Journal Pre-proof

Performance evaluation of a two-echelon inventory system with network lost sales

Chaaben Kouki, Joachim Arts, M. Zied Babai

PII: S0377-2217(23)00766-X
DOI: https://doi.org/10.1016/j.ejor.2023.10.009
Reference: EOR 18705

To appear in: *European Journal of Operational Research*

Received date : 14 December 2022
Accepted date : 6 October 2023

Please cite this article as: C. Kouki, J. Arts and M.Z. Babai, Performance evaluation of a two-echelon inventory system with network lost sales. *European Journal of Operational Research* (2023), doi: https://doi.org/10.1016/j.ejor.2023.10.009.

Performance evaluation of a two-echelon inventory system with network lost sales

Chaaben Kouki *[1], Joachim Arts[2], and M. Zied Babai[3]

[1]Essca School of Management, Angers, France, chaaben.kouki@essca.fr
[2]Luxembourg Centre for Logistics and Supply Chain Management, Université du Luxembourg , Luxembourg, joachim.arts@uni.lu
[3]Kedge Business School, Bordeaux, France, mohamed-zied.babai@kedgebs.com

*Corresponding author *e-mail*: chaaben.kouki@essca.fr

Performance evaluation of a two-echelon inventory system with network lost sales

Chaaben Kouki *[1], Joachim Arts[2], and M. Zied Babai[3]

[1]Essca School of Management, Angers, France, chaaben.kouki@essca.fr
[2]Luxembourg Centre for Logistics and Supply Chain Management, Université du Luxembourg , Luxembourg, joachim.arts@uni.lu
[3]Kedge Business School, Bordeaux, France, mohamed-zied.babai@kedgebs.com

Abstract

Inventory systems are largely analyzed in the literature under the common assumption of backorders due to the complexity of lost sales. In this paper, we consider a two-echelon inventory system composed of a central warehouse and multiple local warehouses subject to lost sales. The demand faced by each local warehouse is a Poisson process and the stock in the warehouses is controlled according to a continuous review base-stock policy. This system has been analyzed in the literature under deterministic or exponential lead-times at the central warehouse, deterministic lead times at the local warehouses and approximate performance evaluations have been proposed for two cases: (1) the demand is lost if no items are available in the local warehouse, the central warehouse, or in the pipeline in between (i.e., a waiting time threshold for incoming demand equal to the local warehouse lead time), and (2) when there is a waiting time threshold less than the local warehouse lead time. Based on a queuing network representation of the system, we extend the performance analysis of the system in the first case by considering generally distributed lead times both at the central and local warehouses and by providing the exact closed-form expressions for the inventory performance measures. In the second case, we provide new approximate solutions under generally distributed lead times at the central warehouse. We numerically show that our exact and approximate solutions perform equally or better than those presented in the literature under deterministic lead times.

Keywords: Inventory, Lost sales, Queueing network, Emergency orders.

*Corresponding author *e-mail*: chaaben.kouki@essca.fr

1 Introduction and motivation

The unplanned downtime of a capital asset, such as an MRI-scanner, aircraft, or production line, has severe consequences for the users and owners of such assets. The quickest way to restore an asset's operational condition is to replace the failed component with a spare part. The inventory system that distributes spare parts to the different locations of assets is essential to mitigate the effects of unplanned downtime. Unfortunately, spare parts are often expensive, such that holding inventory is quite costly (Noorwali et al., 2023). When a stock-out causes continued downtime, the consequences are often severe in terms of the opportunity costs. It is therefore crucial to design an inventory system that can trade-off the cost of downtime due to inventory shortages with the cost of holding inventories.

Given the high cost of downtime, in multi-echelon inventory systems, advanced technologies, such as connected sensors and information systems, can be used to track orders in real time between a central warehouse and local warehouses to reduce the downtime and costs. For instance, when lead-times are known and constant, under a centralized inventory system, information on the remaining time until an order arrives at a local warehouse can be used to better manage the flow of orders and determine the stock levels. Depending on the remaining lead time, it may be possible to place emergency orders to reduce critical machine downtime. Moreover, under a decentralized inventory system, the customers at a local warehouse may not be able to wait more than the imposed maximum time within which it should be delivered, meaning emergency orders to an external supplier should be placed to quickly react to the system downtime. This maximum acceptable waiting time can be determined through service agreements between the central warehouse and local warehouses.

This context motivates our study that considers a two-echelon inventory system composed of a central warehouse and multiple local warehouses in which each local warehouse faces stochastic demand. Demand that arrives at a local warehouse must be served either immediately from stock, or within a waiting time threshold by items in the pipeline between the central and local warehouse. Demand is lost when this cannot be done. We analyze this system by considering a stochastic lead time at the central warehouse and deterministic or stochastic lead times at local warehouses, depending on the system settings. In fact, our model only provides an exact solution in the case where $w_j = l_j$ and an approximation in the case where $w_j < l_j$.

In the last decades, a body of literature has analyzed multi-echelon inventory systems under stochastic demand, and an overview of the literature is provided by Bijvank et al. (2014). However, the literature dealing with multi-echelon inventory systems under lost sales is relatively scarce compared to inventory systems under backorders since the METRIC model of Sherbrooke (1968) has been developed. The most relevant studies to our work are those of Alvarez and van der Heijden (2014) and Johansson and Olsson (2018). In particular, Alvarez and van der Heijden (2014) analyze a two-echelon inventory system where demand is lost if there is no stock available at the local warehouse, central warehouse, and the pipeline in between. They assume a constant lead time at the local warehouse and an exponentially distributed lead time at the central warehouse. Based on the age of the products in the pipeline or in stock, they provide an approximate performance evaluation of the system when the waiting time threshold is equal to the local warehouse lead time. We refer to this case as "network lost sales", which differs from the classic case where lost sales occur if there is no stock at the local warehouse (i.e., waiting time threshold equal to zero) analyzed in Andersson and Melchiors (2001). The same inventory system is analyzed in Özkan et al. (2015) where lost sales occur if there is no stock at the local warehouse and no possible emergency supply from the central warehouse. More recently, Johansson and Olsson (2018) analyzed a two-echelon inventory system where demand is Poisson distributed and all lead times in the inventory system are constant. They consider the case of a waiting time threshold less than or equal to the local warehouse lead time. They propose an approximate solution based on the derivation of the age of products in the pipeline or in stock. We note that these studies provide an approximate analysis of the inventory system under deterministic lead times at the local warehouses and deterministic or exponential lead times at the central warehouse. The objective of our paper is (1) to extend the findings of Alvarez and van der Heijden (2014) to generally distributed lead times at both the central and the local warehouses with an *exact* performance evaluation , and (2) to extend the findings of Johansson and Olsson (2018) to generally distributed lead times at the central warehouse and improve the approximation accuracy. In particular, we consider a two-echelon performance evaluation of a two-echelon inventory system with network lost sales where the inventory in the warehouses is controlled according to a continuous review base-stock policy. The demand faced by each local warehouse is Poisson distributed. We model the system as a queuing network and analyze its performance in terms of four performance measures: the fill rate, the expected stock on-hand, the expected backorders in the local warehouses, and the expected waiting time. Hence, the contributions of this paper are twofold.

1. We generalize the findings of Alvarez and van der Heijden (2014) by considering generally distributed lead times at both the central and the local warehouses. Moreover, we provide exact closed-form expressions for the evaluation of the inventory system in terms of the fill-rate, the fraction of demand lost, the mean stock on-hand, and the mean waiting time. We observe that this inventory system can be recast as a special type of queueing network that admits a product form solution. We also prove that the steady-state probability distribution of the number of orders in the inventory system depends on the lead time distribution only through its mean,

showing that the steady-state probabilities computed by the method of Alvarez and van der Heijden (2014) can also be used to obtain closed-form expressions for the performance measures.

2. We model the inventory system considered by Johansson and Olsson (2018), as a network of loss queues (or tandem queues) with state-dependent arrival rates and generally distributed lead times at the central warehouse. We propose new approximate solutions that are closer to the exact simulation model (developed in Rockwell Arena software) compared to Johansson and Olsson (2018) and Özkan et al. (2015) under deterministic lead times. The relevance for industry of this lost sales case is highlighted in Howard et al. (2015) using the example of the Volvo's spare parts inventory system.

The remainder of the paper is organized as follows. Section 2 provides an overview of the literature dealing with multi-echelon inventory systems under a lost sales assumption. We describe in Section 3 the inventory system we consider and the underlying assumptions, presenting the system modelling using queuing networks in the threshold waiting time cases of being equal to or less than the local warehouse lead times. In Section 4, we provide the performance analysis and the expressions for the performance measures in both threshold waiting time cases. Section 5 is dedicated to the numerical investigation where we assess the performance of our solutions compared to those proposed in the literature. The conclusions and avenues for future research are presented in Section 6.

2 Literature review

There is a considerable body of literature dealing with multi-echelon inventory models since the 1960s. This literature can be divided into two streams depending on how the system reacts to excess demand, i.e., backorders or lost sales. The analysis of multi-echelon inventory systems in the case of backorders has attracted most attention. This literature stream started with the seminal work of Sherbrooke (1968) and the well-known METRIC-model, with many of the results consolidated in books e.g., Sherbrooke (2004), Muckstadt (2005), and Van Houtum and Kranenburg (2015). Exact performance evaluation of a system under a base-stock policy is possible with analyses using service measures, such as truncated waiting times, e.g., Dreyfuss and Giat (2017), Dreyfuss and Giat (2018), Topan et al. (2017) and Dreyfuss and Giat (2019). The Poisson demand assumption, commonly used in the METRIC-based literature, has been relaxed, for example in Costantino et al. (2018) by considering a Zero-Inflated Poisson distribution to deal with irregular demand patterns.

However, the literature dealing with multi-echelon inventory models under lost sales is relatively scarce. The early study of Nahmias and Smith (1994) analyzes a two-echelon inventory system composed of a central warehouse and multiple retailers where the stock is controlled according to a periodic order-up-to-level policy. They assume the lead times in the system are equal to zero, and demand at the retailer is lost when there is no stock in this echelon. Andersson and Melchiors (2001) extend the analysis with the same inventory system under constant lead times and a base-stock inventory control policy. The authors propose an approximate solution to derive the cost in each echelon and the total cost of the system. An overview of the literature dealing with multi-echelon inventory systems under lost sales is provided in Bijvank and Vis (2011).

More recently, the performance evaluation of multi-echelon inventory systems under lost sales has attracted the attention of Alvarez and van der Heijden (2014), Özkan et al. (2015), and Johansson and Olsson (2018), studies that are highly relevant to our paper. Alvarez and van der Heijden (2014) consider a two-echelon inventory system with a central warehouse and multiple local warehouses. The authors assume that demand is lost if there is no stock available at the local warehouse, central warehouse, and the pipeline in between. The lead time distribution is assumed deterministic at the local warehouse, and exponential at the central warehouse. Based on the age of the products in the pipeline or in stock, they provide an approximate performance evaluation of the system when the waiting time threshold is equal to the local warehouse lead time. They also show that their approximation leads to better performance compared to that of Andersson and Melchiors (2001). Özkan et al. (2015) consider a two-echelon inventory system where demand facing an out-of-stock situation at the local warehouse is met from the central warehouse through an emergency order. They propose an iterative procedure to find the fractions of demand satisfied by local and central warehouses and the external supplier. However, in their model, the pipeline between the local warehouse and the central warehouse is never used to meet demand even if there are items in the pipeline that are not yet assigned to any existing backorder. The model we propose in this paper is different from that of Özkan et al. (2015) because it takes into account the remaining lead time of pipeline orders through the modelling of the system with tandem queues. This information can be very useful to reduce the lost sales because sometimes it is better to wait for an order that arrives from the pipelines in a threshold time than to place an (expensive) emergency order that might even arrive later. Johansson and Olsson (2018) analyze the same system as Alvarez and van der Heijden (2014), but assume that the waiting time threshold at the local warehouse can be less than or equal to the warehouse lead time. Moreover, they assume that all lead times in the inventory system are constant. They propose an approximate solution based on the derivation of the age of the products in the pipeline or in stock. In Özkan et al. (2015), the emergency lead time can be set equal to the maximum acceptable waiting time at the local warehouse to find the model setting of Johansson and Olsson (2018). However, to the best of our knowledge, none of the papers in the literature consider generally distributed lead times at the central warehouse. Moreover, there are no exact performance evaluation results in the literature dealing with multi-echelon inventory systems in the network lost sales case.

