)$ are introduced to linearize the transportation cost function $\phi(\cdot)$ and represent the number of trucks of capacity V necessary to transport the total amount of products over link (i, j) at time step q . To model the step-wise constant transportation cost, these variables have to be restricted to integer values; otherwise a continuous relaxation can be used to model linear transportation costs.

For all time steps prior to t , the values of variables are known so that they serve as deterministic data in the above limited look-ahead problem. The same holds for the values of supplies $\mathcal{S}_i^k(q)$ and demands $\mathcal{D}_j^k(q)$ within the considered horizon $\{t, \dots, t+\Omega-1\}$. Consequently, the problem is a mixed-integer quadratic program (in case of stepwise-constant transportation costs) or a quadratic program (in case of linear transportation costs). The model parameters are nicely arranged in Table 1.

2.2. Tactical decision making

In a given collection of suppliers, warehouses, and consumers, let \mathcal{L} be the union of all potential links (directed edges) from suppliers to warehouses, warehouses to consumers, and suppliers to consumers. The tactical problem can then be defined

Table 1. Model parameters.

Parameter	Definition
T	Number of discrete time slots considered
K	Number of product types
S	Number of suppliers in the network
W	Number of warehouses in the network
D	Number of consumers in the network
$\mathcal{S}_i^k(t)$	Amount of products of type k supplied at time t by supplier i
$\mathcal{D}_j^k(t)$	Amount of products of type k demanded at time t by consumer j
τ_w	Delay in warehouse w due to consolidation activities
Ω	Horizon length
V	Capacity of a truck
c_{ij}	Costs for sending a truck from node i to node j
h_w	Costs per day for storing a product in warehouse w
β^k	Costs for not delivering product k in time

as choosing a subset $\mathcal{U} \subseteq \mathcal{L}$ of a given cardinality L with smallest expected operational cost. In doing so, we tacitly assume that there exists some mapping γ_μ that assigns the expected operational costs $\gamma_\mu(\mathcal{U})$ to the network of chosen links \mathcal{U} when using a particular control policy μ . As it is unclear how this expectation can be computed exactly in the present set-up, we resort to a variant of sample average approximation by Monte Carlo simulation, where we simulate the process for a fixed number of time steps while applying the given policy μ . Here, we use the approximate limited look-ahead policy $\hat{\mu}$ defined above, and refer to the resulting estimator of the expected operational cost as $\hat{\gamma}_{\hat{\mu}}(\mathcal{U})$.

In a bi-objective formulation, this tactical problem is to be solved for any value of $L \in \{L_{\min}, \dots, |\mathcal{L}|\}$, where L_{\min} is the smallest number of links such that no node is isolated and each warehouse is connected to at least one supplier and one consumer. Our conjecture about this trade-off between the number of chosen links and the operational costs is that the operational costs are only marginally sensitive to a reduction of the number of links from the fully connected network until a critical value, after which the costs increase sharply. In this case, the cost curve would form a pronounced knee. Hence, our particular goal is to determine a representative network from this region, as this would constitute in a sense a best-possible bi-objective approximation. Such a network could be considered “cost-effective” as it would yield close-to-minimal operational costs with a close-to-minimal number of links.

To determine the cost curve, we first develop a bi-objective branch and bound method that computes the minimal cost topology for any given number of links simultaneously. As this method still searches a large part of the solution tree, we additionally propose a heuristic, which is able to generate a satisfactory solution for larger instances.

2.2.1. Bi-objective branch and bound method

We develop a branch and bound algorithm that defines a tree search over a set of binary decision variables z_{ij} associated with each link $(i, j) \in \mathcal{L}$, where

$$z_{ij} = \begin{cases} 1 & \text{if the link from facility } i \text{ to facility } j \text{ is present,} \\ 0 & \text{otherwise.} \end{cases}$$

A pseudo-code description of the algorithm is given below (Algorithm 1). The algorithm starts with a fully connected network and no bound on the minimal cost and maintains and updates a current approximation to the cost curve $\hat{\gamma}_{\hat{\mu}}$.

During the search, the L -entry is updated whenever a network z with L links is identified to have lower operational cost than the current best network with L links (lines 14–16). At each branching operation (lines 19–27), a natural lower bound is given by setting all undecided variables to one as all solutions in this subtree can only be composed of a subset of those links. The subtree can be pruned if this lower bound is not lower than the current best value for any number of links L for which

Algorithm 1. Branch and Bound

```

1: let  $z := (\star, \star, \dots, \star)$ 
2: let  $\text{LB}(z) := -1$ 
3: initialize the list of unexplored nodes  $A$  with  $z$ 
4: while  $A$  is not empty do
5:   pick  $z$  as the last element from  $A$ 
6:   delete  $z$  from  $A$ 
7:   let  $p$  be the number of ones in  $z$  and  $q$  be the number of zeros in  $z$ 
8:   if  $p + q = |\mathcal{L}|$  then
9:     if  $\text{LB}(z) = -1$  then
10:      let  $c := \hat{\gamma}_{\hat{\mu}}(z)$ 
11:     else
12:      let  $c := \text{LB}(z)$ 
13:     end if
14:     if  $c < \text{opt}(p)$  then
15:       let  $\text{opt}(p) := c$ 
16:       let  $Z_{\text{opt}}(p) := z$ 
17:     end if
18:   else
19:     if  $\text{LB}(z) = -1$  then
20:       let  $z'$  be a copy of  $z$  where all  $\star$  are replaced by 1
21:       let  $\text{LB}(z) := \hat{\gamma}_{\hat{\mu}}(z')$ 
22:     end if
23:     if  $\text{LB}(z) < Z_{\text{opt}}(L)$  for some  $L$  with  $p \leq L \leq |\mathcal{L}|$  then
24:       let  $z''$  be a copy of  $z$  where the first  $\star$  is replaced by 0 and append  $z''$ 
        to  $A$ 
25:       let  $z'''$  be a copy of  $z$  where the first  $\star$  is replaced by 1 and append  $z'''$ 
        to  $A$ 
26:       let  $\text{LB}(z'') := -1$  and  $\text{LB}(z''') := \text{LB}(z)$ 
27:     end if
28:   end if
29: end while

