ARTICLE IN PRESS
JID: EOR [m5G;February 4, 2022;1:11]

European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Production, Manufacturing, Transportation and Logistics

Operational planning in service control towers – heuristics and case study

B. Gerrits*, E. Topan, M. C. van der Heijden

University of Twente, P.O. Box 217, Enschede 7500 AE, the Netherlands

ARTICLE INFO

Article history:
Received 15 February 2021
Accepted 14 January 2022
Available online xxx

Keywords:
Supply chain
Inventory
Spare parts
Operational planning
Case study

ABSTRACT

We study performance improvement in multi-echelon, closed loop spare part supply chains using operational interventions based on real-time status information. Our objective is to minimize the total cost relevant costs, consisting of intervention costs and the backorder costs. In this paper, we focus on proactive interventions, aiming to avoid stockouts. We assume that all reactive interventions are fixed. Proactive interventions that we study include lateral transshipments, emergency shipments, stock reservations, expediting part repairs, and early new buys of parts. These interventions are invoked by using alert generation, when the supply chain status deviates from the plan. We propose heuristic rules to generate alerts. We also develop heuristic rules for the choice of operational interventions. We model and test our heuristics in a simulation test bed, based on data of a global IT company by using the case data in Germany. Numerical experiments reveal the following key insights: (i) downstream interventions – proactive lateral and emergency shipments – have most impact in reducing costs, (ii) communicating losses in the supply chain (no returns, failed repairs) for early new buys has positive impact on fill rates at negligible costs, and (iii) expedite repair and stock reservations using the proposed rules is not profitable.

1. Introduction and contribution to literature

Capital goods are advanced technical systems that are crucial for delivering goods and services. Examples are airplanes, medical equipment, and industrial computers. Their availability is very important for the users. When such an asset is not available, this may result in severe revenue loss and customer dissatisfaction. To provide after-sales services and maintenance to users, service providers of capital goods typically operate complex spare parts supply networks, and through service level agreements (SLAs), they offer detailed service packages to ensure high uptime for users. Spare part supply networks differ from manufacturing supply chains in many aspects, see Cohen, Agrawal and Agrawal (2006). Relevant for this research are the inclusion of return, repair and disposal of failed components, low demand rates, and fast response requirements. Also, the network should be capable of SLAs that may vary between customers (Kutanoglu & Mahajan, 2009; Tiemessen, Fleischmann, van Houtum, van Nunen & Pratsini, 2013), with varying prices for the corresponding contracts to provide differentiated services.

It is a challenge to minimize the costs of running the spare parts network while attaining the various SLAs. To achieve this, planners make decisions at strategic level (e.g., where to locate warehouses), tactical level (e.g., the stock levels per part and per location), as well as day-to-day operational planning activities. At the operational level, planners must respond to deviations from tactical plans when monitoring the supply chain, e.g., a part request cannot be fulfilled from the nearest warehouse, or the current stock on hand falls below a threshold. In such a case, planners consider interventions when they face an acute issue (e.g., urgent demand at a warehouse that is out of stock), or to anticipate on a high stockout risk in the short run. The first category are *reactive* interventions, the second category are *proactive* interventions. Examples of proactive interventions are (i) initiating an emergency shipment from upstream the supply chain to a warehouse nearby customer sites, (ii) relocating a part between stock locations at the same level in the supply chain (i.e., a transshipment between local warehouses), and (iii) expediting return and repair of a failed part.

To support planners making such decisions, a recent trend is to use a *service control tower* (SCT). According to Accenture (2015), an SCT "acts as a centralized hub that uses real-time data from a company's existing, integrated data management and transactional systems to integrate processes and tools across the end-to-end supply service chain and drives business outcomes". In a typical SCT,

* Corresponding author.
E-mail address: b.gerrits@utwente.nl (B. Gerrits).

https://doi.org/10.1016/j.ejor.2022.01.025

the supply chain status is closely monitored by collecting a lot of real-time data, and alerts (triggers or exception messages) are generated in case of possible supply chain anomalies. Planners often analyze such anomalies manually, but decision support could be offered by presenting multiple intervention options with their estimated impact. The amount of real-time data facilitates operational planning is a key aspect that differentiates SCTs from more traditional centralized inventory control systems focusing on tactical decisions that have been studied extensively in the past decades (see the review of de Kok et al., 2018).

In recent work, Topan, Eruguz, Ma, van der Heijden and Dekker (2020a) reviewed the literature on operational planning in service control towers, and confronted this literature with practice to identify challenges and needs in after-sales service logistics. Some of the findings in practice include:

- Operational decisions in spare part supply chain are seldom executed automatically. Thresholds for alerts are often somewhat arbitrary. Often, too many alerts are generated, such that planners do not have time to check all these alerts.
- Despite the fact that companies typically offer a range of service contracts with various SLAs, literature on customer differentiation in operational planning is scarce.

We observe these issues in a case study at a global IT company for hardware solutions that is headquartered in Europe. The forward supply chain of the company consists of a central warehouse and several local warehouses. Failed parts are returned and repaired in the reverse supply chain, after which they are suitable for reuse in the forward part of the supply chain. Together, we have a closed loop spare part supply chain, where new buys compensate losses due to parts that are not returned or repaired.

Customers have different SLAs, namely same-day deliveries (premium customers) or next-day deliveries (nonpremium customers). Nonpremium customers are served from the central warehouse, and premium customers can be served from multiple (but not all) local warehouses within the required time window. Consequently, there are subsets of same-day customers that overlap and cannot be easily decomposed over local warehouses.

Inspired by the case study, we develop methods for alert generation and heuristics to find good proactive interventions that improve SLA fulfillment and reduce the need for expensive reactive interventions. Moreover, reactive interventions are rather trivial in this setting. When a same-day delivery cannot be satisfied from a nearby local warehouse, it is satisfied from the central warehouse via an emergency shipment. If the central warehouse is out of stock, it waits until new material arrives. Yet proactive interventions to avoid future stockouts remain as an issue. Therefore, we focus on proactive interventions aiming to improve supply chain performance. We develop a discrete-event simulation model with real-life data from the case study as testbed for the decision rules that we will develop.

The contribution of our paper is as follows:

- We develop heuristic decision rules for operational planning, i.e., interventions for a multi-echelon, closed loop spare part supply chain that do <u>not</u> allow for a decomposition over local warehouses.
- We include customer differentiation, i.e., we have two different customer service levels, expressed as time windows within which spare parts should be delivered at the customer site.
- We use our model and solution methods to study the impact of generating alerts and interventions in a real-life case study for the hardware solutions of a multinational IT company. We develop several insights including the trade-off between the number of alerts being generated for interventions and the short-term performance of the supply chain.

The remainder of this paper is structured as follows. In Section 2 we review the related literature. We provide a detailed explanation of the case study that motivated our research in Section 3. Section 4 deals with interventions and alert generation that are applicable to the case study. We introduce our mathematical model and notation in Section 5. In Section 6, we develop various decision rules for interventions. We evaluate the decision rules in a simulation study based on the case and provide key insights in Section 7. Finally, we present our conclusions and directions for further research in Section 8.

2. Related literature

Although strategic and tactical planning of spare parts is widely discussed in the literature, there is a limited number of papers on operational planning. Our paper contributes to the latter stream of research. For a recent review of literature regarding operational interventions in service control towers, we refer to Topan et al. (2020a). Our paper is also related to the papers that addresses some of the interventions in our paper individually, e.g., expediting (e.g., Song & Zipkin, 2009), lateral transshipments (Paterson, Kiesmüller, Teunter & Glazebrook, 2011). Yet, our paper is different from these papers because we consider multiple interventions. We further differentiate our paper form vast majority of papers by selecting the interventions among many alternatives based on real time information (instead of assuming fixed decisions and assessing the impact of these decisions in the long term), and furthermore by considering a multi-echelon setting. The papers which consider multiple interventions in a multi-echelon setting is summarized in Table 1.

Hoadley and Heyman (1977) proposes a mathematical model to determine emergency shipments, lateral transshipments, return allocation, and characterize the properties of the cost function. Fisher (1989) considers priority scheduling of repairs and cannibalization, and proposes different rules and policies and tests them in a simulation. Pyke (1990) and Abell, Miller, Neumann and Payne (1992) have a similar approach, but they also include lateral transshipments. Caggiano, Muckstadt and Rappold (2006) consider repair capacity, stock allocation and emergency shipments from the central warehouse. They propose an MILP formulation and an effective heuristic to find the optimal interventions. Grahovac and Chakravarty (2001) consider lateral transshipments and emergency shipments only. Meissner and Senicheva (2018) apply approximate dynamic programming to find lateral transshipment decisions in a two-echelon system for retail products with short selling seasons, negligible lead times and lost sales. Howard, Marklund, Tan and Reijnen (2015) investigate lateral transshipments, emergency shipments, and proactive and reactive stock allocation interventions, using queuing theory. Topan and van der Heijden (2020b) consider a similar setting and propose an MILP formulation to include both event based and periodic interventions. They find that joint interventions reduce total downtime considerably, and that proactive emergency shipments contribute the most.

Our work differs from these papers in the following ways: (1) we explore a broader range of proactive operational interventions, i.e., using pipeline stock, emergency shipments from inventory upstream, lateral transshipments, as well as reallocation of returns, expediting repairs, and reservation at the upstream location. (2) We consider a large closed-loop supply chain including repair and return processes, where other papers typically focus on a part of the supply chain only (typically a part of the forward supply chain). (3) We combine alert generation (triggers) and interventions. (4) We cover losses in the supply chain due to part return and repair. (5) The central warehouse deals with different types of demand: direct demand (next day and same day emergency)

Table 1
Summary of papers on operational planning with multiple interventions in a multi-echelon setting.

Related Papers	Intervention											Intervention type		Information use		Review type			Problem/model setting						Use of Case Data
	Emergency shipments from the central warehouse	Lateral transshipments	Pipeline stock	Forward stock allocation	Return stock allocation	Capacity allocation	Cannibalization	Priority scheduling repairs	Expediting repairs	Backorder clearing	Stock reservation at the central warehouse	Proactive	Reactive	Return yield	Verification yield	Continuous / Event based	Continuous / Trigger based	Planned, Periodic review	Number of items	Number of echelons	Planning horizon	Demand	Supply	Tactical parameters (base stock levels)	
Bertrand and Bookbinder (1998)		x			x							x				x			Single	Multi	Finite	General distribution	Deterministic replenishment, zero lateral	Base stock levels fixed	
Caggiano et al. (2006)	x			x		x						x						x	Multi	Multi	Finite (rolling)	General distribution	Deterministic	Not mentioned	
Fisher (1989)			x				x	x				x	x			x			Multi	Multi	Infinite	Poisson	Exponential repairs	Base stock levels fixed	US Navy
Grahovac & Chakravarty (2001)	x	x	x									x	x			x			Single	Multi	Infinite	Poisson	Deterministic	Base stock levels also optimized	
Hoadley & Heyman (1977)	x	x			x							x	x					x	Single	Multi	Finite	General distribution	Zero lead time	Not mentioned	
Howard et al. (2015)	x	x	x	x						x			x			x			Single	Multi	Infinite	Poisson	Deterministic	Base stock levels also optimized	Volvo Parts
Pyke (1990)		x		x			x	x				x	x			x			Multi	Multi	Finite (rolling)	General distribution	Exponential distribution	Base stock levels fixed	
Abell et al. (1992)		x		x			x	x		x		x	x					x	Multi	Multi	Finite (single)	General distribution	Deterministic	Base stock levels fixed	
Topan and van der Heijden (2020)	x	x	x	x						x		x	x			x		x	Multi	Multi	Finite (rolling)	General distribution	Deterministic	Base stock levels fixed	Global semiconductor manufacturer
Our paper	x	x	x		x				x		x	x		x	x		x		Multi	Multi	Finite (rolling)	Poisson	Deterministic	Base stock levels fixed	Global IT manufacturer

and replenishment orders of local warehouses. (6) Finally, our paper differs from most of these papers by testing our approach in a real life case with actual data, whereas most of the papers in this field are theoretical contributions.