Before closing this section, we note the literature stream that considers stochastic network representations to analyze inventory systems under more complex inventory control policies, demand processes, and lead times. Such representations often lead to product form solutions (Kouki et al., 2019, 2020). For example, Song and Zipkin (2009) use a queuing network representation to perform a closed-form performance analysis of an inventory system with two supply sources facing Poisson demand under the dual-index policy. This technique is extended to evaluate multi-echelon inventory networks with dual sourcing (Drent and Arts, 2021), also used when demand is a Markov-modulated Poisson process in asymptotic regimes of interest (Arts et al., 2016). For networks with emergency shipments that depend on the state in a complicated way, an exact analysis has also been obtained in some cases (e.g., Howard et al., 2015), and an approximate analysis in others (e.g., Boucherie et al., 2018; Özkan et al., 2015). We use this queuing network representation in our paper to provide an exact performance evaluation of the inventory system of Alvarez and van der Heijden (2014) under generally distributed lead times at both the central and the local warehouses. We also use it to propose new and more accurate approximations for the system of Johansson and Olsson (2018) under generally distributed lead times at the central warehouse and deterministic lead times at the local warehouses.

3 System description and modelling

In this section, we first describe the inventory system we analyze and then explain the modeling method used to derive the performance measures.

3.1 System description and assumptions

We consider a two-echelon inventory system composed of a central warehouse and J local warehouses. Demand at local warehouse j is a Poisson process with rate λ_j for $j \in \{1, \cdots, J\}$, and the total arrival rate is $\lambda_0 = \sum_{j=1}^{J} \lambda_j$. The assumption of a Poisson demand process is appropriate for slow-moving items and it is widely used in the inventory control literature for spare parts, as in Howard et al. (2015), Song and Zipkin (2009) and Johansson and Olsson (2018). In addition, the Poisson arrival assumption allows us to cast the two-echelon inventory system as a network of queues for which we can derive the steady state probabilities. The stock at each local warehouse is controlled with a base-stock inventory policy where the base-stock level at local warehouse j is denoted with S_j for $j \in \{1, \cdots, J\}$. The central warehouse uses an installation base-stock policy with base-stock level S_0. If an item is

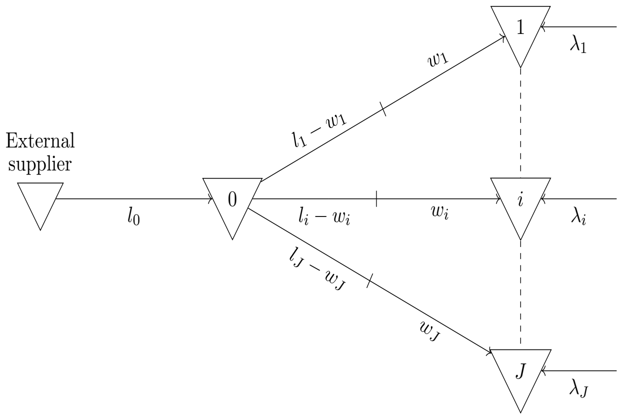

Figure 1: The multi-class loss network

available at warehouse j, demand is immediately met and an order is placed at the central warehouse. The lead time of the central warehouse is generally distributed with mean l_0. We assume that the lead times at local warehouses are stochastic under the Alvarez and van der Heijden (2014)'s settings and deterministic under the Johansson and Olsson (2018)'s settings. The deterministic assumption in the latter case is needed for modelling purposes. We denote the mean lead time by l_j for $j \in \{1, \cdots, J\}$. When demand arrives at local warehouse j, if it is not met within a threshold waiting time, w_j (for $j \in \{1, \cdots, J\}$), it is lost. It is worth pointing out that under deterministic lead times at local warehouses, the condition to accept demand under the Alvarez and van der Heijden (2014) settings is "to have an uncommitted item that is available in the local warehouse, the central warehouse, or the pipeline in between". This condition is equivalent to "having a waiting time threshold that is equal to the local warehouse lead time". Under stochastic lead times, one cannot refer to a waiting time threshold because customers typically require a guaranteed response time to satisfy their demand and the true waiting time may exceed the required response time due to the stochastic nature of the lead times. Consequently, we use in this paper the condition that a demand is lost when there is no uncommitted inventory in the pipeline to local warehouse j or at the central warehouse. This condition is independent of the customer waiting time at the local warehouse when there is a stock out. For simplification purposes, the terminology $w_j = l_j$ is used to refer to the Alvarez and van der Heijden (2014) settings for both deterministic and stochastic lead times. The case where w_j takes values less than l_j is studied in Johansson and Olsson (2018) for deterministic lead times. Both cases $w_j = l_j$ and $w_j < l_j$ occur in practice, and the case $w_j = l_j$ does not require that the remaining lead-times of items in the pipeline can be tracked for practical implementation. Note also that the results of the case $w_j = l_j$ are required to analyze the case $w_j < l_j$.

Our objective is to determine in both cases, $w_j = l_j$ and $w_j < l_j$, the on-hand stock and the number of backorders at warehouse j, denoted with I_j^+ and I_j^- respectively, the fraction of demand directly satisfied from on-hand at warehouse j, denoted with α_j, the fraction of demand that is satisfied

Table 1: Table of notation ($i = 1, 2, \cdots, J$)

	System parameters
J	Total number of local warehouses
λ_j	Demand rate at local warehouse j
λ_0	Total demand rate, $\lambda_0 = \sum_{j=1}^{J} \lambda_j$
S_j	Base stock level at warehouse j
S_0	Base stock level of the central warehouse
$S^{tot} = S_0 + \sum_{j=1}^{J} S_j$	Maximum number of orders in the pipeline to the central warehouse
l_j	Lead time from the central warehouse to warehouse j
w_j	Waiting time of demand at warehouse j
l_0	(mean) lead time from an external supplier to the central warehouse
	State definition and probabilities
n	Total number of orders in the pipeline to the central warehouse or the local warehouse
$Q(n)$	Steady-state probability of having n items in transportation from the external supplier to the central warehouse
$P(n)$	Steady-state probability of having n orders in the pipeline to the local warehouse
$\mathcal{S}$	State space of number of orders in the pipeline to the central warehouse
Ω	State space of backorder levels at the central warehouse
$B_0^j(t)$, B_0^j	Backorder level at the central warehouse at time t and in steady-state associated with local warehouse j, $j = 1, ..., J$, respectively
$N_j(t)$, N_j	Number of orders in the pipeline to the central warehouse at time t and in steady-state due to warehouse j, $j = 1, ..., J$, respectively
$D_j(l_j)$	Demand during lead time l_j at warehouse j, $j = 1, ..., J$.
	Performance measures
$\mathbb{E}[I_j^+]$	Expected inventory level at warehouse j
$\mathbb{E}[I_j^-]$	Expected number of backorders at warehouse j
$\mathbb{E}[W_j]$	Expected waiting time for backordered demand at local warehouse j
α_j	Fraction of demand satisfied without waiting from on-hand inventory at warehouse j
β_j	Fraction of demand satisfied but delayed at warehouse j
δ_j	Fraction of lost demand at warehouse j

after a delay at warehouse j, denoted with β_j, and the fraction of lost demand, denoted with δ_j, for $j = 1, 2, \cdots, J$. Note that $\alpha_j + \beta_j$ is sometimes referred to as the time window fill-rate where w_j is the length of the time window. In Table 1, we summarize the notation used throughout this paper.

3.2 System modelling

To derive expressions for the performance measures associated with local warehouse j, we need the steady-state probability $P(n)$, $n \geq 0$ of having n orders in the pipeline to local warehouse j ($j \in \{1, \ldots, J\}$). Note that $P(n)$, $n \geq 0$ depends on the state of the central warehouse, and can be determined, depending on the central warehouse stock, as follows:

- If the central warehouse has stock on hand, then we can find $P(n)$, $n \geq 0$ by viewing the number of orders in the pipeline between the central warehouse and local warehouse j as a queue with Poisson arrivals with intensity λ_j, and a generally distributed service time with mean l_j. $P(n)$, $n \geq 0$ can therefore be computed using the well-known steady-state probability of the occupancy level in an $M/G/\infty$ queue in conjunction with the state of the number of orders in the pipeline to the central warehouse.

- If the central warehouse is out of stock, the steady-state probability $P(n)$, $n \geq 0$ depends on the system settings:

  Case 1 ($w_j = l_j$): Since the demand is lost when there are no items available at the local warehouse or in the pipeline in between, $P(n)$, $n \geq 0$ depends on the number of backorders B_0^j

associated with local warehouse j and already waiting in the central warehouse. Note that B_0^j cannot exceed S_j. We can find $P(n)$, $n \geq 0$ by modelling the number of orders in the pipeline to each local warehouse j as a tandem queue, as shown in Figure 2. The first node in Figure 2 represents the orders in the pipeline to the central warehouse, forming an $\bullet/G/S^{tot}$ queue with an input rate λ_j that depends on the number of backorders B_0^j, service rate $1/l_0$, and total number of severs S^{tot}. The second node is an $\bullet/G/\infty$ queue with service rate $1/l_j$. For this tandem queue, we show how it can be analyzed exactly, enabling us to derive the steady-state probability $P(n), n \geq 0$, by conditioning on the number of backorders B_0^j. Although we need the distribution of B_0^j to find $P(n), n \geq 0$, the opposite is not true. In other words, we can find the steady-state probability of B_0^j without knowing $P(n)$, $n \geq 0$, since each time demand arrives, it will always be accepted as long as $B_0^j < S_j$, regardless of the number of orders in the pipeline between the central and local warehouses. This result is crucial to derive a closed-form expression for $P(n)$, $n \geq 0$. The steady-state probability of the number of backorders B_0^j is derived by analyzing the central warehouse separately, regardless of the number of orders in the pipeline between the central and local warehouse, in other words, regardless of the state of orders in the pipeline at queue 2. Once the steady-state probability of B_0^j is found, we can derive $P(n), n \geq 0$. Note that the model of Özkan et al. (2015) does not assume a threshold waiting time at the local warehouses. We approximate the steady-state probability $P(n)$, $n \geq 0$ of the Özkan et al. (2015) model using our results under $w_j = l_j$.