```

there are potential solutions in the subtree; otherwise, two children are created by setting the first undecided variable to zero and one, respectively, and added to the list of unexplored nodes A .

2.2.2. Heuristic

Since the branch and bound method evaluates a large number of possible topologies, only small problems can be solved. Hence, we propose a heuristic that evaluates less than $|\mathcal{L}|$ topologies and still yields a very good solution. The heuristic starts from

Algorithm 2. Heuristic

-
- 1: let \mathcal{U} be the fully connected network \mathcal{L}
 - 2: **while** the network is not minimally connected **do**
 - 3: apply the policy $\hat{\mu}$ during K time steps and determine $\hat{\gamma}_{\hat{\mu}}(\mathcal{U})$
 - 4: determine usage P_{ij} for each link in \mathcal{U}
 - 5: delete the link(s) from \mathcal{U} with minimal usage
 - 6: **end while**
-

the fully connected network and approximates the operational costs $\hat{\gamma}_{\hat{\mu}}(\mathcal{L})$ when using the control policy $\hat{\mu}$ over the fixed number of time steps. Then the link(s) that are least used over this time period are deleted. In determining the usage of a link, the value of the product type(s) through this link is taken into account by a weight factor. We define the weighted average daily usage P_{ij} of the link between facilities i and j as

$$P_{ij} = \frac{1}{T} \sum_{t=1}^T \sum_{k=1}^K \frac{\beta^k}{\beta_{\max}} \cdot x_{ij}^k(t), \quad \text{where } \beta_{\max} = \max_k \beta^k. \quad (6)$$

In this way, links that provide just-in-time delivery of high value products are not (easily) deleted. This procedure of deleting links is repeated until a minimally connected network remains. A pseudo-code description of the algorithm is given in Algorithm 2.

3. Computational Study

In the previous section, we proposed two algorithms to find a network topology with a close to minimal number of links that is still cost-effective on the operational level. In this section, the performance of both of these algorithms is evaluated and compared. Experiments suggest that the heuristic is very effective in constructing a network topology with a small number of links that yields close to minimal operational costs. However, large real-life problems are still intractable. Hence, we replace the stepwise cost function by a linear cost function and find similar results much faster. Some large real-life instances are solved for the linear cost structure and results are presented.

To evaluate the performance of the heuristic, a bi-objective relative approximation factor $(\varepsilon_c, \varepsilon_L)$ is used as a performance measure. The factor denotes the relative deviation from the ideal point, the point composed of the single-objective optima. In this case, those values are known, so we can define the two components as

$$\varepsilon_c(\mathcal{U}) := \frac{\hat{\gamma}_{\hat{\mu}}(\mathcal{U})}{\hat{\gamma}_{\hat{\mu}}(\mathcal{L})} \quad \text{and} \quad \varepsilon_L(\mathcal{U}) := \frac{|\mathcal{U}|}{L_{\min}}, \quad (7)$$

denoting the operational costs of the considered topology \mathcal{U} relative to the operational costs in the fully connected network, and the number of links $|\mathcal{U}|$ relative to the number of links in a minimally connected network. Even though we are not

able to give some *a priori* performance guarantee of the heuristic, this factor can be used to bound the approximation quality a posteriori.

In addition to evaluating the performance of the heuristic approach, a computational study is performed to investigate the robustness of the generated network topology with respect to changes in the supply and demand distributions. Results suggest that a “heuristic network” with a satisfactory $(\varepsilon_c(\mathcal{U}), \varepsilon_L(\mathcal{U}))$ value is cost-effective as long as the means of supply and demand distributions remain the same. Finally, we formulate the hypothesis that increasing the initial inventory can compensate for the backlog costs embedded in a particular supply and demand time series. The hypothesis is confirmed by experimental results.

For the logistics networks in this paper, we consider S suppliers, W warehouses, D consumers and the distribution of K product types. The networks considered typically consist of few suppliers and warehouses and a lot of consumers, such that $W < S \ll D$.