There are several papers on service differentiation (Kranenburg & van Houtum, 2008), customer differentiation and stock reservation at the central warehouse (Axsäter, Olsson & Tydesjö, 2007), and expediting (Arts, Basten & Van Houtum, 2016). However, these papers focus on tactical planning (e.g., finding the optimal target or base stock levels). We differ from these papers by focusing on operational planning (e.g., in our paper base stock levels are fixed). There are methods based on decomposing the complex multi-echelon problem into simpler single-echelon problems. However, these methods are applicable to tactical planning, e.g., finding optimal base stock levels and threshold levels (e.g., Howard et al., 2015, Kranenburg & van Houtum, 2009), and not to operational planning problems with a finite horizon. In a recent paper, Topan et al. (2020b) develop operational interventions for spare part supply chains and test their method to a different case than ours. Compared to this paper, we also include (i) the return and verification process (ii) losses in the supply chain due to missing returns and unsuccessful repairs with various yields, (iii) a new buy channel to compensate for these losses, (iii) customer differentiation (different service levels for premium and nonpremium demand), (iv) the impact of alert generation.

3. Case study

A large manufacturer of IT-hardware, middleware and software provides after sales services to its customers to support their daily operations. To reduce downtime due to spare part unavailability, the company has a global supply chain for its spare parts with local warehouses close to the customer sites. In this section, we introduce the current supply chain, spare parts flows, and operational interventions at the manufacturer. We use the case study to build our model.

3.1. Current supply chain structure

Fig. 1 shows the product flows In our case study. We focus on the part of the supply chain that serves customers of the manufacturer in Germany. Spare parts are stored in 11 *local warehouses*, which are replenished from a central warehouse in The Netherlands. There are ten *customer groups*, each of which is an aggregated set of individual customers.

3.2. Forward part flows

Each customer group consists of *premium* and *nonpremium* demand. *Nonpremium* demand is satisfied from the central warehouse, as customers can always be delivered in time (next day) from there, provided that the part is on stock. In this way, the manufacturer benefits from pooling demand uncertainty over customer groups, but also differentiates its customer service for the two customer groups.

Premium demand is primarily satisfied from the nearest local warehouse, the so-called *primary local warehouse*. If that local warehouse is out of stock, the part is supplied from a set of alternative local warehouses in a predetermined fixed sequence, insofar they can supply parts in time (same day). If this is not feasible, demand is satisfied the next day using an emergency shipment from the central warehouse. These interventions to meet demand are called *reactive* interventions.

Fig. 2 shows the local warehouses that can supply same-day deliveries to each customer group (yet not the predetermined fixed sequence of local warehouses, and for this, we keep a fixed sequence of local warehouses for each customer). Local warehouses can deliver the same day to between one and four customer groups (e.g., primary local warehouse 7 can provide same-day deliveries to customer groups 6, 7, 8 and 10). This complex structure with overlapping subsets of local warehouses for different customer groups with same-day deliveries complicates the analysis. Therefore, a simple analysis based on decomposition over local warehouses is not feasible.

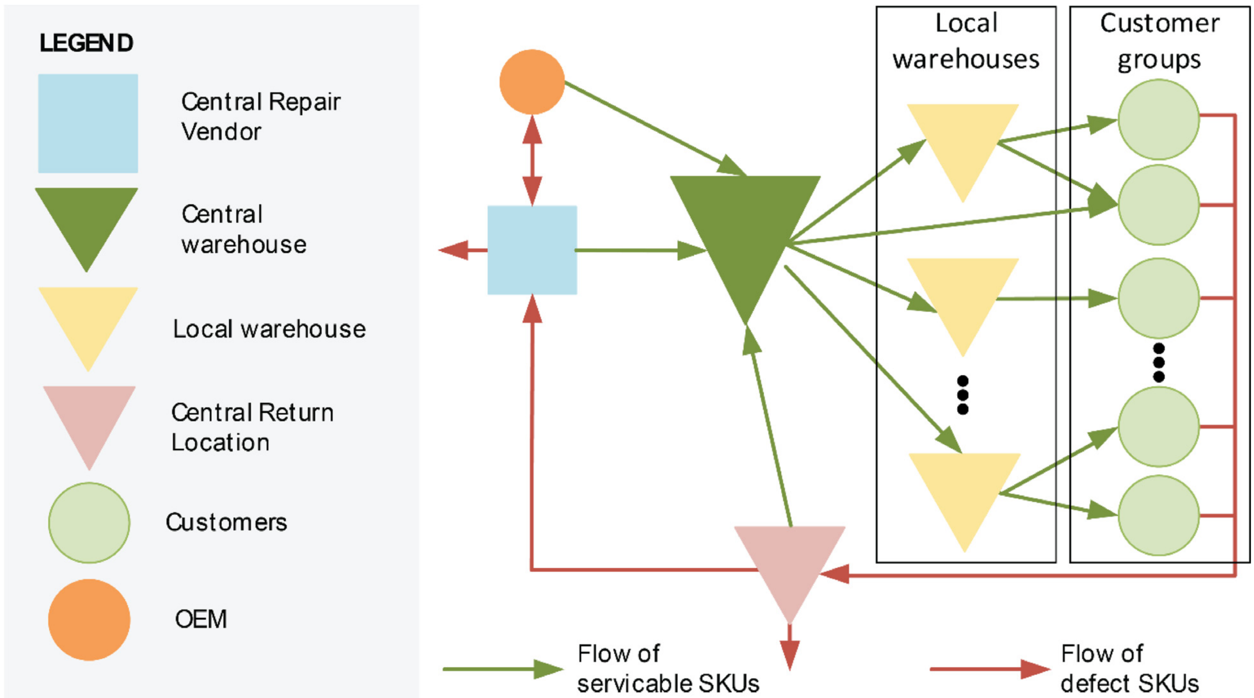

Fig. 1. Supply chain in case study (SKU = Stock Keeping Unit; OEM = Original Equipment Manufacturer).

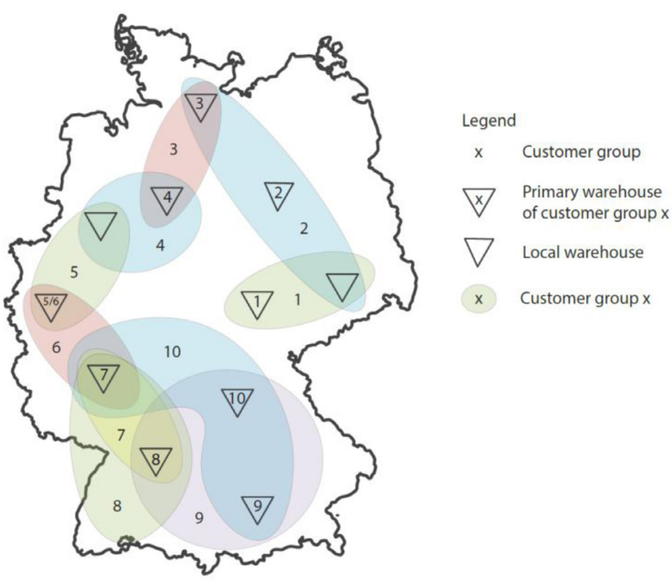

Fig. 2. Customer groups and local warehouses in Germany.

3.3. Reverse part flows

Demand for a spare part is always associated with a part failure at the corresponding customer location. All parts are in principle technically and economically repairable. Failed parts are sent once per week to the Central Return Location (CRL, cf. Fig. 1) for further processing, inspection and dispatching. Some failed parts may contain privacy sensitive data and thus a fraction is not returned to the CRL, resulting in a part-specific *return yield*. Remaining parts are inspected and classified as either *repairable*, *non-repairable* or *no-fault-found* (i.e., serviceable). No-fault founds are sent periodically (batched) to the central warehouse. Non-repairable parts are scrapped at the CRL. Repairable parts are sent twice per week to the Central Repair Vendor (CRV, cf. Fig. 1), where they are consolidated and verified. The verification process checks for possible reasons why repair should not be executed (e.g., expired warranty or a new part successor). Thus, a fraction is not forwarded for repair, resulting in a part-specific *verification yield*. Parts that pass the verification process are forwarded to the Original Equipment Manufacturer (OEM) for repair. At the OEM, parts are further inspected by experts. Then repair may turn out to be infeasible, resulting in a part-specific *repair yield*. Those parts that are infeasible to repair are returned to the CRV to be scrapped. The remaining parts are repaired at the OEM. After repair, they are returned to the central warehouse as *serviceable stock*. These stocks are used to meet non-premium demand, replenish local warehouses, and to satisfy premium demand unmet by any local warehouse. This all occurs on a First Come, First Serve (FCFS) basis.

3.4. Inventory control

At a tactical level, the inventories are controlled as follows. All the local warehouses use a periodic review, order-up-to level policy (cf. Silver, Pyke & Thomas, 2017) with a review period of two days. The order-up-to levels are determined such that the on-time delivery percentage is approximately 95% for each SKU. As we have discrete demand with low rates, we cannot attain exactly 95% for every SKU. We found the order-up-to levels using some preliminary simulation experiments proceeding from an initial setting based on a basic periodic review inventory model.

The central warehouse uses a continuous review reorder point, (s, S) installation stock policy. That is, the central warehouse places a replenishment order at the OEM (a so-called *new buy*) when its installation inventory position drops to or below the reorder point s. Then the size of the replenishment order raises the installation inventory position to the order-up-to level S. The installation inventory position of the central warehouse is defined as the on-hand stock minus the backorders from the local warehouses and nonpremium customers, plus the new buy parts in the pipeline from the OEM, plus the repaired parts in the pipeline between CRV and central warehouse, plus the pipeline stock of no-fault founds between the CRL and the central warehouse. The reorder point at the central warehouse is chosen such, that the on-time delivery percentage is about 95% for each SKU. Similar to the order-up-to levels of the local warehouses, we accomplished this using preliminary simulation experiments.

3.5. Operational interventions

Currently, the company does not apply proactive interventions. The reactive interventions in place have been described in Section 3.2.

4. Improvement potential: interventions and alerts

In this section, we focus on *proactive* interventions to avoid stockout risks. Based on the case study, we identify two types of proactive interventions: proactive interventions for premium demand, and those that are mainly made for nonpremium demand.

4.1. Proactive interventions for premium demand

Same-day deliveries involve a set of local warehouses that can fulfill same-day requests. We consider the following interventions:

- A *proactive emergency shipment* from central warehouse to a local warehouse connected to customer groups facing a high stockout risk. In this way, we expedite replenishment of inventories at local warehouses.
- A *proactive lateral transshipment* from one local warehouse to another local warehouse. In this way, we rebalance inventories at local warehouses.
- *Reserving parts at the central warehouse for premium demand*. Typically, a stockout for premium customers (i.e., same day delivery) is considerably worse than a stockout for nonpremium customers (i.e., next day delivery). In that case, it may make sense to prioritize premium demand (if local warehouses are out of stock) and replenishment orders from local warehouses over nonpremium demand, when the central warehouse is low on inventory. Of course, this will have negative impact on the nonpremium service level, so we have to investigate whether this makes sense. Note that currently these two demand streams are handled FCFS.

4.2. Proactive interventions that are mainly for non-premium demand

For nonpremium demand, we need to look at interventions that decrease the stockout risk at the central warehouse. This also influences stockout risks of premium requests, because the central warehouse faces replenishment orders, as well as emergency shipments from the local warehouses for premium demand.