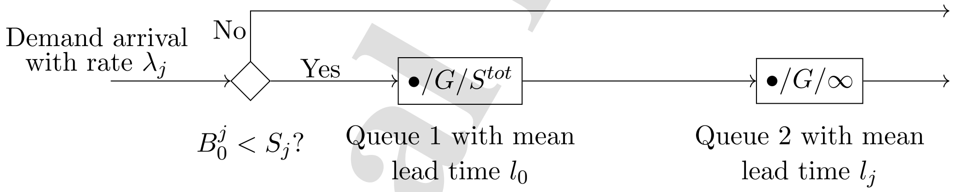

Figure 2: Tandem queues of number of orders in the pipeline to local warehouse j when $w_j = l_j$

Case 2 ($w_j < l_j$): In this case the number of orders in the pipeline to the central warehouse and local warehouse depend on each other. We model the inventory system by using three queues as shown in Figure 3. The first queue represents the number of orders being routed to the central warehouse. The second and third queues represent the number of orders destined for the local warehouse j with a mean processing lead time $l_j - w_j$ and w_j, respectively. Hence, when demand arrives, it will be accepted if the number of backorders associated with local warehouse j and waiting at the central warehouse is strictly less than S_j (i.e., $B_0^j < S_j$), or if the total number of backorders B_0^j plus the number of orders to local warehouse j at the second queue is strictly less than S_j. Note that the case $w_j < l_j$ is hence modelled as a tandem queuing network with blocking, where demand arrival at the local warehouse is blocked either because $B_0^j = S_j$, or because the number of orders to local warehouse j at the second node is S_j. It should also be noted that a tandem queue with a blocking mechanism at the first queue that depends on the state of the second queue is difficult to analyze exactly since we need to track all pipeline orders to the warehouses, i.e., the occupancy level of each node in the tandem queues. In a small setting, with exponential lead times it is possible to study the dynamics of the network of queues, but in the general case, such dynamics suffers from the curse of dimensionality and the network analysis becomes extremely complex. Thus, we find an approximate steady-state probability $P(n)$, $n \geq 0$ of having n orders in the pipeline to the local warehouse by decomposing the queuing network into two independent sub-networks: one related to the local warehouse consisting in node 1 only, the other related to the local warehouse consisting in nodes 2 and 3 together. We will elaborate more on this in Section 4.2.

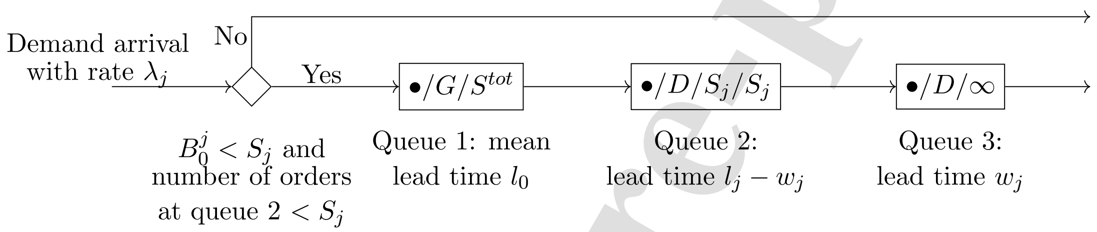

Figure 3: Tandem queues of the number of orders in pipeline to local warehouse j when $w_j < l_j$

In summary, to compute $P(n), n \geq 0$, we should first find the marginal probability density of the number of backorders associated with local warehouse j and waiting at the central warehouse, B_0^j, which requires determining the steady-state probability of having n orders in the pipeline to the central warehouse, denoted with $Q(n)$, $n \in \{0, \ldots, S^{tot}\}$. This is the objective of the next section, where we also express the performance measures.

4 Inventory performance analysis

In this section, we derive the the steady state probabilities of having n orders in the system and we provide the expressions for the performance measures α_j, β_j, δ_j, $\mathbb{E}[I_j^+]$, $\mathbb{E}[I_j^-]$ and $\mathbb{E}[W_j]$. Note that calculating the performance measures α_j and β_j enables the straightforward calculation of the expected waiting time of demand at local warehouse j using Little's law as $\mathbb{E}[W_j] = \frac{\mathbb{E}[I_j^-]}{(\alpha_j+\beta_j)\lambda_j}$. To do so, we calculate in the two cases: $w_j = l_j$ and $w_j < l_j$, the probability of having b_j backorders at the central warehouse associated with local warehouse j, $\mathbb{P}\{B_0^j = b_j\}$, and the steady-state probability of having n orders in the pipeline to the central warehouse (i.e., $Q(n)$, $n \in \{0, \ldots, S^{tot}\}$).

4.1 Analysis of the case $w_j = l_j$

4.1.1 Central warehouse analysis

In this section, we provide a closed-form expression for the steady-state probability of having a total of n orders in the pipeline to the central warehouse, i.e., $Q(n)$, $n \in \{0, \ldots, S^{tot}\}$. Our analysis is built on the following two observations:

1. The steady-state probability $Q(n)$ does not depend on l_j for $j = 1, \cdots, J$.

2. If $n < S_0$, then the demand arrival rate at the central warehouse is λ_0.

The first observation means that the demand arrival process at the central warehouse generated from local warehouse j remains stationary as long as the number of backorders at the central warehouse associated with local warehouse j is less than S_j. Indeed, each time demand from local warehouse j arrives, and the central warehouse has stock on-hand or does not have stock on-hand but the number of backorders is less than S_j, demand will always trigger an order at the central warehouse. This event is independent of whether the local warehouse has stock or the number of orders in progress between the central warehouse and local warehouse j. Demand at local warehouse j is lost only when the number of backorders at the central warehouse associated with local warehouse j reaches S_j. This means that the lead time l_j has no impact on the demand arrival process at the central warehouse. This property does not hold for models such as those of Özkan et al. (2015) and Johansson and Olsson (2018) because we need to know the number of orders in progress to the local warehouse to decide whether to accept or reject an incoming demand. In contrast, in our framework with $w_j = l_j$, we only need information about the number of backorders at the central warehouse associated with local warehouse j to decide whether to accept or reject an incoming demand.

The second observation means that the number of items on-order between the external supplier and the central warehouse is distributed as in a loss queue with a Poisson arrival process with rate λ_0, a service time of expected duration l_0, and S_0 servers. This observation enables us to find the probability $Q(n)$ for any $n \leq S_0$. We will elaborate on this further on in the paper.

To find the probability $Q(n)$, $n > S_0$, we need to know the number of backorders at the central warehouse associated with local warehouse j. However, the backorder level to local warehouse j held at the central warehouse depends on the number of orders triggered by local warehouse j and accepted by the central warehouse. We therefore need to define two random variables, one representing the number of orders already accepted at the central warehouse, and the other representing the number of backorders associated with each local warehouse j.

To do so, let $N_j(t)$ denote the number of orders in the pipeline to the central warehouse that were triggered by demand at local warehouse $j \in \{1, \ldots, J\}$. Note that demand at local warehouse j causes $N_j(t)$ to increase by 1 provided that the waiting time for this item will be less than l_j. Moreover, note that the waiting time w_j is always less than l_j as long as the number of backorders in the pipeline to the central warehouse does not exceed S_j. Otherwise, there must be an item in the pipeline between the central warehouse and local warehouse j that will arrive within l_j units of time and can thus be allocated to incoming demand. Furthermore, after demand acceptance, there must be $N_j(t) \leq S_0 + S_j$ and $\sum_{j=1}^{J} N_j(t) \leq S_0 + \sum_{j=1}^{J} S_j$ at any time t. We denote with $\mathbf{N(t)} = (N_1(t), \ldots, N_J(t))$ the stochastic vector of the total number of orders in the pipeline to the central warehouse at time t. Its associated stationary stochastic process is denoted with $\{\mathbf{N(t)}\}$. Let $\mathbf{N} = (N_1, \ldots, N_J)$ be the random variable denoting the number of orders in the pipeline in steady-state. The state space of $\{\mathbf{N(t)}\}$ is given by $\mathcal{S}$:

$$\mathcal{S} = \left\{ \mathbf{n} := (n_1, n_2, \ldots, n_J) \mid n_j \leq S_j + S_0, \ j = 1, \ldots, J, \text{ and } \sum_{j=1}^{J} n_j \leq \sum_{j=1}^{J} S_j + S_0 \right\}. \tag{1}$$

Note also that the dynamics of the process $\{\mathbf{N(t)}\}$ are not sufficient to determine the steady-state probability $Q(n)$, $n \in \{0, \ldots, S_0 + \sum_{j=1}^{J} S_j\}$. In fact, since the number of backorders at the central warehouse cannot exceed S_j, the steady-state of the number of orders in the pipeline to the central warehouse cannot be determined completely through the vector $\mathbf{N}$ in steady-state. Consider an example of two local warehouses setting the parameters $S_0 = 3$, $S_1 = 1$, and $S_2 = 2$. Assume that the description of the state of orders in the pipeline to the central warehouse is only via vector $\mathbf{N}$. Thus, if the orders in the pipeline are $\mathbf{N} = (2, 3)$, two orders are triggered by local warehouse 1 and three orders by local warehouse 2. The state $(2, 3)$ also means there are 2 backorders in the system. If these two backorders are associated with local warehouse 1, then any demand from local warehouse 1 should be rejected because the maximum number of backorders due to local warehouse 1 has already been reached. On the other hand, if we ignore the threshold level S_j of backorders associated with warehouse j, we still have $N_j \leq S_0 + S_j$, and demand at warehouse j is accepted at the central warehouse, whereas it should not be.

Therefore, to find the steady-state probabilities of the number of orders in the pipeline to the central warehouse, we need to know the steady-state probabilities of backordered demand at the central warehouse stemming from all local warehouses, $j = 1, \ldots, J$. For this purpose, we consider our second random variable $B_0^j(t)$ as the number of backorders in the central warehouse associated with the local warehouse j at time t. The state of total backorders at time t is denoted with $\mathbf{B_0(t)} = (B_0^1(t), \ldots, B_0^J(t))$ and its associated steady-state stochastic process with $\{\mathbf{B_0(t)}\}$ with a finite state space Ω

$$\Omega = \{\mathbf{b} := (b_1, \ldots, b_J) | b_j \leq S_j\}. \tag{2}$$

We denote with $\mathbf{B_0} = (B_0^1, \ldots, B_0^J)$ the vector of all backorders in steady-state. Note that $\sum_{j=1}^J N_j \leq S_0$ implies that $B_0^j = 0$ for all local warehouses $j \in \{1, \ldots, J\}$, and for any $\mathbf{N} \in \mathcal{S}$, $B_0^j \leq N_j$, $j \in \{1, \ldots, J\}$. Therefore, $\sum_{j=1}^J B_0^j > 0$ only for

$$\mathbf{n} \in \mathcal{S}(n), \quad \mathcal{S}(n) = \left\{ \mathbf{n} \in \mathcal{S} \mid n = \sum_{j=1}^{J} n_j > S_0 \right\},$$

We now derive the steady-state probability of the vector $(\mathbf{b}, \mathbf{n})$, which we denote with $\pi(\mathbf{b}, \mathbf{n})$ as follows:

$$\pi(\mathbf{b}, \mathbf{n}) = \mathbb{P}\{\mathbf{B_0} = \mathbf{b}, \mathbf{N} = \mathbf{n}\} = \mathbb{P}\{\mathbf{B_0} = \mathbf{b} | \mathbf{N} = \mathbf{n}\} \mathbb{P}\{\mathbf{N} = \mathbf{n}\}, \quad \mathbf{b} \in \Omega, \mathbf{n} \in \mathcal{S}(n) \tag{3}$$

where

$$\mathbb{P}\{\mathbf{B_0} = \mathbf{b} | \mathbf{N} = \mathbf{n}\} = \lim_{t \to \infty} \mathbb{P}\left\{B_0^1(t) = b_1, \ldots, B_0^J(t) = b_J | N_1(t) = n_1, \ldots, N_J(t) = n_J\right\}, \quad \mathbf{b} \in \Omega, \mathbf{n} \in \mathcal{S}(n)$$

and

$$\mathbb{P}\{\mathbf{N} = \mathbf{n}\} = \lim_{t \to \infty} \mathbb{P}\{N_1(t) = n_1, \ldots, N_J(t) = n_J\}, \quad \mathbf{n} \in \mathcal{S}(n)$$

The vector $(\mathbf{B}_0, \mathbf{N})$ is a reversible process and has a product form solution as shown in Proposition 1.