The supply and demand time series $\mathcal{S}_i^k(t)$ (supply of product type k by supplier i at day t) and $\mathcal{D}_j^k(t)$ (demand of product type k by consumer j at day t) are randomly distributed according to a joint distribution that is defined by the following sampling procedure: for each combination of product type, supplier and consumer, we assume a uniform distribution on an interval $[\mu_i^k(t) - \hat{\sigma}_i^k(t), \mu_i^k(t) + \hat{\sigma}_i^k(t)]$ with given mean $\mu_i^k(t)$ and half-length $\hat{\sigma}_i^k(t)$. Next, T preliminary samples out of each of the $K \cdot (S + D)$ distributions are taken. The final samples are then determined by adding a constant either to each supply or each demand sample for each product type to enforce that for each product type, the sum of T supplies equals the sum of T demands over the horizon. We generate stochastic time series of $T \cdot S$ supplies and $T \cdot D$ demands from uniform distributions with random means and variances for 100 time steps. We assume that supplies and demands are only known 2 days in advance, so the window size Ω is equal to 3. Moreover, we focus on networks of relatively small geographical extent. Therefore, it is assumed that products can be transported from any arbitrary node to another within one day. In addition, we assume that the delay for going through warehouse n is equal to one day, which implies that $\tau_w = 1$. The costs are chosen such that the proportion of transportation and storage costs $c_{i,j} : h_n = 1 : 0.6$ on average and the early/late delivery costs β^k are randomly chosen between the bounds of 0.001 and 100.

3.1. Comparison: Branch and bound versus heuristic

First, we compare the results of the branch and bound method and the results of the heuristic applied to the stepwise cost model. Hence, we generate different problem instances (supply and demand distributions as well as spatial location of network nodes) for many small network configurations and run the branch and bound method as well as the heuristic. Comparison of the resulting cost curves of operational cost versus number of links show that the heuristic generates a topology that is close to optimal.

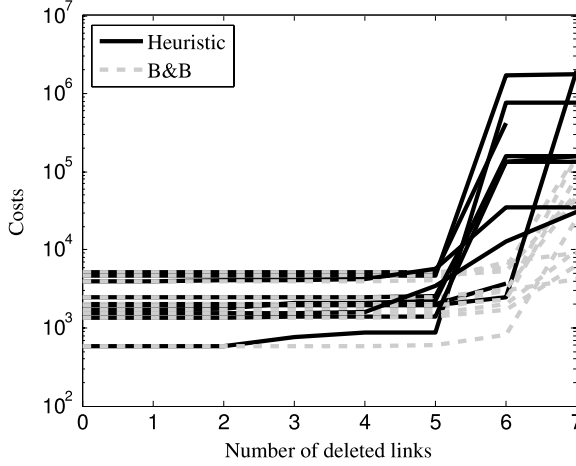
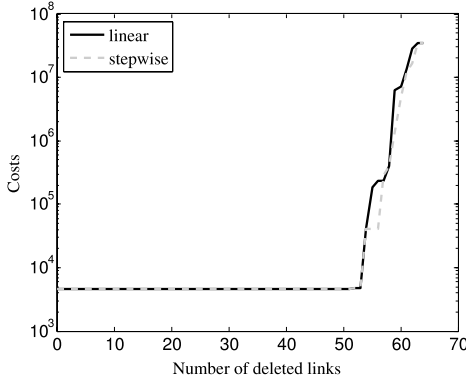


Fig. 2. Comparison of branch and bound with heuristic for 10 supply and demand instances, $S = 2$, $W = 1$, $D = 3$, $K = 1$, Branch and bound: $\varepsilon_c \leq 1.01$, $\varepsilon_L = 1.5$, Heuristics: $\varepsilon_c \leq 1.01$, $\varepsilon_L = 1.63$.

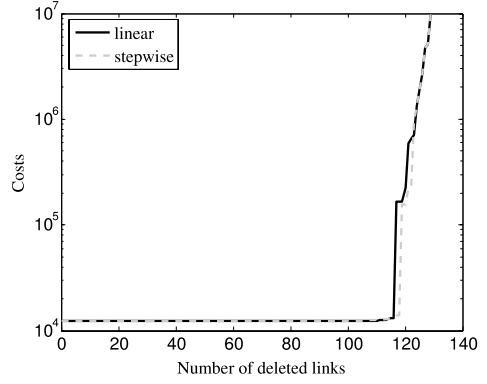
To illustrate this we use a network with $S = 2$, $W = 1$, $D = 3$, and $K = 1$. A hundred sets of supply and demand time series and graphical supply, demand and warehouse locations are generated. Next, for each of these sets, a network topology is constructed with both the branch and bound, and the heuristics method.

Results for both the branch and bound method and the heuristic are shown in Fig. 2. For clarity of presentation we only depict results for ten different parameter sets, which give a proper representation of all hundred performed experiments. As can be seen from this figure, both methods generate equal operational costs for the fully connected network (number of deleted links equals zero). From Fig. 2, it also appears that the critical value in the efficient frontier found by the heuristic occurs at 5 or 6 deleted links, whereas it always occurs at 6 deleted links for the networks generated by the branch and bound approach. The difference in the number of links in the resulting network for these instances is thus at most 1 link. This suggests that the heuristic finds a satisfactory solution for small network instances. Since the CPU time of the branch and bound grows extremely fast with increasing network size, we have to rely on the heuristic for larger instances.

The heuristic, applied to the stepwise cost model, enables to find a close to optimal topology for a network with up to 4 suppliers, 3 warehouses, 25 consumers and 6 product types ($S = 4$, $W = 3$, $C = 25$, and $K = 6$, respectively). Due to the non-linearity, larger real-life problems are still intractable. We therefore remove the integer variables u_{ij} from the model and introduce linear transportation costs ($\phi(x) \equiv Vx$). We apply the heuristic to this linear model, which is now able to construct cost-effective topologies for large real-life network sizes. In the next section, results of the heuristic for the stepwise and linear model are compared. Additionally, we present results of the heuristic for the linear model for large real-life network sizes.