- *Expediting the repair process* (i.e., shortening the lead time). Repairs of failed parts at the OEM can be expedited by giving that repair job priority. Expediting further upstream in the supply chain, e.g., expediting returns by selecting a faster delivery mode is less interesting. Previous research has shown that the impact of shortening lead times decreases when we move away from downstream locations serving customer demand (Van der Heijden, Alvarez & Schutten, 2013).
- *Ordering new buys using real-time return and repair information*. In the spare parts supply chain that we consider, losses occur because not all parts are returned or repaired. In the current situation, information on losses is not immediately taken into account for new buy decisions at the central warehouse. An alternative is to integrate information about the installation inventory position of the central warehouse and losses to make the replenishment decisions.
- *Reserving parts at the central warehouse for direct demand*. This is the similar to the third intervention for premium demand. Fulfilling nonpremium demand or emergency requests for premium demand, is in the short run more important than replenishments of local warehouses. The replenishment orders may have lower priority as they do not immediately affect service levels for premium demand. Likely, at most one of the two stock reservation policies that we introduce may add value depending on the parameter settings, but not both.

4.3. Alert generation

Planners are triggered to monitor the supply chain and to decide upon an operational intervention if needed. These alerts are typically activated when a performance indicator exceeds a prespecified threshold. The main challenge is to select a good threshold: an excessive number of alerts (system nervousness) may arise from loose thresholds, whereas it can be too late for a proactive intervention when thresholds are tight. Also, planners are able to handle a limited number of exception messages only, so generating many messages cause planners to make their own selection of messages they will handle, and that choice is not necessarily the best one and it depends on the experience and skills of the planner.

The alerts are related to the interventions under consideration. For most interventions, an alert is invoked if the inventories in the forward part of the supply chain (at central warehouse or the local warehouses) are low. We use the following trigger mechanism: *the Probability of stockout (POS) at a warehouse exceeds a certain threshold*. The POS is defined as the probability that demand exceeds the on-hand stock level during the remainder of the replenishment lead-time (i.e., until the next part is scheduled to arrive).

The second intervention in Section 4.2, ordering new buys using real-time return and repair information, requires a different trigger. When we observe less returns and repairs than average, we can use this to justify buying additional parts, hence advancing new buys. In that way, inventory shortages at the central warehouse can be prevented. To use this, we set a threshold for the gap between the expected and the actual number of retuned and repaired parts in the last period. In Section 6 we explain how we incorporate decision rules for the interventions. In Section 7.4 we also address the impact of alert generation.

5. Notation and model assumptions

Although we will test our decision rules in a simulation of a multi-item system, we use a single item model: the multi-item problem can simply be decomposed over parts. We will measure the performance in terms of fill rates for premium and nonpremium customers. As these are long-term measures, we base our operational decision rules on backorder costs that in the long-run lead to the fill rates targeted.

We consider a two-echelon supply chain, consisting of a set J of warehouses. We use index $j=0$ for the central warehouse, and $j=1, 2, \ldots, |J|-1$ for the local warehouses. Demand arises from a set I of customer groups, indexed as $i \in \{1, \ldots, |I|\}$. Each customer group may generate demand in two priority classes: *premium demand* (in our case: same-day), denoted by $m=1$, that needs to be satisfied within fixed time window T_1, and *nonpremium* demand (in our case: next-day), denoted by $m=2$, that needs to be satisfied within time window $T_2>T_1$. The demand of customer group i and priority class m follows a Poisson distribution with mean λ_{im}

Table 2
Overview of key notation for decision rules.

Symbol	Description
Supply chain related	
I	Set of customer groups ($i = 1, \ldots, \lvert I \rvert$)
J	Set of warehouses ($j = 0, \ldots, \lvert J \rvert - 1$)
m	Demand class ordered in decreasing priority ($m = 1, 2$)
T_m	Time window within which demand class m has to be filled ($m = 1, 2$)
λ_{im}	Mean demand of customer group i for class m
λ_P	Total mean premium demand (demand class 1) over all customer groups
λ_{NP}	Total mean nonpremium demand (demand class 2) over all customer groups
λ_{TOT}	Total mean demand over all customer groups and demand classes
j_i^*	Primary warehouse j of customer group i
H_i	Set of secondary warehouses of customer group i, where $H_i \subset J \setminus \{0, j_i^*\}$
Lead-times	
L_{LW}	Shipment lead-time for replenishments from central warehouse to a local warehouse
$L_{CRL,i}$	Return lead-time for replaced parts from customer group i to CRL
L_{CRV}	Shipment lead-time from CRL to CRV, including verification time
L_{OEM}	Shipment lead-time of unserviceable parts from CRV to OEM
L_{rep}	Repair lead-time at the OEM
L_{R_CRV}	Shipment lead-time of repaired parts from OEM to CRV
L_{CW}	Shipment lead-time from CRV to central warehouse
L_{NB}	New buy lead time
Yields	
γ	Fraction of failed parts that are returned to the CRL (return yield)
δ	Fraction of returned parts that arrive serviceable at the CRL ("no fault found")
ε	Fraction of returned parts that arrive failed but repairable at the CRL
φ	Verification yield at the CRV
ω	Repair yield at the CRV
Supply chain status when an alert is generated and an intervention needs to be considered	
OH_j	On-hand stock at warehouse j, with OH_0 representing the central warehouse stock
EIP_{LW}	Aggregate echelon inventory position at the local warehouse level
$P_{1,j}$	Expected time until the next part arrives at local warehouse j (equal to $L_{LW\,j}$ if no part is in the pipeline between central warehouse and local warehouse j)
P_2	Expected time until the next part arrives at the central warehouse from either the new buy or repair stream. If both pipelines are empty, this is equal to L_{NB}.
Cost factors	
C_E	Costs of pro-active emergency shipment from central to local warehouse
C_L	Cost of a pro-active lateral transshipment between two local warehouses
C_{ERn}	Cost of expediting repair at stage n.
C_{B1}	Backorder costs for premium demand per item, no matter the duration of the delay
C_{B2}	Backorder costs for nonpremium demand per item, no matter the duration of the delay
Performance	
$EBO_m(s, t)$	Expected backorders for class m demand over period t if we have s parts available
$F_{Pois}(s\|\lambda)$	*Probability that a Poisson distributed random variable with mean λ is $\leq s$*

per period. Every customer group i has a *primary* local warehouse j_i^*, and when it is out of stock we may source from a fixed set $H_i \subset J \setminus \{0, j_i^*\}$ of *secondary* local warehouses from which premium demand can still be satisfied within time window T_1. We refer to Table 2 for an overview of notation.

We use the following *model assumptions:*

1. All (return, repair and shipment) lead times are constant.
2. The lead times from central warehouse to local warehouses are identical for all local warehouses.
3. All demand has to be satisfied (i.e., we have backordering and not lost sales).
4. The nonpremium demand is satisfied from the central warehouse, and the premium demand is satisfied from local warehouses when there is sufficient on hand stock.
5. Local warehouses use (S-1, S) inventory control policies with order-up-to-levels not higher than one.
6. The central warehouse applies an installation stock (s, S) policy.

In the supply chain, we have the following lead times. Upon failure, the defective SKU is removed and is returned to a Central Return Location (CRL) for inspection. Returns are made periodically (in our case: once per week) with lead-time $L_{CRL,i}$ for customer group i. Some parts are not returned, resulting in a return yield γ. The parts that are returned turn out to be (i) properly functioning (serviceable parts), a fraction δ, (ii) defective but repairable, a fraction ε, (iii) defective and not repairable, a fraction $1 - \delta - \varepsilon$. Repairable parts are sent to the central repair vendor (CRV) periodically, and after consolidation and verification to the OEM for repair. Recall that all fractions are part-specific, but that we omit the part index for simplicity of notation.

The lead time of a shipment from the CRL to the CRV, including verification is denoted by L_{CRV}. A fraction does not pass verification and is scrapped at the CRV, resulting in a verification yield φ. The remaining parts are sent to the OEM for repair with lead-time L_{OEM}. After a repair lead-time L_{rep} at the OEM, the part is returned to the CRV with lead-time L_{R_CRV}. Some parts that arrive at the CRV after inspection at the OEM are faulty and these are scrapped, resulting in a repair yield ω. The repaired parts are sent to the central warehouse (CW) with lead-time L_{CW}, where parts are kept on stock or used to clear backorders. We denote the shipment lead time from the central warehouse to a local warehouse by L_{LW}. When a new part needs to be bought from the OEM, we face a new buy lead-time denoted by L_{NB}. Note that we need all these lead times separately and cannot aggregate them, since the interventions that we consider – such as expediting - depend on the stage in the closed loop supply chain. Also, the information available depends on that stage: we have various yields at various points in the reverse supply chain. Thus the uncertainty in supply decreases when moving further in the supply chain, when more stages with each their own lead time have been passed.

Furthermore, we have intervention costs. For premium demand, we denote the costs associated with an emergency shipment from the central warehouse by C_E and a lateral transshipment between any local warehouses by C_L. The costs associated with nonpremium demand interventions are denoted by C_{ER} for expediting repair. Using real-time return information and stock reservation for nonpremium demand do not have costs associated with them. The backorder costs are denoted by C_{B1} for premium demand and C_{B2} for nonpremium demand. To describe the status of the supply chain at a certain point in time, we use the following notation. We define $P_{1,j}$ as the lead-time until the next part arrives at the local warehouse (i.e., a maximum of the regular replenishment lead-time if nothing is in the pipeline). P_2 is defined as the lead-time until the next part arrives at the central warehouse (from either the new-buy or the repair stream). If both pipelines are empty for a certain SKU k, it holds that $P_2 = L_{NB}$ (i.e., the new buy lead-time). We measure the status of the local warehouses by the aggregate echelon inventory position at the local warehouse level EIP_{LW}, defined as the sum of all local warehouse stocks plus all parts in the pipeline between the central warehouse and all local warehouses, minus premium backorders (if any).

6. Operational planning

This section provides heuristic rules for proactive interventions for premium (Section 6.1) and nonpremium demand (Section 6.2).

6.1. Interventions for premium demand

Whenever there is an alert, the model allows choosing between multiple interventions. As discussed in Section 4, we include as pro-active interventions: (i) an emergency shipment from the central warehouse, (ii) a lateral transshipment from a local warehouse, and (iii) reserving parts at the central warehouse for same-day demand. The heuristic rule for the first two interventions is developed in a single model (6.1.1). For the third intervention we propose a separate heuristic rule in 6.1.2.

6.1.1. Proactive emergency and lateral transshipments

At any point in time, the on hand stock at local warehouse j equals OH_j, and the time until the next replenishment (order size one) arrives equals $P_{1,j}$. If there is no replenishment order in the pipeline, we take $P_{1,j} = L_{LW}$, as this is the earliest time for new material to arrive at the local warehouse. Our decision rules try to minimize the total expected costs during $P_{1,j}$, consisting of (i) intervention costs, and (ii) expected stockout costs. As we discussed, a decomposition over local warehouses is not feasible (cf. Section 1 and 3). Therefore, estimating the expected demand per local warehouse – and thus the expected backorder costs – during $P_{1,j}$ is not straightforward. We therefore deploy two heuristics to estimate the expected backorders: a look-ahead approach and a Markov chain-based approach. The look-ahead approach ignores demand fulfillment by reactive lateral transshipments and redirection of demand among local warehouses during stockout. The Markov chain-based approach does incorporate demand fulfillment by reactive lateral transshipments and redirection of demand. However, it includes additional assumptions and limitations (see below), such as a maximum order-up-to level of one. The look-ahead approach can be applied in a wider range of settings, including larger order-up-to levels.

Recall that we use the following logic for reactive interventions for both heuristics: if a local warehouse is out of stock, we check a subset of other local warehouses in a fixed sequence until we found a local warehouse that has a part on hand. If local warehouses who can fulfill demand within one day are out of stock, we use and emergency shipment from the central warehouse.