Proposition 1. *The steady-state probability of* $(\mathbf{b}, \mathbf{n})$ *is given by*

$$\pi(\mathbf{b}, \mathbf{n}) = C \frac{\binom{n_1}{b_1}\binom{n_2}{b_2} \cdots \binom{n_J}{b_J}}{\binom{\sum_{j=1}^J n_j}{\sum_{j=1}^J b_j}} \prod_{i=1}^{J} \frac{(\lambda_i l_0)^{n_i}}{n_i!}, \quad \mathbf{b} \in \Omega, \mathbf{n} \in \mathcal{S}(n),$$

where C *is the normalizing constant.*

Proof. See Appendix 1. ∎

Proposition 1 allows providing a simple expression for $Q(n)$, $n \in \left\{0, \ldots, S_0 + \sum_{j=1}^J S_j\right\}$ using the following corollary.

Corollary 1. *The steady-state probability of having* n *orders in the pipeline of the central warehouse* *is given by*

$$Q(n) = \begin{cases} A\sum_{\substack{\mathbf{n}\in\mathcal{S} \\ n_1+\ldots+n_J=n}} \prod_{j=1}^{J} \frac{(\lambda_j(l_0))^{n_j}}{n_j!} = \frac{(\lambda_0 l_0)^n}{n!} & \text{if } 0 \leq n \leq S_0, \\ A\sum_{\substack{\mathbf{b}\in\Omega, \\ b_1+\ldots+b_J=n-S_0, \\ \mathbf{n}\in\mathcal{S}(n)}} \pi(\mathbf{b},\mathbf{n}) & , \\ = \sum_{\substack{\mathbf{b}\in\Omega \\ b_1+\ldots+b_J=n-S_0}} \left(\frac{(n-S_0)!(\lambda_0 l_0)^{S_0}}{(n)!} \prod_{i=1}^{J} \frac{(\lambda_i(l_0))^{b_i}}{b_i!}\right) & \text{if } n > S_0, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where A is the normalizing constant.

Proof. See Appendix 1. ■

The result of Corollary 1 also shows that the steady-state probability $Q(n)$, $n \in \left\{0,\ldots,S_0+\sum_{i=1}^{J} S_i\right\}$ is insensitive to the lead time distribution, except through its mean l_i, $i \in \{1,\ldots,J\}$. This is in line with the literature related to Jackson queuing networks where the steady-state probability of the network state is insensitive to the distribution of the processing lead time and depends only on its mean (see Boucherie and van Dijk (2011)).

A careful inspection of the preceding proposition and the equations in Alvarez and van der Heijden (2014) will reveal that they are identical, as stated in the following corollary.

Corollary 2. *The expression for the steady-state probability derived by Alvarez and van der Heijden (2014) under constant lead times is exact and given by*

$$Q(n) = \begin{cases} K\frac{(\lambda_0 L_0)^n}{n!}, n \leq S_0, \\ K\frac{(l_0)^n}{(n)!}\prod_{y=0}^{n-1} M(y), \; n > S_0, \end{cases} \tag{5}$$

where

$$M(n) = \begin{cases} \lambda_0, & n \leq S_0 \\ \sum_{j=1}^{J} \lambda_j \left(1 - \mathbb{P}\left\{B_0^j = S_j | n - S_0\right\}\right), & n > S_0 \end{cases}. \tag{6}$$

K is the normalizing constant and $\mathbb{P}\left\{B_0^j = S_j | n - S_0\right\}$ is computed using a truncated multi-nominal distribution detailed in Appendix 2.

Proof. See Appendix 2. ■

Corollary 2 expresses the steady-state probability using the method of Alvarez and van der Heijden (2014). This expression suffers from the curse of dimensionality, which makes the computation of $Q(n)$

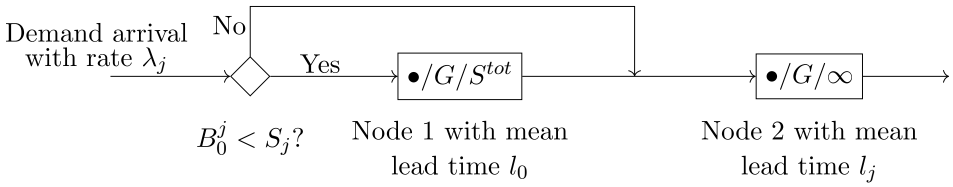

Figure 4: Network with overflow bypass

numerically tedious when the number of classes J and $S_j, j = 1, ...J$ increases. Alvarez and van der Heijden (2014) themselves proposed an approximate method to calculate this probability.

4.1.2 Performance measures

A crucial observation for the purpose of our analysis is also used in Howard et al. (2015). In fact, consider an alternative system where demand that would be lost in the original system is not lost but placed with an outside emergency supplier that also has a lead time of l_j. Note that the behavior of the on-hand inventory in this alternative system is identical to the behavior of the original system. Both systems have identical dynamics for demand filled directly by the system. The on-hand inventory level in the alternative system can be obtained as follows. Let $D_j(l_j)$ denote demand during l_j at warehouse j, and note that $D_j(l_j)$ has a Poisson distribution with mean $\lambda_j l_j$. Let I_j denote the inventory level at warehouse j in the original system, and $\tilde{I}_j$ the inventory level in the alternative system. The above observation implies that $\tilde{I}_j^+ = I_j^+$ where $x^+ = \max(0, x)$. With this observation, we can determine the performance measures.

The alternative system described above for a specific local warehouse j can also be seen as follows. The entire pipeline of products that will be delivered to local warehouse j is depicted in Figure 2 when there are backorders at the central warehouse. The first station in this Figure 2 corresponds to the orders in the pipeline to the central warehouse that are destined for location j. The second station is number of orders in the pipeline between the central warehouse and location j. Unfortunately, the original network has no known product form solution. The alternative system is depicted in Figure 4. This network assumes that products that would normally be lost are placed as an order to an outside supplier with the same lead time as between the central warehouse and local warehouse j. Obviously the network does not represent our network, but does have a product form solution, as also observed by Song and Zipkin (2009) and Howard et al. (2015), and can be exploited, as we will do next.

On-hand stock. Using the observation above, we can write directly that

$$\begin{aligned}\mathbb{P}\left\{I_j^+ = x\right\} &= \mathbb{P}\left\{\tilde{I}_j^+ = x\right\} = \mathbb{P}\left\{(S_j - B_0^j + D_j(l_j))^+ = x\right\} \\ &= \mathbb{P}\left\{B_0^j + D_j(l_j) = S_j - x\right\} = \sum_{y=0}^{S_j} \mathbb{P}\left\{B_0^j + D_j(l_j) = S_j - x \mid B_0^j = y\right\} \mathbb{P}\left\{B_0^j = y\right\} \\ &= \sum_{y=0}^{S_j} \mathbb{P}\left\{D_j(l_j) = S_j - x - y\right\} \mathbb{P}\left\{B_0^j = y\right\} = \sum_{y=0}^{S_j} e^{\lambda_j l_j} \frac{(\lambda_j l_j)^{S_j - x - y}}{(S_j - x - y)!} \mathbb{P}\left\{B_0^j = y\right\}\end{aligned} \tag{7}$$

The expected on-hand stock can easily be computed from the distribution as $\mathbb{E}[I_j^+] = \sum_{x=0}^{S_j} x\mathbb{P}\{I_j = x\}$.

However, note that $\tilde{I}_j^- \neq I_j^-$ and $x^- = \max(0, -x)$, so that the expected backorders cannot be determined in similar fashion. We defer this analysis to later.

Fraction of demand filled immediately, after waiting time, and lost. Demand in our network can be either filled immediately from stock on-hand, after a waiting time from stock elsewhere in the network, or lost entirely. From the steady-state probability $Q(n)$, $n \leq S_0 + \sum_{j=1}^{J} S_j$, we obtain the fraction of demand that is lost to the network for each local warehouse, that is δ_j. Let

$$P^j(k) = \mathbb{P}\left\{B_0^j = k\right\} = \sum_{\mathbf{b}^j \in \Omega_j} \left(\frac{(k + \sum_{i=1, i\neq j}^{J} b_i)!(\lambda_0 l_0)^{S_0}}{(S_0 + k + \sum_{i=1,i\neq j}^{J} b_i)!} \frac{(\lambda_j l_0)^k}{k!} \prod_{i=1, i\neq j}^{J} \frac{(\lambda_i l_0)^{b_i}}{b_i!} \right), \tag{8}$$

be the probability of having k backorders at the central warehouse associated with local warehouse j, and Ω_j the multidimensional state space that excludes the number of backorders triggered by local warehouse j:

$$\Omega_j = \left\{ \mathbf{b}^j = (b_1, \ldots, b_{j-1}, b_{j+1} \ldots, b_J) | b_j \leq S_j, j = 1, \ldots, J, \sum_{i=1, i\neq j}^{J} b_i \leq \sum_{i=1,i\neq j}^{J} S_i \right\}. \tag{9}$$

Because Poisson arrivals see time averages (Wolff, 1982), it follows that the fraction of demand lost by the network at warehouse j, is given by

$$\delta_j = P^j(S_j). \tag{10}$$

Similarly, the fraction of demand that can be filled immediately from the inventory on-hand at local warehouse j can be computed directly as

$$\alpha_j = \left(\sum_{k=0}^{S_0 - 1} Q(k) \right) \sum_{n=0}^{S_j - 1} \frac{(\lambda_j l_j)^n e^{-\lambda_j l_j}}{n!} + \sum_{k=0}^{S_j - 1} \sum_{n=0}^{S_j - k - 1} \left(\frac{(\lambda_i l_j)^n e^{-\lambda_j l_j}}{n!} P^j(k) \right). \tag{11}$$

Since $\alpha_j + \beta_j + \delta_j = 1$ by definition, we finally obtain

$$\beta_j = 1 - \alpha_j - \delta_j. \tag{12}$$

Expected backorders and on-hand inventory for local warehouses. For the expected backorders at local warehouse j, we use Little's law and the relation of the original system to the alternative system described at the beginning of this section. Note that I_j^- denotes the number of backorders at local warehouse j as $x^- = \max(0, -x)$. We follow the idea of Howard et al. (2015). A critical observation at this point is that $\tilde{I}_j^- = I_j^- + \tilde{X}$ where $\tilde{X}$ is the number of customers with backorders that will be filled by the outside supplier that delivers after l_j time units in the alternative system. According to Little's law, we have $\mathbb{E}[\tilde{X}] = \delta_j \lambda_j l_j$. Therefore, $\mathbb{E}[I_j^-]$ can be determined when $\mathbb{E}[\tilde{I}_j^-]$ is known. Fortunately, this quantity can be found in a similar way to the on-hand stock:

$$\begin{aligned} \mathbb{E}[\tilde{I}_j^-] &= \mathbb{E}[(D_j(l_j) + B_0^j - S_j)^+] = \mathbb{E}[D_j(l_j) + B_0^j - S_j] + \mathbb{E}[(S_j - D_j(l_j) - B_0^j)^+] \\ &= \lambda_j l_j + \mathbb{E}[B_0^j] - S_j + \mathbb{E}[\tilde{I}_j^+] \end{aligned} \tag{13}$$

Here, $\mathbb{E}[\tilde{I}_j^+]$ follows from Equation (7) and $\mathbb{E}[B_0^j]$ can be computed directly from Equation (8). Now the expected backorders at local warehouse j can be expressed as

$$\mathbb{E}[I_j^-] = \mathbb{E}[\tilde{I}_j^-] - \mathbb{E}[\tilde{X}] = \mathbb{E}[\tilde{I}_j^-] - \lambda_j \delta_j l_j. \tag{14}$$

4.2 Analysis of the case $w_j < l_j$

In this section, we propose new approximations of two models in the literature. The first concerns the model that Özkan et al. (2015) studied. Their work is similar to ours in Subsection 4.1.1 with one difference: when demand cannot be satisfied with available stock or the central warehouse, it is lost, while in our first model, demand is satisfied after a waiting time from the regular channel as long as the number of backorders at the central warehouse does not exceed S_j, $j \in \{1, \ldots, J\}$.