(a) $S = 3$, $W = 2$, $D = 15$, $K = 5$,
 $\varepsilon_c \leq 1.01$, $\varepsilon_L = 1.65$.

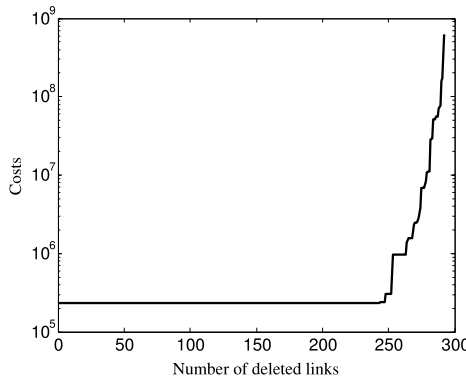


(b) $S = 4$, $W = 3$, $D = 20$, $K = 5$,
 $\varepsilon_c \leq 1.01$, $\varepsilon_L = 1.57$.

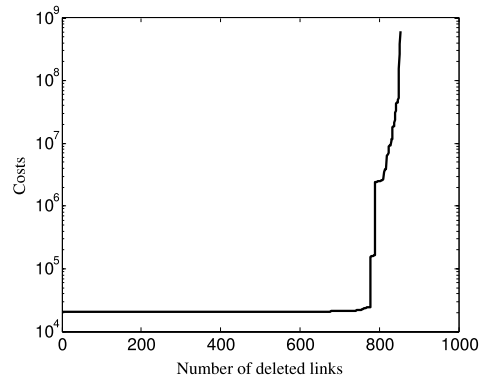
Fig. 3. Comparison heuristics results for the linear and stepwise model, $T=100$, $\tau_w=1$, $\Omega=3$. The cost start to increase at about (a: 165%, b: 157%) of the links of a minimally connected network.

3.2. Performance of the heuristic for large networks

Results for both the linear and the stepwise model are presented in Fig. 3 for two different network configurations. For larger network sizes, we have to rely on the heuristic model with linear costs. Two cost curves for large real-life networks are presented in Fig. 4. Costs are plotted on a logarithmic scale versus the number of deleted links. In these plots, zero deleted links corresponds to the fully connected network, whereas the maximal number of deleted links corresponds to a minimally connected network.



(a) $S = 5$, $W = 3$, $D = 40$, $K = 6$,
 $\varepsilon_c \leq 1.01$, $\varepsilon_L = 2.35$.



(b) $S = 8$, $W = 4$, $D = 75$, $K = 10$,
 $\varepsilon_c \leq 1.01$, $\varepsilon_L = 1.97$.

Fig. 4. Heuristics results for the linear model, $T = 100$, $\tau_w = 1$, $\Omega = 3$. The cost start to increase at about (a: 235%, b: 197%) of a minimally connected network.

All these curves (in Figs. 3 and 4) exhibit the same typical characteristics from the left to the right: at first, a large number of links can be deleted without affecting the costs significantly: Up to 70% of the links are not used at all in the first heuristic iteration and are thus deleted at once. Still, about 15% of all the links can subsequently be deleted without substantially affecting costs. Then, after about 85% of the total number of links that can be deleted, costs start increasing slightly with the number of deleted links. Finally, when about 90–95% of the total number of links that can be deleted are actually deleted, costs start increasing dramatically. Below the figures the ε_L -value are given for an upper bound on the ε_c -value of 1.01, which means that the heuristic is stopped if deleting one more link would lead to operational costs that are more than 1% higher than the operational costs in the fully connected network.

The comparison between the results of the linear and stepwise model show that in the fully connected network, as expected, the costs for the stepwise model are slightly higher than the costs for the linear model (due to the logarithmic scale this can hardly be noticed in Fig. 3). Furthermore, Fig. 3 shows that for both models the dramatic cost increase starts at approximately the same number of links. Further analysis (not depicted in the figures here) shows that for the performed experiments, the set of deleted links at the critical value is approximately the same for both models. A large number of experiments, not presented in this paper, exhibit the same characteristics, suggesting that the heuristic for the linear model is a quite good approximation of the heuristic with the stepwise model.

Apparently, a topology with only a small number of links suffices to gain close to minimal operational costs. Since only a small number of links remain we can also conclude that the resulting topology is also close to optimal. In [9], similar characteristics are presented for a manufacturing system with deterministic supplies and demands. The results in [9] suggest that a small flexibility (each plant produces only a couple of product types) can almost achieve the benefits of total flexibility (each plant produces all product types). Furthermore, it should be noticed that the curves in Figs. 3 and 4 exhibit approximately horizontal plateaus. The reason for this is that not one, but a group of links has to be deleted to significantly affect the costs.

3.3. Statistical robustness analysis

We now formulate the hypothesis that networks generated for a particular supply and demand time series are insensitive to changes of second order moments of supply and demand distributions. Thus, if information about the individual means of supplies and demands is available *a priori*, we only need to generate a stochastic time series with these means to find a cost-effective topology for other scenarios with these means. An additional experiment suggests that these stochastic time series, rather than deterministic ones, are required to generate a robust topology. Furthermore, we perform an experiment suggesting that the heuristics applied to the linear model finds a topology, which is efficient with respect to the economies of

scale principle (consolidation of product flows to induce full truckload shippings). Finally, as we optimize over a finite period of time T , we have to address the influence of the initial inventory. We demonstrate that the steady state value of the costs in the fully connected network converges to the minimal transportation plus storage costs, as initial inventory increases.