Below we discuss these two heuristics to estimate expected backorders.

Method 1: look-ahead approach

We focus on a single local warehouse, and therefore we suppress the local warehouse index j. We calculate the expected backorders as follows. Given that we have s parts available to satisfy premium demand in the next period with length t, we compute the expected premium backorders by:

$$EBO_P(s,t) = \sum_{n=s+1}^{\infty} (n-s)\frac{(\lambda_P t)^n e^{-\lambda_P t}}{n!} \tag{1}$$

We can rewrite this as

$$EBO_P(s,t) = \lambda_P t * \{1 - F_{Pois}(s-1|\lambda_P t)\} - s * \{1 - F_{Pois}(s|\lambda_P t)\} \tag{2}$$

where we define the cumulative Poisson distribution as $F_{Pois}(s|\lambda_P t) = 0$ if $s < 0$.

We set $t = L_{LW}$ if no replenishment order is in the pipeline, and $t = P_1$ otherwise. As a simple decomposition method is not feasible for our case study, we estimate the demand faced by a local warehouse solely with the original demand rates dedicated to this warehouse.

Method 2: Markov chain-based approach

We use a discrete time Markov chain to evaluate the impact of proactive interventions decisions considering reactive interventions that can take later at all local warehouses over a period L_{LW}. Here, we evaluate particularly the expected total number of backorders $EBO_P(s, L_{LW})$ and costs associated with the interventions. Additional to the main model assumptions, we further assume:

a) Premium demand never exceeds one per period in the entire network.
b) We neglect other proactive shipments decisions during evaluation period L_{LW} at the local warehouses.
c) The central warehouse has ample stock.

Assumption (a) is the consequence of a negligible probability of having more than one unit of demand at each period at any local warehouse. Note that this is justified also because order-up-to-levels are not higher than one in our setting. Assumption (b) is needed to model our evaluation using a Markov chain. Note that otherwise we would need a Markov decision model to incorporate proactive decision that can be made later. The transitions represent the inventory status changes due to demand and the reactive interventions. We include the proactive intervention in the first period that we aim to evaluate, and ignore any further proactive interventions over L_{LW}.

We define the starting state vector $(x_1, \ldots, x_J)$, where x_j denotes the number of discrete time units until the next replenishment order of local warehouse j will arrive. To exemplify, $x_1 = 2$ denotes that the replenishment order of warehouse 1 is due in 2 days. Also, $x_1 = 0$ denotes that the part is already on stock and there is no replenishment order in the pipeline. This means that we can only use warehouse j for (direct or lateral) demand fulfillment when state $x_j = 0$. We illustrate the Markov chain for a primary warehouse j where we distinguish between $x_j = 0$ and $x_j > 0$. In both cases, demand with size 1 may occur at warehouse j in the next period, or no demand occurs. This leads to the following transitions:

- A demand arrives with probability $1 - e^{-\sum_{i \in I} \lambda_{i1}}$. As demand per period is at most one by assumption, we face customer class i demand with probability $\frac{\lambda_{i1}}{\sum_{i \in I} \lambda_{i1}}$.
  - The demand is satisfied from stock if the warehouse has positive stock, e.g., $x_{j_i^*} = 0$. Considering that all pipeline stocks will advance 1 period (day) forward and a new replenishment order is placed for warehouse j, the transition is from $(x_1, \ldots, x_J)$ to $(x_1 - 1, \ldots, x_{j_i^*} = L_{LW}, \ldots, x_J - 1)$.

  - If $x_{j_i^*} > 0$, the demand at warehouse j is satisfied from the first secondary local warehouse in H_i (so according to the "precedence list") that has positive stock. Let this warehouse be j'. Given the advancement of pipeline stock and a replenishment order for warehouse j, the transition is from $(x_1, \ldots, x_J)$ to $(\max(x_1 - 1, 0), \ldots, \max(x_{j_i^*} - 1, 0), \ldots, x_{j'} = L_{LW}, \ldots, \max(x_J - 1, 0)$.
  - If none of the secondary local warehouses H_i have positive stock, demand is satisfied from the central warehouse by a reactive emergency shipment. Then, the transition is from $(x_1, \ldots, x_J)$ to $(\max(x_1 - 1, 0), \ldots, \max(x_J - 1, 0))$. This means a stockout for customer group i since demand is not satisfied the same-day.
- No demand arrives with a probability $e^{-\sum_{i \in I} \lambda_{i1}}$. Then, the transition is from $(x_1, \ldots, x_J)$ to $(\max(x_1 - 1, 0), \ldots, \max(x_J - 1, 0))$.

We calculate $EBO_{L_{LW}}$ by using n-step transitions of the Markov chain where $n = 1, \ldots, L_{LW}$. From assumption (a), the probability that a demand occurrence at a local warehouse without stock ($x_j > 0$) at any discrete time period gives the expected number of backorders at the end of each discrete time period. Since we are interested in the total backorders over L_{LW} periods, we have L_{LW} discrete time periods, and thus, L_{LW} such probabilities. To calculate each, we calculate probabilities that a demand occurrence at a warehouse without stock at the end of day n using n-step transitions for $n = 1, \ldots, L_{LW}$. Their sum gives $EBO_P(s, L_{LW})$. To evaluate each intervention, we calculate $EBO_P(s, L_{LW})$ and costs associated at all local warehouses over a period L_{LW} of starting with an initial state that the intervention will lead to. An intervention is selected if the decrease in expected costs exceeds the intervention costs.

6.1.2. Reserving parts at the central warehouse for premium demand

If backorder costs for premium demand are considerably higher than those for nonpremium demand, it makes sense to prioritize shipments for premium demand, i.e., replenishment orders by the local warehouses and emergency shipments. When a nonpremium customer arrives, we check the supply chain status and estimate the expected shortage costs until the next replenishment arrives at the central warehouse to select among the two options: to reject the nonpremium demand or to fill the nonpremium demand. We reject nonpremium demand if we find a (significant) reduction in total shortage costs.

The key status information here consists of (i) the expected time until the next part(s) arrive at the central warehouse P_2, (ii) the on-hand stock level at the central warehouse OH_0, and (iii) the total number of parts available in the downstream (at or on transport to the local warehouses, minus premium backorders) just before the nonpremium customer arrives, represented by the aggregate echelon inventory position EIP_{LW}.

If we block a nonpremium customer, we have one nonpremium backorder for sure. We also have one part extra to satisfy premium demand in the period until the next part(s) in the pipeline of the central warehouse arrive at the local warehouses, so during $P_2 + L_{LW}$. We estimate the premium backorder reduction by aggregating demand and inventories over all local warehouses. That is, we assume that any part available at the downstream can be used to satisfy premium demand at every local warehouse. Obviously, this assumes we use the parts at the downstream perfectly. This assumption will be closer to reality if we apply proactive lateral and emergency shipments as discussed in 6.1.1.

Blocking the nonpremium customer implies that we have OH_0+EIP_{LW} parts available to satisfy premium demand during P_2+L_{LW}, whereas we have OH_0+EIP_{LW}-1 parts available if we do not block. When we plug this in (2), the impact of blocking the nonpremium customer on premium backorders equals

$$\Delta EBO_P = EBO_P(OH_0 + EIP_{LW} - 1, \lambda_P(P_2 + L_{LW})) - EBO_P(OH_0 + EIP_{LW}, \lambda_P(P_2 + L_{LW})) \quad (3)$$

This can be rewritten as (see also equation (2.7) in Van Houtum & Kranenburg, 2015),

$$\Delta EBO_P = 1 - F_{Pois}(OH_0 + EIP_{LW} - 1 | \lambda_P(P_2 + L_{LW})) \quad (4)$$

The nonpremium customer is blocked if the expected cost reduction in premium backorders exceeds the increase in backorder cost of nonpremium customers due to the rejection, i.e.,

$$C_{B1} \Delta EBO_P > C_{B2} \quad (5)$$

6.2. Interventions for nonpremium demand

Interventions for nonpremium demand focus on the central warehouse. Recall that the central warehouse (i) satisfies nonpremium demand, (ii) replenishes local warehouses, (iii) fills emergency requests from the local warehouses, (iv) receives repaired parts and (v) orders new buys from the OEM. As discussed in Section 4, we include the following pro-active interventions: (i) expediting the repair process (6.2.1), (ii) ordering new buys using real-time information on losses in the return and repair process (6.2.2), and (iii) reserving parts for direct demand (6.2.3).

6.2.1. Expediting the repair process

We may intervene at three stages during the repair process: (stage 1) before repair has started, we prioritize the repair job, thereby reducing waiting time before the start and during the execution of repair, (stage 2) after repair and still at the OEM, using an expedited shipment from the OEM to the CRV, and (stage 3) after repair and at the CRV, using an expedited shipment the CRV to the central warehouse. We can reduce the lead time in stage n to a fraction κ_n of the original lead time ($0 < \kappa_n < 1$). We may reduce lead times in multiple stages. If we deploy expediting at all three stages, we reduce the total return lead-time to $\kappa_1 L_{rep} + \kappa_2 L_{R\_CRV} + \kappa_3 L_{CW}$. If the part is at stage 2 of the repair process, the expedited repair lead-time is $\kappa_2 L_{R\_CRV} + \kappa_3 L_{CW}$ and if the part is at stage 3, the expedited repair lead-time is $\kappa_3 L_{CW}$. Expediting makes most sense for the part that is scheduled to arrive at the central warehouse earliest.

We propose the following decision rule. Let the next part that is expected to arrive at the central warehouse be in repair stage n. Then, we can shorten the expected remaining lead time (either from new buy or from repair) from P_2 to $P_{e,n}$, $n = 1,\ 2,\ 3$. Here the input parameter $P_{e,n}$ denotes the expedited remaining repair lead time in stage n. In the numerical experiments (Section 7), we will set this equal to a fraction of the original remaining repair lead time P_2. Note that expediting repair does not have any impact in the short run if a new buy order arrives before expedited repair. Also, we note that an early arrival of part at the central warehouse means less nonpremium backorders and less delay in fulfilling (replenishment orders of) premium demand. As it is difficult to distinguish the impact on nonpremium and premium backorders and the majority is nonpremium demand, we use only nonpremium backorder costs to estimate the impact of expediting. Analogously to Section 6.1.2, we find expected backorders at the central warehouse over time t by:

$$EBO_{CW}(s, t) = \lambda_{TOT} t * \{1 - F_{Pois}(s - 1 | \lambda_{TOT} t)\} - s * \{1 - F_{Pois}(s | \lambda_{TOT} t)\} \quad (6)$$

We expedite repair if the marginal benefit of expediting exceeds the marginal intervention cost C_{ERn} in stage n, i.e.,

$$[EBO_{CW}(OH_0, P_2) - EBO_{CW}(OH_0, P_{e,n})] * \frac{C_{B2}}{C_{ERn}} > 1. \quad (7)$$

Table 3
Expected and actual number of parts leaving the supply chain.

Location →	CRL	CRV	OEM
Actual number of parts leaving the supply chain	Q_F - Q_R − Q_{NFF}	Q_R − Q_V	Q_V − Q_C
Expected number of parts leaving the supply chain	(1 - γ - ε) Q_F	(1 - φ)Q_R	(1 - ω)Q_V

With a similar logic, we could also develop decision rules for any further parts scheduled to arrive from the repair pipeline. Likely the marginal returns of these interventions will decrease with the number of repairs to expedite, so this only makes sense if our numerical experiments reveal that expediting the first arriving parts adds sufficient value (which will turn out to be not the case).