The second model we address is that of Johansson and Olsson (2018) where an emergency replenishment is triggered based on the remaining lead time of orders in the pipeline to each local warehouse $j, j = 1, ..., J$. Let us denote with w_j the maximum waiting time a customer is willing to accept if no products are available in the local warehouse. According to Johansson and Olsson (2018), when a customer arrives at the local warehouse, an emergency replenishment is requested from an outside supplier if the remaining lead time of the orders in the pipeline that have not yet been allocated to an existing backorder exceeds w_j. In the model of Johansson and Olsson (2018), there is a possibility that the emergency order cannot be executed, but we relax this assumption by assuming that all emergency replenishments are performed with probability 1. If we set w_j equal to the lead time l_j, then Johansson and Olsson (2018)'s model is reduced to ours for deterministic lead times. In this case, we can analyze such a system exactly. In the case of $w_j = 0$, we refer to the model of Andersson and Melchiors (2001). The difficulty in providing an exact solution to these two models lies in the fact that orders to the central warehouse depend on the pipeline between the central warehouse and the local warehouse, which is not the case in the model studied in Section 4.1.

The approximations we propose are easy to understand. In fact, we use the exact steady-state probability in the case where $w_j = l_j$, together with the steady-state probability of the pipeline between the local warehouse and central warehouse. Under the assumption of the model of Özkan et al. (2015), assume (this is an approximation) that the orders in the pipeline between the central warehouse and local warehouse j form an $M/G/S_j/S_j$ queue with a Poisson process with arrival rate λ_j and service time with mean l_j, thus the steady-state probability of having n products during lead time l_j is simply:

$$\begin{cases} p(n, l_j) = \mathbb{P}\{D_j(l_j) = n\} = \frac{\frac{(\lambda_j l_j)^n}{n!}}{\sum_{k=0}^{S_j} \frac{(\lambda_j l_j)^k}{k!}}, \\ P(n, l_j) = \mathbb{P}\{D_j(l_j) \le n\} = \sum_{k=0}^{n} p(k, l_j). \end{cases} \tag{15}$$

For the model of Johansson and Olsson (2018), we make a similar assumption related to the arrival process in the pipeline between the central and local warehouse, i.e., Poisson process with rate λ_j. In the observation of Howard et al. (2015), as stated in Section 4.1.2, the pipeline between the central and local warehouse forms tandem queues $M/G/S_j/S_j \rightarrow \bullet/G/\infty$, where the first queue has an arrival rate λ_j and service rate with mean $l_j - w_j$, and the second queue has a mean service rate w_j. If the first queue is full, incoming demand is satisfied by emergency replenishment with a lead time w_j by overflowing the first node and going directly to the second node. We can write the steady-state probability of having n items during lead time l_j with

$$\begin{cases} p(n, l_j) = \frac{1}{\sum_{k=0}^{S_j} \frac{(\lambda_j(l_j - w_j))^k}{k!}} \sum_{m=0}^{\min(n, S_j)} \frac{(\lambda_j(l_j - w_j))^m}{m} \frac{(\lambda_j w_j)^{n-k}}{(n-k)!} e^{-\lambda_j w_j}, \\ P(n, l_j) = \mathbb{P}\{D_j(l_j) \le n\} = \sum_{k=0}^{n} p(k, l_j). \end{cases} \tag{16}$$

We are now ready to express the steady-state probability at the central warehouse. For the Özkan et al. (2015) model,

$$Q(n) = \begin{cases} A\frac{(\lambda_0 l_0)^n}{n!} & \text{for } 0 \leq n \leq S_0, \\ A\sum_{\substack{\mathbf{b}\in\Omega, \\ b_1+\ldots+b_J=n-S_0}} \left(\frac{(n-S_0)!(\lambda_0 l_0)^{S_0}}{(n)!}\prod_{j=1}^{J}\frac{(\lambda_j(l_0))^{n_j}}{b_j!}P(S_j-b_j,l_j)\right) & \text{for } n > S_0, \\ 0 & \text{otherwise,} \end{cases} \quad (17)$$

which is the same steady-state probability of our original model conditioned by probability $P(S_j - n_j, l_j)$. For Johansson and Olsson (2018), we also need to approximate the arrival rate at the central warehouse, since the orders triggered at the central warehouse depend on the stock available at warehouse j (if any) and on the number of orders in the pipeline during w_j.

$$Q(n) = \begin{cases} A\frac{(\lambda_0 l_0)^n}{n!} & \text{for } 0 \leq n \leq S_0, \\ A\sum_{\substack{\mathbf{b}\in\Omega, \\ b_1+\ldots+b_J=n-S_0}} \left(\frac{(n-S_0)!(\lambda_0 l_0)^{S_0}}{(n)!}\prod_{j=1}^{J}\frac{(\lambda_j(l_0))^{b_j}}{b_i!}P(S_j-b_j,l_j-w_j)\right) & \text{for } n > S_0, \\ 0 & \text{otherwise.} \end{cases} \quad (18)$$

where $P(S_j - b_j, l_j - w_j)$ is given by (15) and λ_0 is approximated by

$$\lambda_0 = \sum_{j=1}^{J} \lambda_j \left(1 - \frac{\frac{(\lambda_j(l_j-w_j))^{S_j}}{S_j!}}{\sum_{k=0}^{S_j}\frac{(\lambda_j(l_j-w_j))^k}{k!}}\right).$$

We can now express all the performance metrics we need, but since system downtime is the most critical factor to assess, we focus on the fraction of demand satisfied directly from available stock, the fraction satisfied after a waiting time (if any), and the fraction of demand lost. We will obviously use the same analysis technique as our original model. Evaluating these fractions requires the probability that there are n orders in the pipeline to local warehouse j, which we denote with $P(n)$, $n \in \{0, \ldots, S_j\}$. This probability is given by $p(n, l_j)$ if the central warehouse has stock on hand. However, if the central warehouse is out of stock, $P(n)$ should be conditioned on the number of backorders at the central warehouse. In this case $P(n) = \sum_{k=0}^{n} \mathbb{P}\left\{B_0^i = k\right\} \mathbb{P}\left\{D_i(l_i) = n - k\right\}$. Using the same notation as in the original model, we can deduce the steady-state probability of having n orders in the pipeline to local warehouse i as follow.

Under the Özkan et al. (2015) and Johansson and Olsson (2018) models:

$$\begin{aligned} P(n) \quad &= \quad \left(\sum_{k=0}^{S_0-1} Q(k)\right) p(n, l_j) & (19) \\ &+ \sum_{\mathbf{b}^j\in\Omega_j} \left(\sum_{k=0}^{min(n,S_j)} \frac{(k+\sum_{i=1,i\neq j}^{J} b_i)!(\lambda_0 l_0)^{S_0}}{(S_0+k+\sum_{i=1,i\neq j}^{J} b_i)!}\frac{(\lambda_j l_0)^k}{k!}p(n-k,l_j)\prod_{i=1,i\neq j}^{J}\frac{(\lambda_i l_0)^{b_i}}{b_i!}P(S_i-b_i,l_i-w_i)\right), \end{aligned}$$

where $w_i = 0$ for Özkan et al. (2015) and $p(n-k, l_j)$ is obtained with Equations (15) and (16) for the Özkan et al. (2015) and the Johansson and Olsson (2018) models, respectively. It should be noted that when using $p(n-k, l_j)$ under the Johansson and Olsson (2018) setting, the number of pipeline orders to the local warehouse waiting at the central warehouse plus those at the second queue with a mean lead time $l_j - w_j$ cannot exceed S_j .

From Equation 19, we can write for the Özkan et al. (2015) model

$$\begin{cases} \alpha_j = \sum_{n=0}^{S_j-1} P(n), \\ \beta_j = \left(\sum_{k=0}^{S_0-1} Q(k)\right) p(S_j, l_j), \\ \delta_j = 1 - \alpha_j - \beta_j, \end{cases} \tag{20}$$

and for the Johansson and Olsson (2018) model

$$\begin{cases} \alpha_j = \sum_{n=0}^{S_j-1} P(n), \\ \delta_j = \left(\sum_{k=0}^{S_0-1} Q(k)\right) p(S_j, l_j) \\ \sum_{\mathbf{b}^j \in \Omega_j} \left(\sum_{k=0}^{S_j} \frac{(k+\sum_{i=1,i\neq j}^{J} b_i)!(\lambda_0 l_0)^{S_0}}{(S_0+k+\sum_{i=1,i\neq j}^{J} b_i)!} \frac{(\lambda_j l_0)^k}{k!} p(S_j - k, l_j - w_j)\right) \prod_{i=1,i\neq j}^{J} \frac{(\lambda_i l_0)^{b_i}}{b_i!} P(S_i - b_i, l_i - w_i)\right) \\ \beta_j = 1 - \alpha_j - \beta_j, \end{cases} \tag{21}$$

where $p(S_j - k, l_j - w_j)$ and $P(S_i - b_i, l_i - w_i)$ are given by Equation (15). We note that the fraction of lost demand δ_j does not take into account the pipeline state during w_j. In fact, each lost demand can be satisfied with an emergency replenishment with lead time w_j, which does not influence the fraction of lost demand in our system. The waiting time w_j here plays the role of l_j when considering the Alvarez and van der Heijden (2014) model (i.e., $l_j = w_j$). Indeed, lead time l_j does not influence the fraction of lost demand. Furthermore, Equation 21 reduces to the exact steady-state probability of having n orders in the pipeline to local warehouse i, when $l_j = w_j$, $j \in \{1, \ldots, J\}$.

We end this section by first noting that the calculation of $\mathbb{E}[I_j^+]$ and $\mathbb{E}[I_j^-]$ can be obtained directly from the expression of $P(n), n \geq 0$ in Equation (19) as $\mathbb{E}[I_j^+] = \sum_{k=0}^{S_j}(S_j - n)P(n)$ for both models, and $\mathbb{E}[I_j^-] = \sum_{k=S_j+1}^{\infty}(n - S_j)P(n) - w_j\lambda_j\delta_j$ for the Johansson and Olsson (2018) model. We do not have $\mathbb{E}[I_j^-]$ under Özkan et al. (2015), since demand is not backordered at local warehouses and, second, by summarizing the main results in Table 2.

Table 2: Summary of the main results. $Q(n)$,$P(n)$ represents the steady-state probability of having n items in the pipeline to the central warehouse and local warehouse, respectively. α_j, β_j and δ_j are the fraction of demand satisfied immediately from stock at local warehouse, after a delay and lost respectively. $E[I_j^+]$ is the expected stock on hand and $E[I_j^-]$ is the expected backorder level for the local warehouse j.