3.3.1. Sensitivity

To allow for an automatic statistical sensitivity analysis we define a representative “heuristic network” to be the smallest network \mathcal{U} determined by the heuristic, whose operational costs are within a factor F of the operational costs of the fully connected network, i.e. where $\varepsilon_c(\mathcal{U}) \leq F$. Deleting one more link will increase the costs above that level. We then perform the following statistical analysis: We choose fixed mean supplies and demands for a network with $S = 4$, $W = 3$, $D = 20$, and $K = 5$ and generate stochastic supply and demand samples (for 100 time steps). Based on this, we determine the “heuristic network” and fix its topology, i.e. its link structure \mathcal{U} . For this network \mathcal{U} , we then compute the operational costs for supply and demand time series from twenty different distributions (with different second order moments but same means) and record the resulting approximation factor $\varepsilon_c(\mathcal{U})$ in each case. Finally, the mean values $\bar{\varepsilon}_c$ and $\bar{\varepsilon}_L$ of the twenty instances are determined.

The above described experiment is repeated for 99 other heuristic networks, each generated from particular means of supply and demand. Figure 5 shows four scatter plots (each for a particular value of F) of the hundred mean values $\bar{\varepsilon}_c$ versus $\bar{\varepsilon}_L$. As expected, the robustness decreases as F increases (since then the number of links in the heuristic network decreases which makes the network more sensitive). We define the network to be robust if the mean value of ε_c (over 20 experiments) is at most 1.2 with a frequency of 95%. The scatter plot in Fig. 5(a) shows that ε_c is strongly clustered near one indicating that, with high likelihood, our process of choosing a heuristic network (with $F = 1.001$) will find a network that is a “very good” network even when the supply and demand time series vary randomly with respect to second order moments. The histogram for the case where $F = 1.01$ [Figure 5(b)] is still satisfactory since more than 95% of the considered cases results in an ε_c value equal to or smaller than 1.2. The results for $F = 1.05$ [Figure 5(c)] and $F = 1.1$ [Figure 5(d)] do not satisfy our definition of a robust network anymore. For these values of F , several means and standard deviations of ε_c are even outside the figure range (larger than 3). We choose as “heuristic network” the one with $F = 1.01$, since (i) the accompanying results fulfill our definition of a robust network, and (ii) the generated heuristic networks have fewer links than the case where $F = 1.001$. These results suggest that the “heuristic network” ($F = 1.01$) indeed is robust to changes in second order moments of supply and demand distributions.

The corresponding values of ε_L suggest that the generated heuristic networks on average consist of 2.5 times the number of links in a minimally connected network.

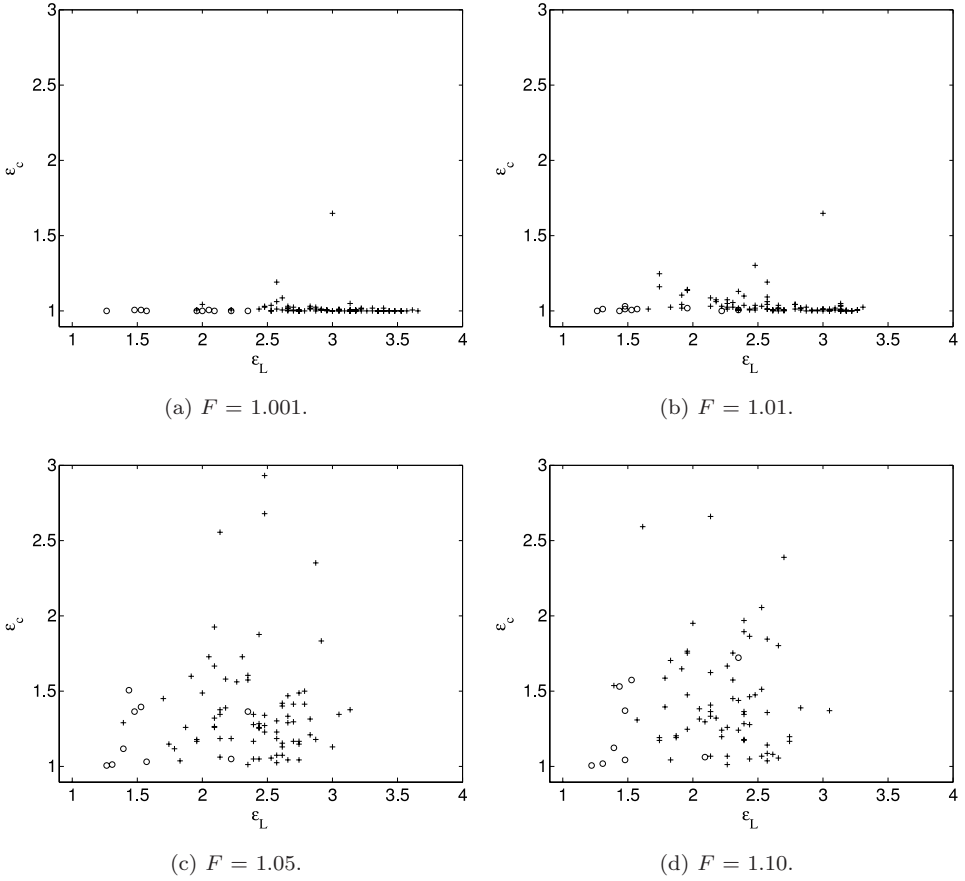


Fig. 5. Sensitivity of "heuristic network" to 2nd order moments. "o" represents a product mix with 1 high value and 4 low value products, "+" product mix with more than 1 high value product.