6.2.2. Ordering new buys using real-time return information

Recall that the installation inventory position at the central warehouse includes the on-hand stock level and the pipeline stock from both the new buy stream and the repair stream and is controlled using an (s, S) installation stock policy. Consequently, information on parts that leave the supply chain (i.e., because they are not returned or repaired) is delayed. Only a fraction of failed parts may arrive in the part of the repair stream that is visible for the central warehouse (i.e., between the CRV and the central warehouse). This reduces the installation inventory position and this may lead to placing a new buy order.

We can reduce the information delay by tracking the actual number of parts leaving the supply chain compared to the expected number of parts. For example, if more parts have left the supply chain than normal, the central warehouse may replenish its inventory earlier than prescribed by its (s, S) policy to compensate the parts left the system (and the other way around). Note that failed parts are transshipped through the supply chain in batches, where some parts in a batch may leave the supply chain earlier at various locations, resulting in a yield. Parts that stay within the supply chain always remain in the same batch. At the CRL (return), CRV (verification) and OEM (repair) the yield may be smaller than 1.

At the CRL batches of parts are returned from the installed base and are checked for repairability. For the CRL we use the following notation:

- Q_F = number of parts that have failed since the previous batch of failed parts has been returned from the installed base
- Q_R = number of parts identified as repairable at the CRL from a batch of returned parts (sent to CRV)
- Q_{NFF} = number of parts identified as “no fault found” at the CRL from a batch of returned parts (“serviceable” parts)

At the CRV, batches of repairable parts arrive from the CRL and are verified:

- Q_V = number of parts verified as repairable at the CRV that enter the repair process at the OEM

At the OEM, we track how many repairs are successful:

- Q_C = number of parts that have been repaired successfully

We track the difference between the actual and expected number of parts leaving the supply chain since the last replenishment order at the central warehouse, see Table 3. If the cumulative actual number minus the expected number exceeds some threshold Δ_{max} (trigger), we decide to order early. That is, we order $S - IP$ if the installation inventory position at the central warehouse IP falls at or below $s + \Delta_{max}$.

6.2.3. Reserving parts at the central warehouse for direct demand

We consider reserving stock for direct demand and backordering replenishment orders from local warehouses when the central warehouse stock is low. Direct demand consists of nonpremium demand and overflow of premium demand (i.e., reactive emergency shipments to meet premium demand that cannot be satisfied from local warehouses). Such a stock reservation can make sense, as delay in satisfying direct demand will immediately lead to penalty costs, whereas a delay in replenishment orders does not immediately lead to penalty costs. We decline a replenishment request arriving from a local warehouse whenever the decrease in total backorder costs for direct demand is higher than increase in the total backorder costs for premium demand. We use the following logic to determine the expected backorders for the different demands streams, which are then multiplied by their respective costs parameter.

Similar to Section 6.1.2, if we reject the replenishment request, we have EIP_{LW} parts available to satisfy premium demand until P_2+L_{LW}, and OH_0 parts to satisfy nonpremium demand and reactive emergency shipments for premium demand until P_2. If we accept the replenishment request, we have one part more for premium demand at the local warehouses (EIP_{LW}+1) and one part less for nonpremium demand and premium emergency shipments (OH_0 − 1). Similar to (3) and (4), we find that the impact of rejecting replenishment order on premium backorders is negative:

$$EBO_P(EIP_{LW}, P_2 + L_{LW}) - EBO_P(EIP_{LW} + 1, P_2 + L_{LW}) \\ = F_{Pois}(EIP_{LW} | \lambda_P (P_2 + L_{LW})) - 1 \tag{8}$$

The nonpremium backorders are affected over the interval [0, P_2]. We first approximate the shortage under the option “reject”. Then we can also (approximately) evaluate the “accept” option if we replacing EIP_{LW} by EIP_{LW}+1, and OH_0 by OH_0 −1.

The nonpremium backorders in the interval [0, P_2] depend on the overflow premium demand that needs to be satisfied from the central warehouse if the local warehouse stock is insufficient. Let T_{PS} be the time until the first shortage of premium demand occurs (and so a request for an emergency shipment is issued to the central warehouse), and $M = min\{T_{PS}, P_2\}$. In the interval [0, M], the central warehouse faces nonpremium demand only. In [M, P_2], the CW also faces emergency shipment requests to satisfy premium demand. T_{PS} equals the time until premium demand equals $EIP_{LW} + 1$. As demand is Poisson distributed, the time between demand events is exponentially distributed with parameter λ_p, and so T_{PS} has an Erlang distribution with parameters $EIP_{LW} + 1$ and λ_p, with mean $E[T_{PS}] = \frac{EIP_{LW}+1}{\lambda_P}$. The probability that $T_{PS} \geq P_2$ is the probability that premium demand until P_2 is not more than EIP_{LW}, so $F_{Pois}(EIP_{LW} | \lambda_P P_2)$. Some calculus reveals that

$$E[M] = \frac{EIP_{LW} + 1}{\lambda_P} \{1 - F_{Pois}(EIP_{LW} + 1 | \lambda_P P_2)\} + P_2 F_{Pois}(EIP_{LW} | \lambda_P P_2) \tag{9}$$

We use two approximations to simplify the calculations.

Approximation 1. *We replace M by its expectation E[M].* Then the nonpremium backorders in $[0, E[M]]$ are simply given by $EBO_{NP}(OH_0, E[M])$, where we use (2) for $EBO_{NP}(s, t)$ replacing λ_P by λ_{NP}.

For the nonpremium backorders in [E[M], P_2], we first approximate the total expected shortage over [0, P_2]. The expected total demand over [0, P_2] consists of nonpremium demand over the

Table 4
SKU-specific properties.

SKU k	Demand rates (per year)		Lead times (days)		Yields at CRV	
	Premium λ_P	Nonpremiumλ_{NP}	New buy $L_{NB,k}$	Repair$L_{rep,k}$	Verification φ_k	Repair ω_k
1	119	878	84	10	1.00	0.970
2	85	850	84	10	1.00	0.975
3	56	607	84	10	0.94	0.975
4	53	636	122	97	0.97	0.970
5	40	600	43	10	0.45	0.975
6	40	576	64	10	0.96	0.975
7	38	357	60	10	0.78	0.975
8	38	239	67	10	0.84	0.975
9	37	168	100	84	1.00	0.910
10	36	309	65	10	0.19	0.975

entire interval, plus emergency orders for premium demand over [E[M], P_2]), so $\lambda_{NP}P_2 + \lambda_P(P_2 - E[M])$.

Approximation 2. *The total demand is Poisson distributed with* *constant* *rate over [0,* P_2*].* Then we approximate the total backorders (nonpremium demand + premium emergency requests) by $EBO_{PNP}(OH_0, P_2)$ using (2), replacing λ_P by $\lambda_{NP} + \lambda_P(1 - E[M]/P_2)$. These are backorders over [0, P_2], so we have to deduct the expected backorders in [0, E[M]]. Then we find the nonpremium backorders in [E[M], P_2] as a fraction $\lambda_{NP}/(\lambda_P + \lambda_{NP})$ of the total expected backorders, so:

$$\frac{\lambda_{NP}}{(\lambda_P + \lambda_{NP})}\{EBO_{PNP}(OH_0, P_2) - EBO_{NP}(OH_0, E[M])\} \tag{10}$$

Then, the sum of expected nonpremium backorders in [0, E[M]] and in [E[M], P_2] equals

$$\frac{\lambda_{NP}}{(\lambda_P + \lambda_{NP})} * EBO_{PNP}(OH_0, P_2) + \frac{\lambda_P}{(\lambda_P + \lambda_{NP})} * EBO_{NP}(OH_0, E[M]) \tag{11}$$

7. Numerical results

Based on the case study and model described, we construct a discrete-event simulation model in Plant Simulation 15.2. We programmed the decision logic using the built-in language SimTalk. The simulation model is used to evaluate the effectiveness of proactive interventions for operational planning on the service level, expressed in fill rates. We first define the basic setting in Section 7.1, and next discuss the key results in Section 7.2. We discuss the impact of alert generation in Section 7.3.

7.1. Experimental design

We used the replication-deletion approach (cf. Law, 2014), where each replication corresponds to a five-year simulation of the closed loop supply chain. In each replication, we generate demand for each stock keeping unit from Poisson distribution with mean values as observed in our case study, as we observed one-by-one demand in all cases. For the yield at each stage we simply use a Bernoulli distribution with yield rates as observed in the case data. Welch's Method revealed that a warm-up period of one year is sufficient, leaving us with four years to collect results. We chose the number of replications such that the half-width of the 95% confidence interval of the fill rate for premium demand and for nonpremium demand is at most equal to 0.01. We found that 30 replications are sufficient in all instances. The average run time per instance is about 35 s on a Dell XPS17 9700, Intel Core i7–10875H, 16GB RAM.

As basis for our experiments, we use the case study as described in Section 3 for $K = 5$ representative SKUs. Table 4 summarizes the key characteristics of these SKUs. All failed parts are returned to the CRL and are repairable ($\gamma=\varepsilon=1$ and $\delta=0$). The shipment lead times for the return flows are $L_{CRL,i} = L_{CRV} = L_{OEM} = L_{R\_CRV} = L_{CW} = 5$ days. The regular replenishment lead-time between central warehouse and each local warehouse equals $L_{LW} = 5$ days. Emergency shipments from the central warehouse and proactive lateral transshipments both take one day. Expediting repair shortens the remaining lead-time by 50%. As key performance indicators, we use the fill rate per demand class (premium, nonpremium). We choose the backorder costs for premium and nonpremium customers as $C_{B1} = 20{,}000$ and $C_{B2}=2000$, irrespective the delay. In this way, we attain decent fill rates, as we will see in the numerical results.

To gain more insight, we also added 15 fictional SKUs. First, we used the ten SKUs from Table 4 and doubled the shipment lead times between central warehouse and local warehouses (SKU 11–20). Our hypothesis H1 is that interventions in the downstream part of the supply chain are more effective if these lead times are longer. Second, we used the first five SKUs with the highest demand rates and generated five more SKUs with double demand rates (SKU 21–25). Our hypothesis H2 is that we have more possibilities to intervene if we have more items in the various pipelines, and therefore may find more impact of operational interventions. Finally, we added experiments for the ten SKUS from Table 4 with 5% less stock at the central warehouse. Our hypothesis H3 is that stock reservations may have more impact if the central warehouse runs out of stock more frequently. We also performed experiments for SKU 21–25 with 5% less warehouse stock *and* with doubled shipment lead-times between the central warehouse and the local warehouses. These experiments show similar results compared to SKUs 1–10. We omitted the results in Table 5 for sake of conciseness, but we included the details in the appendix.

We consider six proactive intervention types as described in Section 4. We use a benchmark scenario to compare our scenarios against to identify which intervention type has the highest positive impact. The benchmark scenario is the one above without proactive interventions (reactive interventions only). Next, we combine interventions to assess which combination makes most sense.

7.2. Numerical results and insights

Table 5 shows they key results from the simulation study. The intervention costs are the average costs of all reactive and proactive interventions per year. We display the weighted average fill rate per demand class over the SKUs as mentioned in the header of column 3 and 4, weighted on the demand rate. We provide detailed results per SKU in the appendix (Table A1-A3). As we need order-up-to levels higher than one at a subset of local warehouses in experiment 22–26, we only apply the look-ahead heuristic for proactive lateral and emergency shipments and not the Markov-chain approach. For a fair comparison, we also show the weighted

Table 5
Key results of the simulation experiments. Black values show no significant change compared to the scenario without interventions (reactive only). Green (red) values show significant positive (negative) changes based on T-test with $\alpha = 0.05$.