Performance measure	Alvarez and van der Heijden (2014) setting	Johansson and Olsson (2018) setting	Özkan et al. (2015) setting
	$w_j = l_j$	$w_j < l_j$	$w_j = 0$
$Q(n)$	Equation 4	Equation 18	Equation 17
$P(n)$	Equation 7	Equation 19	Equation 19
α_j	Equation 11	Equation 21	Equation 20
β_j	Equation 12	Equation 21	Equation 20
δ_j	Equation 10	Equation 21	Equation 20
$E[I_j^+]$	$\sum_{x=0}^{S_j} xP\{I_j = x\}$	$\sum_{k=0}^{S_j}(S_j - n)P(n)$	$\sum_{k=0}^{S_j}(S_j - n)P(n)$
$E[I_j^-]$	Equation 13	$\sum_{k=S_j+1}^{\infty}(n - S_j)P(n) - w_j\lambda_j\delta_j$	Does not apply

5 Numerical investigation

In this section, we compare our model to the two models considered and to an exact simulation model built in Rockwell Arena software. The numerical investigation is first conducted under deterministic lead times to enable the comparison with the benchmark models. Second, we conduct an experiment where stochastic lead times at the central warehouse are considered.

We start by comparing the CPU time to compute the steady-state probability of central warehouse $Q(n)$ with that of Alvarez and van der Heijden (2014), (i.e., Corollary 2) and our method (i.e., Corollary 1). Next, we compare our model to Özkan et al. (2015)'s model where there is no waiting time at the local warehouse. We then focus on evaluating the performance of our model and compare it to the Johansson and Olsson (2018) model. Waiting time is evaluated as a fraction of the minimum lead time of all local warehouses. Note that Johansson and Olsson (2018) only consider symmetric values, i.e., all local warehouses have the same parameters. We study a similar setting but also consider the non-symmetric case where values differ among the local warehouses.

We consider the case of two warehouses with the following input parameters: $l_0 \in \{2, 20\}$, $\{l_1, l_2\} \in \{1, 5\}$, $S_0 \in \{1, 5\}$, $\{S_1, S_2\} \in \{1, 2\}$, $\lambda_1 = 0.1$, $\lambda_2 \in \{0.2, 0.5, 1\}$ and $w \in \{0.25, 50, 0.75\} \times \min(l_1, l_2)$. Note that we have included other values of the parameters considered in Johansson and Olsson (2018) and Özkan et al. (2015), since – as will be shown – for the new values, the performance of these two models can be lower than that under the values presented in their papers. In fact, when we choose values where the lead times of the local warehouses vary by a factor of two or more, we are able to evaluate the similarities and differences in the performance measures of the two models and our model against the simulation. In line with the Özkan et al. (2015) model, we rely on the fraction of satisfied demand from the system and the fraction of lost demand as the main performance measures for comparison. For ease of reading, we denote with JO the model of Johansson and Olsson (2018) and OVS the model of Özkan et al. (2015).

Our setting enables us to make a comparison with the OVS model of 32 (48) instances in the symmetric (asymmetric) case. For the comparison with the JO model, we have 64 (192) symmetric (asymmetric) instances. The relevant results showing the performance of the different models are reported in Tables 3–6.

The computation time of $Q(n)$ using the method of Alvarez and van der Heijden (2014) and our method is reported in Table 3. The computational experiments are realized on a PC with Intel Core i7-11800H running at 2.3 GHz. The CPU used by our model is very reasonable compared to that of Alvarez and van der Heijden (2014). Indeed, their method suffers from the curse of dimensionality, as the computation time increases exponentially when S_j, $j \in \{1, \ldots, J\}$ increases or when the number of local warehouses increases.

Table 3: CPU time required for $Q(n)$ using the Alvarez and van der Heijden (2014) model and our model

S_j, $j = 1, \ldots, J$	CPU time in seconds in Alvarez and van der Heijden (2014) model	CPU time in seconds in our model
1	0	0
2	0.1	0
3	0.5	0
4	2.3	0.1
5	7.3	0.2
6	14.7	0.5
7	40.1	1.1
8	101	2.3
9	197.2	4.2
10	334.8	7.3

$S_0 = 3$, l_0=2, J=5, and $\lambda_j = 1, j = 1, \ldots, J$.

In the following, we start by reporting in Table 4 the comparative results of the different models in the symmetric case (i.e., identical local warehouses) for $w_j \leq 0.5$. The results for $w_j \geq 0.75$ are reported in Table 8 in Appendix 2 since the conclusions drawn from Table 4 also hold for Table 8. We mark in bold the cases where the difference in α_j, $j = 1, 2$ or δ_j, $j = 1, 2$ for the JO and OVS models is higher than 3% compared to the simulation. We believe that a minimum difference of 3% is a reasonable choice to compare the models. In the asymmetric case (i.e., non-identical local warehouses), the results when our model is compared to the simulation and the JO model are reported in Table 5. The comparison to the OVS model is reported in Table 6. Note that in Tables 5-6 to limit the length of the tables for presentation purposes and facilitate further the interpretation of the results, we only report the instances where the difference in the simulation is significant, exceeding 3%.

Considering identical warehouses, under the JO setting, we find 7 instances out of 64 where the difference, to the simulation model, exceeds 3%. Five of these 7 cases are listed in Table 4 (and 2 are listed in Table 8). Under OVS model setting, Table 4 shows that there are only 5 instances (7 instances) with a difference above 3% in our model (in the OVS model), which shows the accuracy of our model compared to the two benchmark models. Furthermore, the results in Table 4, show that the performance of our model is higher than that of OVS (i.e closer to simulation) especially when the central warehouse lead time l_0 is high compared to the local warehouse lead time l_j, $j \in \{1, \ldots, J\}$.

Table 4: Comparison under symmetric instances

Input parameters						Comparison with JO model						Comparison with OVS model					
						Simulation Model		Our Model		JO Model		Simulation Model		Our Model		OVS Model	
l_0	l_1	S_0	S_1/S_2	λ_1/λ_2	w	$\alpha_1(\%)$	$\delta_1(\%)$	$\alpha_1(\%)$	$\delta_1(\%)$	$\alpha_1(\%)$	$\delta_1(\%)$	$\alpha_1(\%)$	$\delta_1(\%)$	$\alpha_1(\%)$	$\delta_1(\%)$	$\alpha_1(\%)$	$\delta_1(\%)$
2	1	1	1	0.1	0	88.87	11.14	88.56	11.44	88.58	11.42	88.85	1.71	88.40	**6.15**	*88.23*	**6.15**
2	5	1	2	0.1	0	91.69	8.31	91.51	8.49	91.52	8.48	91.66	2.61	91.47	**5.19**	91.48	**5.19**
20	1	1	1	0.1	0	47.39	52.57	47.33	52.67	50.30	49.70	47.34	0.07	47.07	0.55	**52.93**	0.74
20	5	1	2	0.1	0	63.88	36.14	63.80	36.20	64.93	35.07	63.86	0.09	63.71	0.30	65.72	0.39
2	1	5	1	0.1	0	90.93	9.11	90.91	9.09	90.91	9.09	90.91	9.05	90.91	9.09	90.91	9.09
2	5	5	2	0.1	0	92.27	7.66	92.31	7.69	92.31	7.69	92.30	7.68	92.31	7.69	92.31	7.69
20	1	5	1	0.1	0	85.19	14.77	84.73	15.27	84.42	15.57	84.36	4.79	83.68	6.32	82.57	6.27
20	5	5	2	0.1	0	89.27	10.70	88.88	11.12	88.78	11.22	88.93	3.29	88.42	5.15	88.26	5.08
2	1	1	1	0.1	0.25	88.55	9.22	88.28	9.48	88.31	9.45						
20	1	1	1	0.1	0.25	46.63	52.20	46.58	52.24	**49.59**	**49.16**						
2	1	5	1	0.1	0.25	90.73	6.98	90.73	6.98	90.73	6.98						
20	1	5	1	0.1	0.25	84.49	13.33	84.13	13.74	83.83	14.05						
2	1	1	1	0.1	0.5	88.22	7.22	88.05	7.43	88.08	7.40						
20	1	1	1	0.1	0.5	45.90	51.77	45.85	51.80	**48.89**	**48.60**						
2	1	5	1	0.1	0.5	90.60	4.76	90.59	4.76	90.59	4.76						
20	1	5	1	0.1	0.5	83.84	11.87	83.55	12.17	83.24	12.49						
2	1	1	1	0.2	0	78.03	22.00	77.42	22.58	77.57	22.43	77.80	2.71	76.86	**7.95**	76.48	**7.96**
2	5	1	2	0.2	0	78.04	21.99	77.63	22.37	77.66	22.34	77.85	5.64	77.35	**9.50**	77.27	**9.49**
20	1	1	1	0.2	0	28.65	71.32	28.67	71.33	**32.00**	**68.00**	28.65	0.05	28.50	0.26	**37.72**	0.50
20	5	1	2	0.2	0	40.12	59.89	40.10	59.90	41.61	58.39	40.06	0.07	40.02	0.15	**43.89**	0.31
2	1	5	1	0.2	0	83.35	16.67	83.33	16.67	83.33	16.67	83.31	16.56	83.33	16.64	83.33	16.64
2	5	5	2	0.2	0	79.97	20.01	80.00	20.00	80.00	20.00	80.03	19.90	80.00	19.97	80.00	19.97
20	1	5	1	0.2	0	61.85	38.15	61.35	38.64	62.25	37.75	60.40	2.89	59.09	4.86	60.24	5.26
20	5	5	2	0.2	0	66.31	33.67	65.79	34.21	66.21	33.79	65.33	2.88	64.20	**4.95**	63.66	5.22
2	1	1	1	0.2	0.25	76.90	19.11	76.44	19.64	76.61	19.46						
20	1	1	1	0.2	0.25	27.56	71.02	27.55	71.04	**30.87**	**67.55**						
2	1	5	1	0.2	0.25	82.69	13.02	82.71	13.05	82.71	13.05						
20	1	5	1	0.2	0.25	59.82	37.15	59.42	37.53	60.49	36.41						
2	1	1	1	0.2	0.5	75.91	16.08	75.58	16.47	75.77	16.26						
20	1	1	1	0.2	0.5	26.46	70.72	26.48	70.74	**29.78**	**67.08**						
2	1	5	1	0.2	0.5	82.26	9.08	82.25	9.10	82.25	9.10						
20	1	5	1	0.2	0.5	57.79	36.19	57.54	36.41	58.78	35.04						

In addition to these observations, we find that when our performance decreases under a small lead time l_0, it also decreases for the OVS model. Even in this latter case, our model leads to a higher performance compared to OVS. We observe similar performance under the JO model setting, in several cases, especially when the lead time of the central warehouse is high, while our model leads to almost the same performance as the simulation. This negative effect of the central warehouse lead time on the performance in both JO and OVS models is due to their modeling approach. In fact, both models rely on estimating a common waiting time for backordered demands at the central warehouse to calculate the fractions of accepted or lost demand and they do not differentiate this waiting time for the different demand levels experienced at the local warehouses. Therefore, considering the same waiting time for all local warehouses has a negative impact on the performance, especially when the lead time increases. This also increases the backorders, which affects the accepted or lost demand fractions. In contrast, in our approach, we do not rely on the waiting time at the central warehouse, and we are able to derive the accepted or lost demand fractions in a distinct way because we rely on the backorder levels at the central warehouse, B_0^j. The backorder levels are obviously different for non-identical local warehouses, which explains why the performance of JO and OVS is poor in this specific case. Our model is thus a good alternative to the two benchmark models, and at least equivalent or better than the OVS and JO models in most cases.