This means that each network node is on average connected to two or three other network facilities. Apparently, this limited amount of connectivity suffices to provide a large portion of the operational efficiency and robustness of the fully connected network. Another interesting result is that the number of links in the "heuristic network" depends on the product mix. A circle in Fig. 5 represents an instance with a product mix with only one high value product and in this case four low value products, whereas a plus represents an instance with a product mix with at least two high value products. The results suggest that more links are required when more high value products are present. We give the following explanation for this phenomenon: the structure of the "heuristic network" is mostly determined by the number and locations of suppliers and consumers of high value products. When more high value products are present, in general the number of suppliers and consumers of high value products increases, and thus more links have to be established to provide just-in-time delivery.

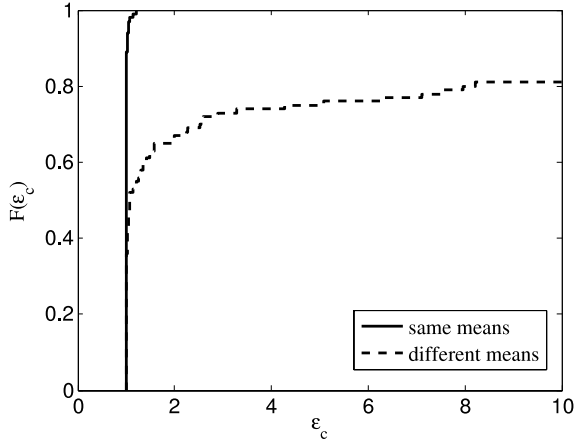


Fig. 6. Empirical cumulative distributions showing the sensitivity of the heuristic networks to changes in (i) variances, and (ii) means of supplies and demands.

To check the sensitivity of the “heuristic network” to the first moments of supply and demand distributions, we perform the following experiment: we take the hundred heuristic networks, where $F = 1.01$, and hundred accompanying results of ε_c from the previous experiments where the means of supplies and demands stay the same, but the variances change [Figure 5(b)]. In addition, for each of the 100 heuristics networks we generate supply and demand time series from newly generated distributions and determine the corresponding ε_c . Figure 6 depicts two empirical cumulative distribution functions of hundred ε_c values generated from (i) time series with different variances, but the same means as the “heuristic network” originally was generated from (solid line), and (ii) time series with different means (dotted line). The solid line again confirms that the “heuristic network” is robust to the variances since the cumulative distribution grows to 100% very fast starting at $\varepsilon_c = 1$. The cumulative distribution function for the instances with different means grows much slower. This is not surprising since the networks have been optimized to deliver the right amount at the right time for an ensemble mean. Changing the mean should change the topology of the optimal network.

3.3.2. Deterministic versus stochastic approach

As the results from the previous subsection suggest, only the supply and demand averages are required to construct a robust topology. This however does not imply that solving the deterministic problem based on these averages results in a robust topology. On the contrary, (random) stochastic supplies and demands are required to create back-up links that are not generated in the deterministic approach. This is illustrated by the following experiment. We consider 3 suppliers, 2 warehouses and 15 customers in a distribution network with 5 product types, i.e. $S = 3$, $W = 2$, $D = 15$, and $K = 5$. We first randomly generate $3 \cdot 5$ supply and $15 \cdot 5$ demand averages.

Now we construct network topologies in the following two ways: the heuristic applied to (i) the deterministic case based on these averages, and (ii) stochastic supplies and demands (with these same averages and additionally randomly generated second order moments). In both cases $\varepsilon_L = 1.01$ for arriving at a topology. Next, 31 new supply and demand time series (of 100 time steps each) with the same averages but different second order moments are generated. The performance of both found topologies for the new time series (relative to the performance of the fully connected network for the new time series) can now be compared by means of the value of ε_c . Figure 7(a) depicts the results of 31 topologies ($F = 1.01$) each subject to 31 time series. Results are expressed by values $(\varepsilon_L, \varepsilon_c)$. In this figure we observe that the (extreme) values of ε_c for the deterministic approach (circles) are much larger than the ones for the stochastic approach (stars). Apparently, the approach based on stochastic supplies and demands leads to much more robust network topologies than the approach based on deterministic supplies and demands.

3.3.3. Performance of the heuristic using linear transportation costs

Since the heuristic using the stepwise model can only solve medium-size instances, we have to rely on the heuristic using linear transportation costs (a cost model without an economies of scale principle). However, we have implemented a kind of a consolidation aspect in our heuristic instead: the utilization of a link is weighed according to the type(s) of product flowing through it. The higher the value of product k , the higher the weigh factor β^k . In this way, we prevent links that transport high-valued product(s) from being deleted and induce links with low-valued products to be deleted. The low-valued products will eventually be added to the