Experiment	Proactive interventions	Weighted fill rate (SKU 1–10)		Intervention costs (yearly)
		Premium	Nonpremium	
1	None (reactive only)	95.2	94.4	€ 0
2	Lateral transshipments (look-ahead)	98.1	94.4	€ 340K
3	Lateral transshipments (Markov chain)	97.6	94.4	€ 161K
4	Emergency shipments	99.5	94.3	€ 330K
5	Stock reservation for premium demand	95.1	94.4	€ 1.5K
6	Expedite repair	95.2	94.4	€ 10K
7	Use return and repair losses for new buys	95.2	94.9	€ 1K
8	Stock reservation for direct demand	92.6	95.5	€ 4K
9	Combination of 2 and 5	98.0	94.4	€ 338K
10	Combination of 2 and 8	96.2	96.4	€ 332K
11	Combination of 2, 4, 6, 7 and 8	92.0	94.1	€ 654K
5% Less central warehouse stock				
12	None (reactive only)	94.6	91.2	€ 0
13	Emergency shipments	98.1	90.5	€ 277K
14	Lateral transshipments (look-ahead)	97.7	91.2	€ 313K
15	Stock reservation for direct demand	90.3	93.0	€ 95K
16	Stock reservation for premium demand	94.6	91.2	(€ 8 K)
Double shipment lead-times				
		Weighted fill-rate (SKU 11–20)		
17	None (reactive only)	95.2	95.3	€ 0
18	Emergency shipments	99.5	95.3	€ 900K
19	Lateral transshipments (look-ahead)	98.5	95.2	€ 276K
20	Stock reservation for direct demand	93.7	96.4	€ 50K
21	Stock reservation for premium demand	95.2	95.3	(€ 1 K)
Double demand rates				
		Weighted fill-rate (SKU 21–25)		
22	None (reactive only)	95.2 [95.5]	95.3 [95.1]	€ 0
23	Emergency shipments	98.8 [99.3]	94.3 [95.2]	€ 195K
24	Lateral transshipments (look-ahead)	98.2 [98.3]	95.3 [95.3]	€ 286K
25	Stock reservation for direct demand	93.5 [93.3]	96.5 [96.4]	€ 12K
26	Stock reservation for premium demand	95.2 [95.5]	95.3 [95.3]	(€ 1 K)

fill-rate of SKUs 1 to 5 in the base setting between square brackets for experiments 22 to 26. We conclude from Table 5:

1. Pro-active lateral transshipments and pro-active emergency shipments have the highest impact on premium demand fill rates at the cost of additional shipments.

Proactive emergency shipments yield a higher fill rate (i.e., 99.5%) than proactive lateral transshipments (i.e., 97.6%), whereas the number of emergency shipments is also much lower than the number of lateral transshipments. The key reason is that using proactive emergency shipments shorten regular replenishment lead times by a factor five proactive lateral transshipments do not have such an effect. Proactive emergency shipments are made when inventories are low only; yet they replace about 20–30% of the regular replenishments. A drawback of proactive lateral transshipments, despite having the same lead-time as emergency shipments, is that they may cause problems at other local warehouses, triggering a possible transshipment there and starting a ripple effect of lateral transshipments. For emergency shipments it holds that stock is brought from upstream, not causing any ripple effects at other local warehouses. As in our case, the central warehouse typically has ample stock, proactive emergency shipments do not negatively impact fill rates at other local warehouses or for nonpremium demand.

2. The look-ahead heuristic for pro-active lateral transshipments performs similarly to the Markov Chain approach at the expense of twice as many shipments.

We observe from Experiment 3 and 4 that the premium fill rates are slightly better for the look-ahead heuristic (i.e., 98.1%) than for the Markov Chain (MC) approach (i.e., 97.6%), but at the expense of 72% more transshipments. The MC approach includes the impact of a lateral transshipment on the entire network, whereas the look-ahead heuristic looks at the impact of a lateral transshipment only on that local warehouse. This is further discussed in Section 7.3. The drawback of the MC approach is that it takes considerably more computation time. That will not be a bottleneck for practical usage, but it is for lengthy simulation runs. Therefore, we combine other interventions with the look-ahead approach only.

3. Stock reservation for direct demand positively impacts nonpremium fill rates and negatively impacts premium fill rates.

As nonpremium demand and reactive emergency shipments of premium customers are prioritized at the central warehouse, regular local warehouse replenishment orders are delayed. As a consequence, the nonpremium fill rate increases and the premium fill rate decreases. This intervention seems counterproductive as the decrease in premium fill rate is larger than the increase in nonpremium fill rate. Further analysis shows that premium demand is satisfied 22% more often the next day via an reactive emergency shipment from the central warehouse. The negative effect is thus somewhat diminished, as the delay is only one day. Note that the decrease in premium fill rates can be counterbalanced by deploying proactive lateral transshipments additionally, keeping the nonpremium fill rate at 96.4%, whilst increasing the premium fill rate to 96.2% (see experiment 10 in Table 5).

4. Stock reservation for premium demand is not fruitful due to low demand rates and fast replenishment.

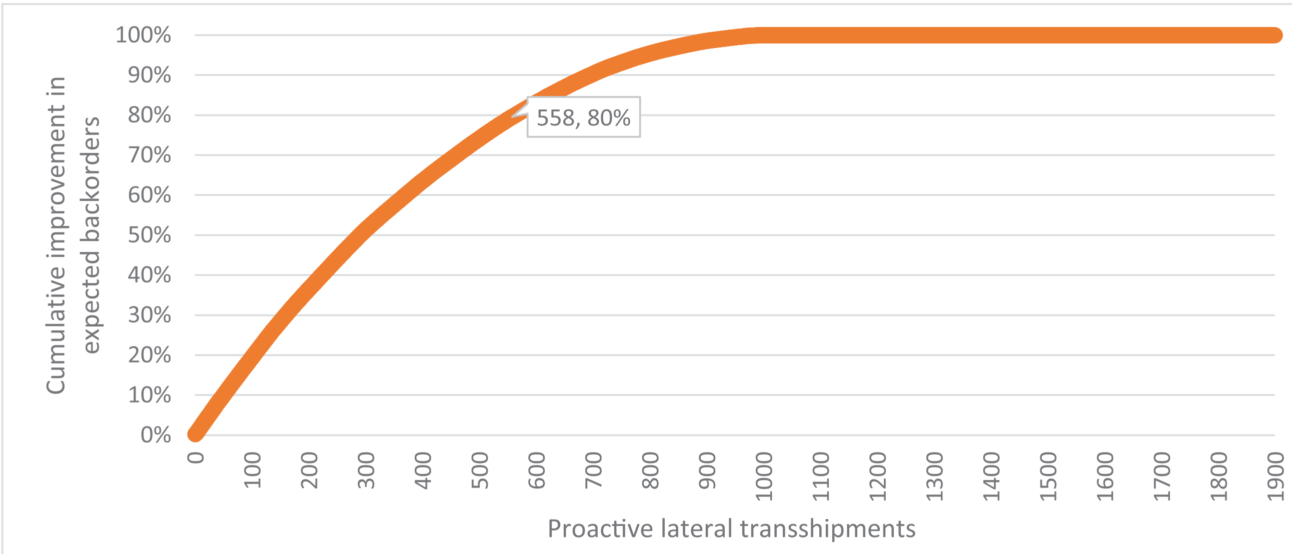

Fig. 3. The cumulative impact of proactive lateral transshipments.

Stock reservation for premium demand only occurs once or twice per year, therefore the impact on fill rates is negligible. In our case, premium demand rates are considerably lower than nonpremium demand rates, resulting in a low probability that a reserved part is actually used to fulfill premium demand. This may also be due to our approximations. First, we assume that all stock in the pipeline to or at the local warehouses can be used to satisfy all premium demand. This is generally not true, although a combination with proactive lateral transshipments may help. Second, we use a myopic rule focusing on the earliest arriving replenishment, either from the new buy stream or the repair stream. Given the yields and batching for new buys, the earliest arriving replenishment is most of the time a single part from the repair channel. This may be insufficient to solve all issues, and we may need a more sophisticated rule that includes more information further upstream the supply chain.

5. Including information on actual losses in the supply chain has a small positive impact on nonpremium fill rates at very little cost

This intervention only requires the registration of scrapped and non-repaired parts and information exchange with the central warehouse, so costs are low. Some parts have higher impact than other, i.e. ranging from 0.3 to 1.4 percent point increase (see Appendix 1 for details). Further research needs to show exactly which factors determine the success of this intervention.

6. Expediting repair is almost never used as an operational intervention

We have fairly frequent stream of repaired parts to the central warehouse. Combined with low repair lead times (e.g., 10 days repair plus 15 days in total for shipment) and high repair yields (e.g., 95%), and that repair process is far from demand fulfillment, expediting repair has no value in our case study.

7. Pro-active emergency shipments in combination with lower central warehouse stock or higher demand rates negatively influences nonpremium fill-rates.

From experiments 12–16 we observe that both pro-active emergency shipments and lateral transshipments improve premium fill rates. However, pro-active emergency shipments negatively influence nonpremium fill rates, due to lower central warehouse stock levels (experiment 13) or due to higher demand rates (experiment 23), whereas this was not the case in the base case setting. As the premium demand rate is roughly 10% of the nonpremium demand rate and 20% to 30% of the regular local warehouse replenishments are replaced by emergency shipments, there is only a slight, but significant impact on the nonpremium fill rates. Moreover, we see that stock reservation for premium demand is still not fruitful for premium fill rates. However, due to this reservation, the number of reactive emergency shipments drop and this reduces costs.

8. Proactive lateral transshipment have more positive impact when increasing shipment lead-times between the central warehouse and the local warehouse.

The impact of pro-active lateral transshipments is further increased when the shipment lead-times between the central warehouse and local warehouses. We would also expect similar behavior for pro-active emergency shipments, but the 99.5% premium fill-rate of experiment 4 is difficult to improve.

The last conclusion confirms our hypothesis H1 for sensitivity analysis as mentioned in Section 7.1. However, we have to reject our hypothesis H2 and H3: higher demand rates do not lead to more impact of interventions, and stock reservations are not more effective under less central warehouse stock.

7.3. Impact of alert generation

Planners using service control towers in practice face the issue that too many alerts are being generated to handle. Therefore, it makes sense to study how many alerts – and thus interventions – account for the largest part of the gain. The decision rule states that whenever the expected total costs (i.e., expected backorders costs plus intervention costs) decrease when performing an intervention, it is executed. This *net improvement* may be close to zero, resulting in an additional shipment or stock reservation with very marginal impact. We show the analysis in Fig. 3.

We observe that 29% of the shipments (558 out of 1909) account for 80% of the improvement when deploying the look-ahead heuristic. A threshold-based approach, which controls the alert generation, is useful to limit the number of proactive transshipments without losing too much performance.

8. Conclusions and further research

This paper studied performance improvement in multi-echelon, closed loop spare part supply chains using operational interventions based on real-time status information in a service control tower. Proactive interventions include lateral transshipments, emergency shipments, stock reservations, expediting part repairs, and early new buys of parts. Numerical experiments based on a case study of a global IT-manufacturer reveal the following insights:

1. Operational interventions make most sense in the downstream part of the supply chain. If we move upstream the supply chain, expediting or prioritization makes less sense, likely because more process steps have to be executed until the part arrives at the final customer.
2. Pro-active lateral transshipments in the local network and pro-active emergency shipments from the central warehouse to the local warehouse have the highest impact on fill rates, at the costs of additional shipments.
3. The proposed look-ahead heuristic for pro-active lateral transshipments performs similarly to the more advanced Markov Chain approach, yet at the expense of twice as many lateral transshipments.
4. Stock reservation for direct demand positively influences non-premium fill rates, particularly when the central warehouse stock is low, at the expense of a decrease in premium fill rates, rendering the intervention counterproductive. This can be counteracted by combining stock reservation for direct demand with proactive lateral transshipments, emphasizing the value of integrating interventions.
5. For proactive lateral transshipments it holds that roughly 30% of the performed interventions account for 80% of the improvement in fill rates, underling the importance of a proper alert generation mechanism to handle only the alerts where interventions are most useful.
6. Including information on losses in the supply chain has a small positive impact on fill rates at little costs.
7. Stock reservation for premium demand and expediting repair do not influence fill rates, also not when lower yields (more losses in the supply chain) are enforced.