In the case of non-identical local warehouses, for the comparison to the JO model, out of the 192

Table 5: Comparison of our model, the Johansson and Olsson (2018) model, and the simulation under asymmetric instances

Input parameters									Simulation				Our model				*JO* Model			
l_0	l_1	l_2	S_0	S_1	S_2	λ_1	λ_2	w	$\alpha_1(\%)$	$\alpha_2(\%)$	$\delta_1(\%)$	$\delta_2(\%)$	$\alpha_1(\%)$	$\alpha_2(\%)$	$\delta_1(\%)$	$\delta_2(\%)$	$\alpha_1(\%)$	$\alpha_2(\%)$	$\delta_1(\%)$	$\delta_2(\%)$
20	1	1	1	1	1	0.1	0.2	0	44.87	30.32	55.13	69.68	44.92	30.27	55.08	69.73	**47.87**	**33.81**	**52.13**	**66.19**
2	1	1	1	1	1	0.1	0.5	0	86.18	60.64	13.82	39.36	86.39	**57.24**	13.61	**42.76**	86.56	**57.61**	13.44	**42.39**
2	5	1	1	2	1	0.1	0.5	0	90.66	60.51	9.34	39.48	90.80	**57.19**	9.20	**42.81**	90.85	57.54	9.15	42.46
20	1	1	1	1	1	0.1	0.5	0	42.81	14.44	57.19	85.56	43.04	14.37	56.95	85.63	45.58	**17.46**	54.42	**82.54**
2	1	1	1	1	1	0.1	1	0	84.56	42.27	15.44	57.73	85.39	**38.77**	14.61	**61.22**	**88.32**	**11.17**	**11.68**	55.31
2	5	1	1	2	1	0.1	1	0	90.34	42.18	9.66	57.82	90.47	**38.73**	9.53	**61.27**	91.43	**11.41**	8.57	55.44
20	1	1	1	1	1	0.1	1	0	42.41	7.66	57.59	92.34	42.37	7.64	57.63	92.36	**47.34**	7.18	**52.66**	**88.21**
20	5	1	1	2	1	0.1	1	0	62.10	7.30	37.90	92.70	62.07	7.26	37.93	92.74	64.50	7.05	35.50	**89.59**
2	1	1	5	1	1	0.1	1	0	90.97	49.94	9.03	50.05	90.89	49.95	9.11	50.05	90.91	**0.00**	9.09	50.00
2	5	1	5	2	1	0.1	1	0	92.22	49.94	7.78	50.05	92.30	49.95	7.70	50.05	92.31	**0.00**	7.69	50.00
20	1	1	5	1	1	0.1	1	0	67.13	21.88	32.87	78.12	67.00	21.51	33.00	78.49	**89.67**	**1.15**	**10.33**	**51.59**
20	5	1	5	2	1	0.1	1	0	80.76	21.14	19.24	78.86	80.46	20.82	19.54	79.18	**91.68**	**1.27**	**8.32**	**51.77**
20	1	1	1	1	1	0.1	0.2	0.25	44.10	29.29	54.83	69.26	44.15	29.13	54.73	69.38	**47.11**	**32.68**	**51.69**	**65.65**
20	5	1	1	2	1	0.1	0.2	0.25	62.85	27.92	36.13	70.67	62.93	27.88	36.09	70.69	64.38	30.82	34.66	**67.60**
2	1	1	1	1	1	0.1	0.5	0.25	85.59	56.79	12.25	35.71	85.90	54.08	11.92	**38.72**	86.10	54.51	11.72	38.23
2	5	1	1	2	1	0.1	0.5	0.25	90.41	56.71	8.75	35.73	90.53	54.05	8.67	**38.75**	90.59	54.46	8.61	38.29
20	1	1	1	1	1	0.1	0.5	0.25	42.15	12.84	56.80	85.41	42.27	12.85	56.66	85.44	44.74	15.80	54.13	**82.10**
20	1	1	5	1	1	0.1	0.5	0.25	69.41	35.38	28.86	59.89	69.36	34.77	28.89	60.60	**63.94**	**30.70**	**34.44**	**65.21**
2	1	1	1	1	1	0.1	1	0.25	83.91	35.42	13.95	54.47	84.75	33.01	13.11	**57.61**	**87.82**	**10.41**	**9.95**	**50.10**
2	5	1	1	2	1	0.1	1	0.25	89.71	35.37	9.44	54.54	90.14	32.99	9.04	**57.64**	91.16	**10.53**	8.05	**50.20**
20	1	1	1	1	1	0.1	1	0.25	41.38	6.05	57.53	92.24	41.60	6.03	57.35	92.26	**46.47**	5.76	**52.35**	**87.58**
20	5	1	1	2	1	0.1	1	0.25	61.29	5.73	37.69	92.64	61.39	5.73	37.63	92.64	63.79	5.63	35.24	**89.08**
2	1	1	5	1	1	0.1	1	0.25	90.75	44.49	6.96	42.87	90.69	44.42	7.01	42.96	90.73	**0.00**	6.98	42.86
2	5	1	5	2	1	0.1	1	0.25	92.28	44.47	7.02	42.90	92.13	44.42	7.12	42.96	92.14	**0.00**	7.10	42.86
20	1	1	5	1	1	0.1	1	0.25	65.31	17.35	33.07	77.70	65.45	17.15	32.90	77.98	**89.26**	**1.10**	**8.48**	**45.07**
20	5	1	5	2	1	0.1	1	0.25	79.81	16.77	19.20	78.47	79.58	16.61	19.47	78.68	**91.44**	**1.14**	**7.79**	**45.16**
20	1	1	1	1	1	0.1	0.2	0.5	43.46	28.10	54.35	68.90	43.40	28.03	54.38	69.03	46.37	**31.59**	**51.25**	**65.09**
20	5	1	1	2	1	0.1	0.2	0.5	62.55	26.95	35.49	70.29	62.28	26.84	35.77	70.33	63.74	29.80	34.34	**67.07**
20	1	1	1	1	1	0.1	0.5	0.5	41.54	11.54	56.41	85.21	41.51	11.49	56.36	85.24	43.91	14.32	53.84	**81.62**
20	5	1	1	2	1	0.1	0.5	0.5	61.24	10.92	36.74	85.94	61.13	10.95	36.91	85.94	62.30	13.33	35.76	**82.88**
20	1	1	5	1	1	0.1	0.5	0.5	67.86	31.85	28.66	59.13	67.88	31.47	28.64	59.59	**55.73**	**21.26**	**41.41**	**72.70**
20	5	1	5	2	1	0.1	0.5	0.5	81.01	30.90	17.17	60.34	80.69	30.67	17.47	60.62	**77.12**	**27.47**	**21.01**	**64.73**
2	1	1	1	1	1	0.1	1	0.5	83.57	29.79	12.29	50.88	84.08	28.30	11.61	53.35	**87.28**	**9.91**	**8.24**	**43.64**
2	5	1	1	2	1	0.1	1	0.5	89.62	29.73	8.76	50.93	89.81	28.29	8.57	53.36	90.88	**9.95**	7.56	**43.68**
20	1	1	1	1	1	0.1	1	0.5	40.83	4.73	57.11	92.14	40.84	4.76	57.06	92.16	**45.61**	4.62	**52.05**	**86.81**
20	5	1	1	2	1	0.1	1	0.5	60.63	4.54	37.36	92.52	60.71	4.53	37.32	92.53	63.09	4.51	34.98	**88.46**
2	1	1	5	1	1	0.1	1	0.5	90.59	40.47	4.72	33.37	90.53	40.29	4.82	33.57	90.59	**0.00**	4.76	33.33
2	5	1	5	2	1	0.1	1	0.5	92.00	40.41	6.56	33.37	91.96	40.30	6.54	33.56	91.98	**0.00**	6.53	33.33
20	1	1	5	1	1	0.1	1	0.5	63.83	13.75	32.92	77.31	63.90	13.66	32.82	77.48	**88.80**	**1.09**	**6.65**	**36.57**
20	5	1	5	2	1	0.1	1	0.5	78.95	13.31	19.18	78.02	78.69	13.23	19.41	78.18	**91.19**	**1.06**	**7.29**	**36.44**
20	1	1	1	1	1	0.1	0.2	0.75	42.57	27.08	54.09	68.56	42.66	26.97	54.02	68.66	**45.64**	**30.55**	**50.80**	**64.50**
20	5	1	1	2	1	0.1	0.2	0.75	61.71	25.89	35.34	69.91	61.63	25.85	35.45	69.97	63.11	28.82	34.02	**66.51**
20	1	1	1	1	1	0.1	0.5	0.75	40.50	10.27	56.26	85.02	40.76	10.28	56.06	85.05	43.07	13.00	53.58	**81.09**
20	5	1	1	2	1	0.1	0.5	0.75	60.43	9.78	36.55	85.75	60.46	9.80	36.60	85.75	61.59	12.10	35.52	**82.40**
20	1	1	5	1	1	0.1	0.5	0.75	66.60	28.61	28.31	58.36	66.40	28.47	28.43	58.58	**49.84**	**15.32**	**46.28**	**77.71**
20	5	1	5	2	1	0.1	0.5	0.75	79.80	27.79	17.33	59.49	79.84	27.76	17.39	59.61	**71.72**	**18.36**	**25.36**	**73.28**
2	1	1	1	1	1	0.1	1	0.75	83.03	25.09	10.34	46.84	83.38	24.44	10.13	48.27	**86.67**	**9.68**	**6.58**	**35.46**
2	5	1	1	2	1	0.1	1	0.75	89.30	25.04	8.27	46.83	89.46	24.44	8.12	48.26	90.57	**9.65**	7.12	**35.42**
20	1	1	1	1	1	0.1	1	0.75	40.10	3.75	56.79	92.05	40.10	3.76	56.77	92.05	**44.74**	3.72	**51.78**	**85.81**
20	5	1	1	2	1	0.1	1	0.75	60.10	3.58	36.90	92.44	60.04	3.58	37.02	92.43	62.38	3.62	34.74	**87.65**
2	1	1	5	1	1	0.1	1	0.75	90.29	37.72	2.50	20.11	90.40	37.51	2.56	20.58	90.51	**0.00**	2.44	20.00
2	5	1	5	2	1	0.1	1	0.75	91.73	37.76	6.03	20.11	91.80	37.52	5.99	20.57	91.84	**0.00**	5.96	20.00
20	1	1	5	1	1	0.1	1	0.75	62.30	10.88	32.78	76.94	62.39	10.87	32.76	76.99	**88.22**	**1.14**	**4.91**	**25.10**
20	5	1	5	2	1	0.1	1	0.75	77.79	10.56	19.27	77.64	77.81	10.53	19.34	77.70	**90.91**	**1.01**	**6.84**	**24.48**
20	5	5	5	2	2	0.1	1	3.75	70.33	1.00	15.80	73.94	70.58	1.03	15.76	73.98	**58.59**	0.49	**27.06**	**84.63**

cases, Table 5 shows that there are about 27% of cases (i.e., 52 instances) where the *JO* model shows a difference in α_j, $j = 1, 2$ or in δ_j, $j = 1, 2$ higher than 3% compared to the simulation (cases marked in bold font), whereas in our model, only 4% (i.e., 8 instances) of cases have a difference greater than 3%. Note that under *JO* model, the difference to the simulation can reach 53% for the fraction of demand lost. Similarly, when comparing our model with that of OVS, the results in Table 6 reveal that for the OVS's model, about 50% of the cases (i.e., 28 instances) have a difference in α_j, $j = 1, 2$ or in δ_j, $j = 1, 2$ higher than 3%, whereas in our model, we find only 8 cases where the difference exceeds 3%. This demonstrates that our approximations of α_j and δ_j are very close to the simulation model. It should be noted however that our model and OVS underestimate the fraction of demand