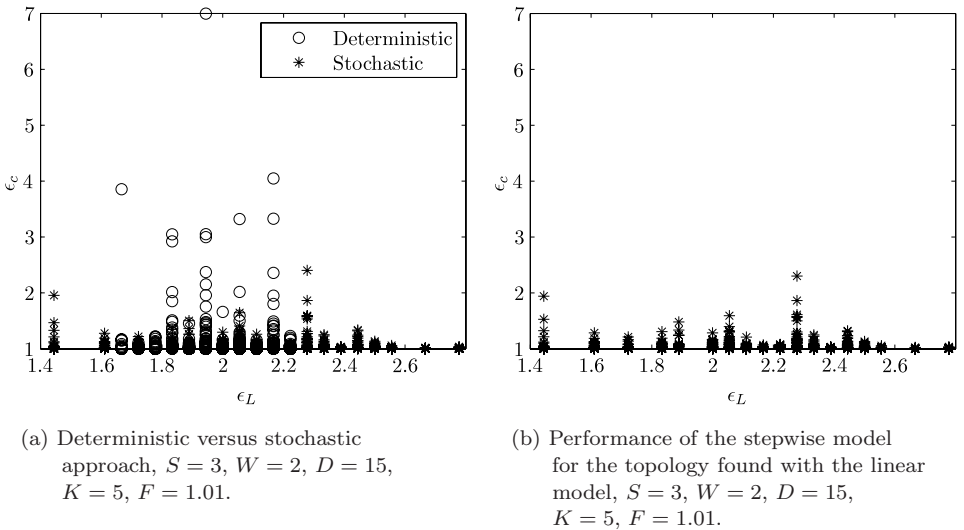


Fig. 7. Validation of the heuristics performance.

links with high value products. Although we realize that this consolidation effect is different from the one in the stepwise model, and as a consequence the daily shipping decisions may be different for both approaches, the constructed network topologies turn out to be very similar. Since this is exactly what we are looking for, namely a good network topology, the heuristic with the linear model can be used for solving large instances, e.g. $S = 8$, $W = 4$, $D = 75$, having $2^{932} \approx 3.63 \cdot 10^{280}$ possible topologies.

The following experiment is developed and performed: We again consider 3 suppliers, 2 warehouses and 15 customers in a distribution network with 5 product types, i.e. $S = 3$, $W = 2$, $D = 15$, and $K = 5$, and use the found 31 heuristic topologies (using linear costs) from the experiment in Sec. 3.3.3. Then, for each of these topologies, we run the stepwise model for (i) the found topology, and (ii) the fully connected network for thirty-one supply and demand time series (the same as in the first experiment) with the original averages, but different second order moments, and compare both costs. Corresponding values of $(\varepsilon_L, \varepsilon_c)$ are shown in Fig. 7(b). We see that the values of ε_c are still satisfactorily small (major part below 1.2) and closely resemble the values for ε_c (stars) in Fig. 7(b). Apparently, running the heuristic using the linear cost model results in a topology, which from an economies of scale perspective can efficiently be used.

3.3.4. Inventory flexibility

In all previously performed experiments the parameter of initial inventory \mathcal{Y}_w^k has been set to zero. It turns out that this initial inventory is the only controllable parameter that can influence the operational costs for running the fully connected network. Consider the black dotted lines in Figs. 8(a) and 8(b), which show results of the heuristic for $S = 4$, $W = 3$, $D = 20$, and $K = 5$, for respectively two different supply and demand time series with equal first moments and $y_w^k(0) = 0$. The fact

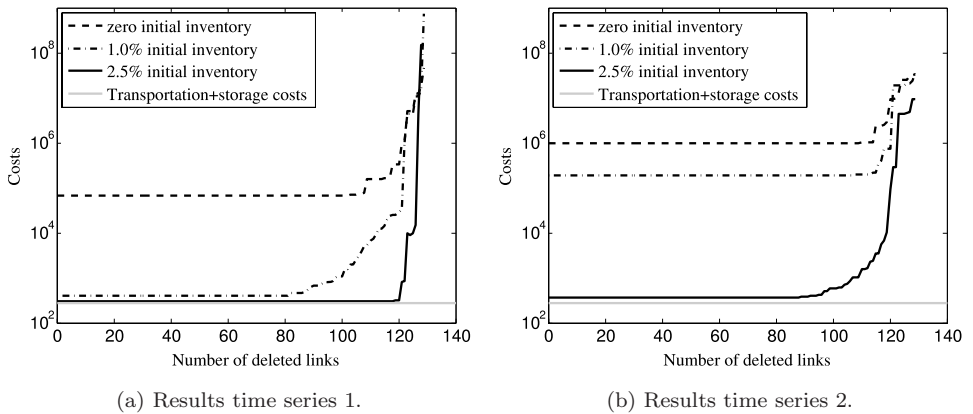


Fig. 8. Initial inventory flexibility. Performance of the network for different values of the initial inventory \mathcal{Y}_w^k .

that the means of supply and demand are equal implies that approximately the same amount of products is shipped along time period T , and thus approximately the same transportation costs are present for both these cases. Nevertheless, a large difference in costs for the fully connected network can be noticed. This difference can therefore only be caused by backlog costs, which are induced by the specific time series. For instance, a time series, which exhibits a peak in demand at the beginning of period T , always leads to backlog costs, even in a fully connected network. Therefore, we pose the hypothesis that the only controllable parameter that influences the costs for the fully connected network is the initial inventory. Increasing the initial inventory will compensate for early backlog, which eventually will lead to convergence to the inevitable transportation costs.

To confirm this hypothesis, we perform the following additional experiments for the two supply and demand time series as mentioned above: we generate a supply and demand time series, set the initial inventory in each of the warehouses to zero and determine the operational costs dependent on the number of deleted links [black dotted lines in Figs. 8(a) and 8(b)]. Next, we perform two additional optimizations for the same supply and demand time series where 1.0% and 2.5% of the total shipped amount of products is present in the warehouses as an initial condition, respectively. The results of these experiments are depicted in Figs. 8(a) and 8(b). As one can see, for both cases, the total costs of the fully connected network indeed converge to the transportation and storage costs as initial inventory increases.