As further research, we include the following topics:

- A more extensive study on a wider range of problem instances, for example of a global supply chain with longer lead times. Possibly, the subset of the proactive interventions for which we did not find significant results yet, can be more useful.
- Extend the approach to allow for compound Poisson demand to capture intermittent demand.
- Improved rules for stock reservation that look further upstream the supply chain and that distinguish more than two demand classes.

Acknowledgment

This work is a part of the project on Proactive Service Logistics for Advanced Capital Goods Next (Project number 438-15-620), which is supported by TKI-Dinalog (Dutch Institute for Advanced Logistics).

Appendix 1. Experimental results

Table A1, Table A2, Table A3.

Table A1
Experimental results for regular settings and 5% less warehouse stock. Numbers in bold indicate a significant difference compared to no interventions (alpha = 0.05). NP = non-premium demand, *P* = premium demand. L.T. = Lateral Transhipment, E.S. = Emergency Shipment, SR-*d* = Stock reservation for direct demand, SR-*p* = Stock reservation for premium demand.

Regular settings	No interventions		L.T.		E.S.		Include return info		Expedite repair		SR-d		SR-p		LT + SR-d		LT + SR-p		All	
SKU	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P
1	95.7%	94.9%	−0.3%	**2.9%**	−0.3%	**4.4%**	**0.3%**	0.0%	−0.2%	−0.1%	**1.0%**	**−1.3%**	−0.2%	0.0%	**0.7%**	**1.1%**	−0.1%	**2.8%**	**−6.9%**	−0.5%
2	95.3%	95.2%	−0.1%	**3.1%**	−0.5%	**3.7%**	0.6%	0.0%	0.4%	−0.1%	**1.0%**	**−3.2%**	−0.1%	−0.1%	**0.8%**	**2.5%**	−0.1%	**3.1%**	**−2.6%**	**1.7%**
3	95.3%	95.5%	−0.2%	**3.3%**	−0.1%	**3.9%**	**0.3%**	0.0%	−0.1%	0.0%	**0.4%**	**−1.7%**	−0.2%	0.1%	**0.4%**	**2.3%**	−0.2%	**3.3%**	**−6.3%**	**4.1%**
4	95.7%	95.4%	−0.2%	**1.4%**	−0.3%	**3.8%**	0.1%	0.0%	0.0%	0.0%	**1.1%**	**−1.5%**	−0.2%	−0.1%	**0.6%**	**1.0%**	0.0%	**1.5%**	**−4.3%**	**−1.5%**
5	95.5%	94.6%	0.1%	**4.1%**	0.0%	**5.3%**	**0.4%**	0.0%	0.1%	0.0%	**0.8%**	**−1.2%**	0.1%	0.0%	**0.7%**	**3.8%**	0.1%	**3.2%**	**−2.1%**	**3.7%**
6	95.2%	95.8%	0.4%	**2.6%**	0.3%	**3.6%**	**1.4%**	0.1%	0.3%	0.0%	**1.3%**	**−1.1%**	0.2%	0.0%	**1.3%**	**2.4%**	0.4%	**2.6%**	0.3%	**1.9%**
7	95.5%	95.4%	0.2%	**2.3%**	−0.1%	**4.1%**	−0.1%	−0.1%	0.3%	0.0%	**1.8%**	**−3.7%**	0.1%	−0.1%	**1.6%**	**1.5%**	0.1%	**2.3%**	**−1.0%**	**−6.8%**
8	95.2%	94.7%	0.6%	**4.4%**	0.2%	**4.5%**	**0.8%**	0.2%	0.5%	0.1%	**1.7%**	**−4.6%**	0.1%	0.1%	**2.1%**	**3.7%**	0.3%	**4.4%**	0.5%	**1.3%**
9	95.0%	95.8%	0.4%	**1.5%**	0.4%	**3.6%**	**0.9%**	0.2%	0.7%	0.1%	**3.2%**	**−4.5%**	0.2%	0.2%	**2.8%**	**0.8%**	0.4%	**1.7%**	0.2%	**−9.7%**
10	95.0%	95.8%	−0.1%	**3.5%**	0.1%	**3.7%**	0.0%	−0.1%	0.1%	−0.1%	**1.3%**	**−7.0%**	−0.1%	0.0%	**1.3%**	**2.7%**	−0.1%	**3.5%**	**−2.7%**	**−1.8%**
Avg	**95.4%**	**95.2%**	0.0%	**2.9%**	−0.1%	**4.1%**	0.5%	0.0%	0.1%	0.0%	**1.1%**	**−2.6%**	−0.1%	0.0%	**1.0%**	**2.0%**	0.0%	**2.8%**	**−3.2%**	−0.3%
5% less warehouse stock																				
SKU	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P
1	91.7%	94.0%	0.3%	**3.2%**	**−1.6%**	**4.2%**	**0.9%**	0.4%	0.3%	0.1%	**1.7%**	**−2.1%**	0.3%	0.1%	**1.2%**	**1.1%**	0.3%	**1.0%**	**−7.0%**	**−1.5%**
2	92.7%	94.4%	−0.1%	**3.7%**	**−0.6%**	**3.6%**	**1.1%**	0.4%	0.4%	0.3%	**2.0%**	**−4.0%**	−0.1%	0.3%	**1.8%**	**2.9%**	0.2%	**3.6%**	**−1.9%**	**1.0%**
3	90.9%	95.1%	0.0%	**3.2%**	0.0%	**3.9%**	**0.7%**	0.1%	−0.1%	0.0%	**0.7%**	**−3.0%**	−0.2%	−0.1%	**0.9%**	**1.9%**	−0.1%	**3.2%**	**−6.8%**	**4.1%**
4	88.1%	94.0%	−0.4%	**1.9%**	**−2.7%**	**1.8%**	0.1%	−0.1%	−0.3%	−0.3%	**2.3%**	**−4.4%**	−0.1%	−0.2%	**2.1%**	**1.0%**	−0.4%	**2.1%**	**−3.2%**	**−5.7%**
5	89.7%	94.3%	−0.2%	**4.3%**	0.1%	**5.2%**	0.5%	−0.1%	0.0%	0.0%	**0.8%**	**−2.6%**	0.0%	0.0%	**0.8%**	**3.6%**	−0.2%	**4.3%**	**−3.1%**	**3.6%**
6	92.8%	95.4%	0.1%	**2.8%**	−0.2%	**3.6%**	**1.1%**	0.3%	0.2%	0.0%	**1.3%**	**−1.9%**	−0.3%	−0.1%	−0.2%	**2.7%**	0.3%	**2.3%**	0.4%	**2.1%**
7	90.8%	95.0%	0.1%	**2.6%**	0.1%	**3.9%**	0.0%	0.1%	0.2%	0.0%	**3.1%**	**−5.9%**	0.0%	0.0%	0.0%	**2.6%**	0.3%	**0.9%**	**−0.8%**	**−11.8%**
8	94.7%	94.7%	0.1%	**4.3%**	0.1%	**4.2%**	0.5%	0.0%	0.0%	−0.1%	**1.9%**	**−5.8%**	−0.1%	−0.2%	0.0%	**4.4%**	0.4%	**3.4%**	−0.5%	0.1%
9	91.4%	95.6%	0.0%	**1.2%**	−0.2%	**3.1%**	0.4%	0.0%	−0.4%	0.0%	**3.7%**	**−8.7%**	0.1%	0.0%	0.0%	**1.2%**	0.4%	**1.9%**	0.3%	**−16.0%**
10	90.8%	95.5%	−0.1%	**3.8%**	−0.4%	**3.6%**	0.1%	0.0%	0.3%	0.0%	**2.0%**	**−10.9%**	0.0%	0.0%	−0.3%	**3.8%**	0.2%	**2.1%**	**−2.6%**	**−4.7%**
Avg	91.2%	94.6%	0.0%	**3.1%**	**−0.7%**	**3.7%**	0.6%	0.2%	0.1%	0.0%	**1.8%**	**−4.3%**	0.0%	0.0%	0.9%	**2.3%**	0.1%	**2.4%**	**−3.2%**	**−2.1%**

Table A2
Experimental results for doubled shipment lead-time between central warehouse and local warehouse (SKUs 11–20). Numbers in bold indicate a significant difference compared to no interventions (alpha = 0.05). NP = non-premium demand, *P* = premium demand. L.T. = Lateral Transhipment, E.S. = Emergency Shipment, SR-*d* = Stock reservation for direct demand, SR-*p* = Stock reservation for premium demand.

Doubled shipment lead-time between central warehouse and local warehouses																				
	No interventions		L.T.		E.S.		Include return info		Expedite repair		SR-d		SR-p		LT + SR-d		LT + SR-p		All	
SKU	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P
11	95.4%	95.0%	−0.1%	**3.4%**	−0.2%	**4.5%**	**0.7%**	0.0%	0.0%	−0.2%	**1.0%**	**−0.9%**	−0.1%	0.0%	**1.0%**	**2.6%**	0.0%	**3.5%**	**−6.5%**	**2.6%**
12	94.8%	94.8%	0.2%	**3.6%**	−0.2%	**4.3%**	**1.1%**	0.1%	0.0%	0.0%	**1.6%**	**−1.6%**	0.3%	−0.1%	**1.6%**	**3.4%**	0.1%	**3.5%**	**−2.4%**	**2.6%**
13	95.0%	94.9%	0.2%	**3.6%**	0.0%	**4.6%**	**0.5%**	0.2%	0.0%	0.1%	**0.7%**	**−1.4%**	0.0%	0.2%	**0.7%**	**1.9%**	0.2%	**3.4%**	**−5.8%**	**3.0%**
14	95.7%	95.4%	−0.2%	**2.0%**	−0.3%	**4.2%**	0.2%	−0.1%	−0.3%	−0.1%	**0.9%**	**−0.6%**	−0.4%	−0.1%	**0.4%**	**1.6%**	0.1%	**2.0%**	**−3.7%**	**2.6%**
15	95.5%	94.8%	0.0%	**4.5%**	0.1%	**5.1%**	**0.5%**	0.2%	−0.1%	0.3%	**0.8%**	**−0.8%**	−0.1%	0.2%	**0.6%**	**4.2%**	0.0%	**4.5%**	**−2.4%**	**3.9%**
16	95.5%	96.4%	−0.1%	**2.1%**	−0.1%	**3.5%**	**1.2%**	0.1%	0.0%	−0.1%	**0.6%**	**−0.4%**	−0.1%	−0.1%	**1.2%**	**1.5%**	0.0%	**1.6%**	**−0.5%**	**1.2%**
17	95.5%	96.1%	0.0%	**1.7%**	−0.2%	**3.7%**	0.1%	0.0%	0.0%	0.0%	**1.4%**	**−1.5%**	0.0%	0.0%	**1.4%**	**0.7%**	0.2%	**2.0%**	**−1.7%**	**−11.0%**
18	95.2%	94.3%	0.3%	**5.2%**	0.2%	**4.8%**	**0.6%**	0.1%	0.3%	0.0%	**1.8%**	**−2.7%**	0.1%	0.0%	**1.7%**	**4.1%**	0.1%	**4.4%**	0.4%	**2.4%**
19	95.6%	96.9%	−0.1%	**1.5%**	−0.2%	**2.6%**	0.1%	0.1%	−0.2%	0.0%	**2.2%**	**−2.3%**	0.2%	0.1%	**2.1%**	**1.0%**	0.1%	**1.6%**	**−0.6%**	**−16.4%**
20	95.1%	94.2%	−0.1%	**5.7%**	−0.1%	**5.6%**	−0.2%	−0.1%	0.1%	0.0%	**1.1%**	**−4.7%**	−0.1%	0.0%	**1.3%**	**4.7%**	−0.2%	**5.3%**	**−3.0%**	**0.8%**
Avg	95.3%	95.2%	0.0%	**3.3%**	−0.1%	**4.3%**	0.6%	0.1%	0.0%	0.0%	**1.1%**	**−1.5%**	0.0%	0.0%	**1.1%**	**2.6%**	0.0%	**3.2%**	**−3.2%**	0.2%

Table A3
Experimental results for regular settings, 5% less warehouse stock and double shipment lead-time between central warehouse and local warehouse and where the demand rates of SKU 1 to 5 have been doubled and renumbered to SKU 21 to 25 accordingly. Numbers in bold indicate a significant difference compared to no interventions (alpha = 0.05). NP = non-premium demand, *P* = premium demand. L.T. = Lateral Transhipment, E.S. = Emergency Shipment, SR-*d* = Stock reservation for direct demand, SR-*p* = Stock reservation for premium demand.

Regular settings	No interventions		L.T.		E.S.		Include return info		Expedite repair		SR-d		SR-p		LT + SR-d		LT + SR-p		All	
SKU	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P
21	95.1%	94.9%	0.1%	**2.8%**	−0.2%	**2.4%**	**0.8%**	0.2%	0.1%	0.0%	**1.4%**	**−2.3%**	0.1%	0.0%	**1.1%**	**2.1%**	0.1%	**2.8%**	**−3.3%**	**2.9%**
22	95.0%	96.6%	0.1%	**1.3%**	0.2%	**1.1%**	**0.6%**	0.1%	0.1%	0.0%	**1.7%**	**−2.6%**	0.1%	0.0%	**1.5%**	**0.6%**	0.1%	**1.4%**	0.0%	**1.6%**
23	95.9%	94.9%	−0.1%	**0.5%**	−0.1%	**3.0%**	**0.6%**	0.1%	0.0%	0.1%	**0.5%**	**−0.9%**	0.0%	0.1%	**0.4%**	0.1%	0.0%	**0.5%**	**−4.3%**	**4.1%**
24	94.8%	95.7%	0.1%	**2.0%**	−0.1%	**2.7%**	0.3%	0.1%	−0.1%	−0.1%	**1.6%**	**−2.1%**	−0.1%	−0.1%	**1.5%**	**1.2%**	0.0%	**2.0%**	**−5.9%**	**1.3%**
25	94.7%	94.4%	0.1%	**4.6%**	0.1%	**4.8%**	**0.7%**	0.0%	0.0%	0.0%	**0.8%**	**−2.4%**	0.0%	0.0%	**0.9%**	**4.3%**	0.1%	**4.6%**	**−3.1%**	**4.4%**
Avg	84.2%	88.3%	0.1%	**2.1%**	0.0%	**2.5%**	**0.6%**	0.1%	0.1%	0.0%	**1.2%**	**−2.1%**	0.1%	0.0%	**1.1%**	**1.5%**	0.0%	**2.2%**	**−3.1%**	**2.7%**
5% less warehouse stock																				
SKU	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P
21	88.6%	93.2%	0.0%	**3.6%**	**−3.2%**	**2.8%**	**1.3%**	0.3%	0.0%	−0.1%	**1.7%**	**−4.9%**	0.0%	−0.1%	**1.5%**	**1.5%**	0.1%	**3.7%**	**−4.9%**	**3.7%**
22	93.4%	96.0%	−0.3%	**1.5%**	**−1.4%**	**1.4%**	**0.8%**	0.3%	−0.1%	−0.2%	**1.8%**	**−3.5%**	−0.1%	−0.2%	**1.6%**	**0.8%**	0.1%	**1.7%**	**−0.7%**	**2.1%**
23	91.1%	94.5%	−0.1%	**0.5%**	**−0.5%**	**3.2%**	**0.9%**	0.1%	−0.1%	0.0%	**0.6%**	**−2.2%**	−0.1%	0.0%	**0.8%**	**0.5%**	0.0%	**0.5%**	**−6.1%**	**3.2%**
24	80.7%	89.7%	−0.2%	**3.5%**	−0.2%	**3.1%**	0.2%	0.4%	0.2%	0.4%	**4.0%**	**−5.7%**	0.2%	0.1%	**2.7%**	**0.8%**	−0.2%	**3.9%**	**−5.0%**	**1.4%**
25	90.6%	94.2%	−0.2%	**4.9%**	0.1%	**5.0%**	**0.8%**	0.1%	0.0%	0.0%	**1.0%**	**−2.9%**	0.0%	0.0%	**1.1%**	**4.5%**	0.1%	**4.9%**	**−5.0%**	**4.6%**
Avg	**89.1%**	**93.7%**	−0.1%	**2.7%**	**−1.2%**	**2.8%**	**0.9%**	0.2%	0.0%	0.0%	**1.8%**	**−4.0%**	0.0%	−0.1%	**1.5%**	**1.4%**	0.0%	**2.9%**	**−4.1%**	**3.0%**
Doubled shipment lead-time between central warehouse and local warehouses																				
	No interventions		L.T.		E.S.		Include return info		Expedite repair		SR-d		SR-p		LT + SR-d		LT + SR-p		All	
SKU	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P	NP	P
21	95.3%	96.0%	−0.1%	**3.0%**	**−2.1%**	**2.8%**	**0.7%**	0.1%	0.1%	0.1%	**1.2%**	**−2.2%**	0.0%	0.1%	**1.1%**	**2.2%**	−0.1%	**2.0%**	**−2.4%**	**0.4%**
22	95.5%	94.7%	−0.1%	**3.6%**	**−1.2%**	**3.7%**	**0.4%**	−0.1%	−0.2%	−0.2%	**1.5%**	**−2.5%**	−0.2%	−0.1%	**1.4%**	**2.5%**	−0.2%	**3.6%**	−0.2%	**1.7%**
23	95.8%	94.8%	0.1%	**2.2%**	**−0.7%**	**4.3%**	**0.7%**	0.2%	0.1%	0.1%	**0.5%**	**−0.5%**	0.1%	0.1%	**0.4%**	**2.0%**	0.1%	**2.2%**	**0.2%**	**4.5%**
24	94.3%	95.4%	0.1%	**1.6%**	**−0.4%**	**3.5%**	**0.7%**	0.1%	0.1%	−0.1%	**1.7%**	**−1.4%**	0.1%	−0.1%	**1.6%**	**1.4%**	0.1%	**2.3%**	**−4.7%**	**2.4%**
25	94.8%	94.7%	0.0%	**4.3%**	−0.1%	**5.0%**	**0.5%**	0.1%	0.0%	0.0%	**0.7%**	**−0.7%**	0.0%	0.0%	**0.9%**	**4.4%**	0.0%	**3.9%**	**−2.9%**	**3.7%**
Avg	**95.2%**	**95.3%**	0.0%	**3.0%**	**−1.0%**	**3.6%**	**0.6%**	0.1%	0.0%	0.0%	**1.2%**	**−1.7%**	0.0%	0.0%	**1.1%**	**2.4%**	−0.1%	**2.7%**	**−1.9%**	**2.0%**

References

Abell, J. B., Miller, L. W., Neumann, C. E., & Payne, J. E. (1992). *DRIVE (Distribution and repair in variable environments): Enhancing the responsiveness of depot repair*. The RAND Corporation, R-3888-AF.

Accenture (2015). Creating a supply chain control tower in the high-tech industry. Retrieved November 7, 2019, from https://www.accenture.com/_acnmedia/accenture/conversion-assets/dotcom/documents/global/pdf/industries_19/accenture-supply-chain-control-tower.pdf

Arts, J., Basten, R., & Van Houtum, G. J. (2016). Repairable stocking and expediting in a fluctuating demand environment: Optimal policy and heuristics. *Operations Research, 64*(6), 1285–1301.

Axsäter, S., Olsson, F., & Tydesjö, P. (2007). Heuristics for handling direct upstream demand in two-echelon distribution inventory systems. *International Journal of Production Economics, 108*(1–2), 266–270.

Caggiano, K. E., Muckstadt, J. A., & Rappold, J. A. (2006). Integrated real-time capacity and inventory allocation for reparable service parts in a two-echelon supply system. *Manufacturing and Service Operations Management, 8*(3), 292–319.

Cohen, M. A., Agrawal, N., & Agrawal, V. (2006). Winning in the aftermarket. *Harvard Business Review, May*, 129–138.

de Kok, A. G., Grob, C., Laumanns, M., Minner, S., Rambau, J., & Schade, K. (2018). A typology and literature review on stochastic multi-echelon inventory models. *European Journal of Operational Research, 269*(3), 955–983.

Fisher, W. W. (1989). An improved simulation model for cannibalization policy performance comparisons in a complex maintenance system. *Simulation, 52*(4), 154–164.

Grahovac, J., & Chakravarty, A. (2001). Sharing and lateral transshipment of inventory in a supply chain with expensive low-demand items. *Management Science, 47*(4), 579–594.

Heijden, M. C. van der, Alvarez, E. M., & Schutten, J. M. J. (2013). Inventory reduction in spare part networks by selective throughput time reduction. *International Journal of Production Economics, 143*(2), 509–517.

Hoadley, B., & Heyman, D. P. (1977). A two-echelon inventory model with purchases, dispositions, shipments, returns and transshipments. *Naval Research Logistics Quarterly, 24*(1), 1–19.

Howard, C., Marklund, J., Tan, T., & Reijnen, I. (2015). Inventory control in a spare parts distribution system with emergency stocks and pipeline information. *Manufacturing and Service Operations Management, 17*(2), 142–156.

Kranenburg, A. A., & Van Houtum, G. J. (2008). Service differentiation in spare parts inventory management. *Journal of the Operational Research Society, 59*(7), 946–955.

Kranenburg, A. A., & Van Houtum, G. J. (2009). A new partial pooling structure for spare parts networks. *European Journal of Operational Research, 199*(3), 908–921.

Kutanoglu, E., & Mahajan, M. (2009). An inventory sharing and allocation method for a multi-location service parts logistics network with time-based service levels. *European Journal of Operational Research, 194*(3), 728–742.

Law, A. M. (2014). *Simulation modeling and analysis* (5th edition). City of publication: New York: McGraw-Hill.

Meissner, J., & Senicheva, O. V. (2018). Approximate dynamic programming for lateral transshipment problems in multi-location inventory systems. *European Journal of Operational Research, 265*, 49–64.

Paterson, C., Kiesmüller, G., Teunter, R., & Glazebrook, K. (2011). Inventory models with lateral transshipments: A review. *European Journal of Operational Research, 210*(2), 125–136.

Silver, E. A., Pyke, D. F., & Thomas, D. J. (2017). *Inventory and production management in supply chains* (4th edition). CRC Press.

Tiemessen, H. G. H., Fleischmann, M., van Houtum, G. J., van Nunen, J. A. E. E., & Pratsini, E. (2013). Dynamic demand fulfillment in spare parts networks with multiple customer classes. *European Journal of Operational Research, 228*(2), 367–380.

Topan, E., Eruguz, A. S., Ma, W., van der Heijden, M. C., & Dekker, R. (2020a). A Review of Operational Spare Parts Planning for Service Control Towers. *European Journal of Operational Research, 282*(2), 401–414.

Topan, E., & van der Heijden, M. C. (2020b). Operational level spare parts planning in a multi-item two-echelon spare parts inventory system. *European Journal of Operational Research, 284*(1), 164–175.

Van Houtum, G. J., & Kranenburg, B. (2015). *Spare parts inventory control under system availability constraints*: Vol. 227. Springer.