4. Conclusions and Recommendations

In this paper we considered transportation networks from the point of view of a third party logistics provider, who deals with the problem of distributing different types of products from suppliers to consumers via transportation links. Warehouses between the suppliers and consumers may be used to compensate for the stochastic behavior of supplies and demands and to consolidate different products. The amounts of these supplies and demands are assumed to be uncontrollable for the logistics provider. With only information about the supplies and demands a few days in advance, the logistics provider has to decide on which transportation links to use for a long period of time (tactical level) and how much of which products to ship through them each day (operational level).

We formulated a bi-level joint network design and network operation problem: At the upper (tactical) level, the network topology has to be constructed, and at the lower (operational) level routings and schedules for daily shipments have to be decided on. A model predictive control with a rolling horizon (MPC) was used as decision model on this operational level. A heuristic was proposed to construct the topology dependent on the operational decisions and compared to a bi-objective branch and bound method. The consolidation of product types so as to leverage on the economies of scale principle is taken into consideration in this network design heuristic.

Experimental results reveal that the cost of such a near-optimally operated network as a function of the number of links in the network stays almost constant as the vast majority of network links are removed and explodes once the link number has decreased below a critical value. The experiments show that our heuristic determines a network that has close to optimal costs with a very low number of links (typically about 10% of all links that can be deleted before a minimally connected network remains).

Furthermore, experimental results suggested that the resulting topology is insensitive to second order moments of the individual supply and demand distributions. Hence, information about the means of supplies and demands over a certain time period suffices to generate a close to optimal network topology robust to any supply and demand scenario with the same means.

Acknowledgments

The work of E. Lefeber is supported by the Netherlands Organization for Scientific Research (NWO-VIDI grant 639.072.072). D. Armbruster was supported by a grant from the Stiftung Volkswagenwerk under the program on Complex Networks and by NSF grant DMS-0604986.

References

- [1] Bertazzi, L., Paletta, G. and Speranza, M. G., Deterministic order-up-to level in an inventory routing problem, *Transp. Sci.* **36**(1) (2002) 119–132.
- [2] Cheung, R. K. M. and Powell, W. B., Models and algorithms for distributions problems with uncertain demands, *Transp. Sci.* **30**(1) (1996) 43–59.
- [3] Cordeau, J. F., Laporte, G. and Pasin, F., An iterated local search heuristic for the logistics network design problem with single assignment, *Int. J. Prod. Eco.* **113**(2) (2008).
- [4] Daskin, M. S., *Networks and Discrete Location: Models, Algorithms, and Applications* (John Wiley and Sons, Inc. 1995).
- [5] Garcia, C. E., Prett, D. M. and Morari, M., Model predictive control: Theory and practice, *Automatica* **25**(3) (1989) 335–348.
- [6] Hax, A. C. and Candea, D., *Production and Inventory Management* (Prentice Hall, Inc., Englewood-Cliffs, 1984).
- [7] Hendriks, M. P. M., Multi-step optimization of logistics networks. strategic, tactical, and operational decisions. Ph. D. thesis, Eindhoven University of Technology (2009).
- [8] Jaillet, P., Song, G. and Yu, G., Airline network design and hub location problems, *Locat. Sci.* **4**(3) (1996) 195–212.
- [9] Jordan, W. C. and Graves, S. C., Principles on the benefits of manufacturing process flexibility, *Manage. Sci.* **41**(4) (1995) 577–594.
- [10] Kwon, S., Park, K., Lee, C., Kim, S. S., Kim, H. J. and Liang, Z., Supply chain network design and transshipment hub location for third party logistics providers, in, *Int. Conf. Computational Science and Its Applications ICCSA'06*, Glasgow, 8–11 May 2006, Lect. Notes Comput. Sci., Vol. 3982 (Springer, Berlin/Heidelberg, 2006), pp. 928–933.
- [11] Leung, J. M. Y., Magnanti, T. L. and Singhal, V., Routing in point-to-point delivery systems, *Transp. Sci.* **24** (1990) 245–261.

- [12] Meepetchdee, Y. and Shah, N., Logistical network design with robustness and complexity considerations, *Int. J. Phys. Distrib.* **37**(3) (2007) 201–222.
- [13] Simchi-Levi, D., Chen, X. and Bramel, J., *The Logic of Logistics*, 2nd edn. (Springer Science+Business Media, Inc., New York, 2005).
- [14] Topaloglu, H., An approximate dynamic programming approach for a product distribution problem, *IIE Trans.* **37** (2005) 697–710.
- [15] Tsiakis, P., Shah, N. and Pantelides, C. C., Design of multi-echelon supply chain networks under demand uncertainty, *Ind. Eng. Chem. Res.* **40** (2001) 3585–3604.
- [16] Wasner, M. and Zapfel, G., An integrated multi-depot hub-location vehicle routing model for network planning of parcel service, *Int. J. Prod. Econ.* **90** (2004) 403–419.
- [17] Zapfel, G. and Wasner, M., Planning and optimization of hub-and-spoke transportation networks of cooperative third-party logistics providers, *Int. J. Prod. Econ.* **78** (2002) 207–220.