Table 6: Comparison of our model, Özkan et al. (2015)'s model, and the simulation under asymmetric instances

Input parameters								Simulation Model				Our Model				*OVS* model			
l_0	l_1	l_2	S_0	S_1	S_1	λ_1	λ_2	$\alpha_1(\%)$	$\alpha_2(\%)$	$\delta_1(\%)$	$\delta_1(\%)$	$\alpha_1(\%)$	$\alpha_2(\%)$	$\delta_1(\%)$	$\delta_1(\%)$	$\alpha_1(\%)$	$\alpha_2(\%)$	$\delta_1(\%)$	$\delta_1(\%)$
2	1	1	1	1	1	0.1	0.2	87.92	79.11	10.48	17.94	87.48	78.04	**7.38**	**12.52**	87.35	77.53	7.52	**13.04**
2	1	5	1	1	2	0.1	0.2	87.85	78.46	10.56	15.26	87.46	77.85	**7.40**	**10.85**	87.20	77.78	7.67	**10.93**
2	5	1	1	2	1	0.1	0.2	91.33	79.06	6.29	17.99	91.17	78.02	4.49	**12.57**	91.18	77.29	4.48	**13.32**
2	5	5	1	2	2	0.1	0.2	91.35	78.41	6.29	15.32	91.16	77.84	4.50	**10.89**	91.18	77.76	4.49	**10.97**
20	1	1	1	1	1	0.1	0.2	44.85	30.32	55.11	69.60	44.70	30.02	55.04	69.50	**53.54**	**36.56**	**46.02**	**62.65**
20	1	5	1	1	2	0.1	0.2	42.90	42.27	57.06	57.58	42.83	42.16	56.98	57.43	**47.48**	**46.93**	**52.20**	**52.38**
20	5	1	1	2	1	0.1	0.2	63.57	28.98	36.37	70.96	63.48	28.76	36.34	70.86	**68.52**	30.82	**31.18**	68.54
20	5	5	1	2	2	0.1	0.2	62.14	41.31	37.82	58.57	62.12	41.18	37.76	58.50	**65.51**	43.58	**34.28**	55.87
20	1	1	5	1	1	0.1	0.2	78.09	66.89	19.27	28.08	77.20	65.40	18.69	27.06	77.67	**63.49**	18.15	28.84
20	1	5	5	1	2	0.1	0.2	77.01	70.62	20.47	24.10	75.91	69.19	20.20	22.25	**73.78**	68.75	22.32	22.65
20	5	1	5	2	1	0.1	0.2	86.11	65.84	12.09	29.31	85.52	64.43	11.12	28.29	86.12	**58.74**	10.51	**33.96**
2	1	1	1	1	1	0.1	0.5	86.14	57.98	12.58	37.07	85.71	55.82	10.99	**32.09**	85.84	**54.81**	10.83	**33.01**
2	1	5	1	1	2	0.1	0.5	86.17	50.04	12.50	38.16	85.80	48.96	10.83	**33.56**	85.55	48.73	11.08	**33.78**
2	5	1	1	2	1	0.1	0.5	90.72	57.92	7.48	37.16	90.58	55.79	6.64	**32.17**	90.62	**54.03**	6.59	**33.87**
2	5	5	1	2	2	0.1	0.5	90.77	49.99	7.40	38.22	90.61	48.95	6.55	**33.63**	90.63	48.67	6.53	**33.90**
20	1	1	1	1	1	0.1	0.5	42.97	14.35	57.01	85.57	42.87	14.18	57.05	85.52	**54.60**	**19.39**	**45.21**	**79.94**
20	1	5	1	1	2	0.1	0.5	41.22	19.85	58.77	80.00	41.16	19.75	58.78	79.96	**48.11**	**23.50**	**51.76**	**75.86**
20	5	1	1	2	1	0.1	0.5	62.47	13.66	37.50	86.28	62.39	13.53	37.55	86.23	**68.61**	15.17	**31.26**	84.26
20	5	5	1	2	2	0.1	0.5	61.11	19.33	38.87	80.55	61.07	19.25	38.89	80.53	**65.60**	21.09	**34.32**	78.39
20	1	1	5	1	1	0.1	0.5	69.39	37.09	29.84	59.67	68.61	35.30	29.88	59.16	**73.12**	35.23	**25.09**	58.18
20	5	1	5	2	1	0.1	0.5	81.81	36.06	17.52	60.88	81.31	34.38	17.48	60.40	83.29	**30.07**	15.28	63.72
2	5	1	1	2	1	0.1	1	90.28	38.75	8.46	55.54	90.13	36.66	8.22	52.60	90.21	**34.88**	8.08	54.04
20	1	1	1	1	1	0.1	1	42.24	7.59	57.75	92.34	42.23	7.51	57.74	92.32	**55.29**	**11.00**	**44.62**	**88.49**
20	1	5	1	1	2	0.1	1	40.72	10.41	59.27	89.46	40.66	10.37	59.32	89.46	**48.45**	12.74	**51.49**	86.80
20	5	1	1	2	1	0.1	1	62.00	7.23	37.98	92.72	62.00	7.16	37.98	92.70	**68.70**	8.24	**31.23**	91.32
20	5	5	1	2	2	0.1	1	60.74	10.13	39.25	89.77	60.74	10.10	39.24	89.76	**65.68**	11.26	**34.27**	88.36
20	1	1	5	1	1	0.1	1	65.56	20.31	34.13	77.66	65.06	19.19	34.33	77.46	**71.70**	20.22	**27.44**	75.05
20	5	1	5	2	1	0.1	1	79.87	19.66	19.79	78.45	79.57	18.62	19.95	78.26	82.23	**16.45**	17.07	79.03

lost for all instances considered. This underestimation is much larger in the OVS model than in ours. This can be explained by an overestimation of the fraction of the demand satisfied from the central warehouse β_j (since $\beta_j = 1 - \alpha_j - \delta_j$). This overestimation is due to the Poisson assumption arrival for the orders in the pipeline between the central warehouse and local warehouse. In fact, such an assumption is valid if the central warehouse has an infinite stock, which is not always true, making the arrival rate less than λ_j and consequently a higher β_j. Moreover, our model outperforms the OVS model in estimating the fraction of demand satisfied with stock on hand. Hence, our model offers a better alternative to the two models developed in the literature, and the results show that our approximations are very close to the simulation.

To conclude the numerical investigation, we conduct an experiment where we consider an inventory system composed of two local warehouses with stochastic lead times at the central warehouse. To analyze the sensitivity to the lead time distribution, we focus on the case where the waiting time threshold is less than the local warehouse lead time. We assume that the central warehouse lead time follows an Erlang distribution with a shape parameter k and a scale parameter k/l_j. Erlangian lead times have been considered in the inventory literature (Johansen, 2005). The Erlang distribution is characterized with a high modelling flexibility, since for $k = 1$, it is equivalent to the exponential distribution and when k increases the distribution becomes less variable, and it tends toward the deterministic distribution for very high values of k. For the purpose of the numerical analysis, we consider a shape parameter $k \in \{2, 4, 8, 16, 32, 64, 128, 256, 512\}$. For the other parameters, we use the same values of the 192 instances presented earlier in this paper, i.e., $\lambda_1 = 0.1$, $\lambda_2 \in \{0.2, 0.5, 1\}$,

$\{l_1, l_2\} \in \{1, 5\}$, $l_0 \in \{2, 20\}$ and $w \in \{0.25, 50, 0.75\} \times \min(l_1, l_2)$. This gives a total of 1728 instances. The percent difference of the performance measures obtained with our model and the simulation are reported in Table 7. A negative value means that our model underestimates the result obtained with the simulation. We report for each value of k, the minimum, the maximum and the average value of the percent difference of α_j and δ_j $(j = 1, 2)$ obtained over the 192 instances. From Table 7, it is clear that our model works very well for a stochastic central warehouse lead time. The error on the percent difference of α_j and δ_j with respect to simulation does not exceed 3.5%, regardless of the shape parameter k of the Erlang distribution of the lead time. Thus, the results obtained under a deterministic lead time are also valid for a stochastic lead time.

Table 7: percent difference of the performance obtained with our model and the simulation under Erlang distributed lead times with shape k

	α_1			α_2			δ_1			δ_2		
k	min (%)	average (%)	max (%)	min (%)	average (%)	max (%)	min (%)	average (%)	max (%)	min (%)	average (%)	max (%)
2	-0.43	0.01	0.57	-0.21	0.27	1.79	-0.57	-0.02	0.47	-1.79	-0.37	0.11
4	-0.56	0.01	0.61	-0.32	0.31	2.16	-0.64	-0.02	0.49	-2.16	-0.44	0.11
8	-0.56	0.01	0.82	-0.39	0.34	2.57	-0.78	-0.01	0.54	-2.57	-0.48	0.11
16	-0.59	0.02	0.87	-0.42	0.38	2.90	-0.87	-0.03	0.58	-2.90	-0.52	0.11
32	-0.72	0.02	1.09	-0.40	0.40	3.12	-1.09	-0.03	0.74	-3.12	-0.55	0.11
64	-0.72	0.03	1.02	-0.39	0.42	3.31	-1.02	-0.03	0.72	-3.31	-0.58	0.11
128	-0.69	0.04	1.03	-0.39	0.43	3.42	-1.03	-0.04	0.70	-3.42	-0.60	0.11
256	-0.77	0.04	1.11	-0.39	0.43	3.44	-1.10	-0.05	0.77	-3.44	-0.61	0.11
512	-0.79	0.05	1.04	-0.34	0.45	3.44	-1.04	-0.05	0.76	-3.44	-0.61	0.11

The results in Table 7 show that our model performs very well compared to the simulation model in most of the cases. The average error on the percent difference of α_j and δ_j increases as the variability of the lead time decreases and these deviations over the 1728 cases does not exceed 1%. For the maximum and minimum observed values, there are few cases where the absolute deviation value reaches 3.5%. This shows that the approximation proposed in our model is of a high quality.

6 Conclusion

We have provided an exact analysis of a two-echelon inventory with general lead times and Poisson demand with network lost sales in the case of waiting time at the local warehouse equal to the lead time between the central and local warehouses. We have found exact expressions for the steady-state distribution of the number of orders in the pipeline to the central warehouse, using this result to find the fraction of demand satisfied directly from local warehouses and the expected number of backorders and on-hand inventory at each local warehouse. We have shown that the steady-state probability of having n orders in the pipeline to the central warehouse is insensitive to the lead time distribution, except through its mean, implying that the expression of the steady-state probability derived in Alvarez and van der Heijden (2014) under constant lead times is exact for any lead time distribution.

When the waiting time is less than the lead time between the central and local warehouse, we derive an approximate expression for the different performance measures and show that our models are a good alternative to the two benchmark models in the literature. Our results fit within the broader literature using stochastic network representations to study complex inventory systems. We have used different techniques from this literature to illustrate the course of the analyses.

Our numerical results show that our model's performance measures are very close to the simulation in the majority of instances considered. In particular, when waiting time at the local warehouse is less than the lead time from the central warehouse, our model outperforms the benchmark models in most cases.

The findings of this work provide some insights to managers into the impact of the lead times' uncertainty on inventory systems under lost sales. Since there is an empirical evidence that replenishment lead times are uncertain in real inventory systems due to supply and transportation issues (Boute et al., 2007; Jakšič et al., 2011; Babai et al., 2022), managers can use our models and performance evaluation solutions with any lead time distribution. Moreover, most of the findings of this paper show that, regardless of the relationship between the waiting time threshold and the lead time, the inventory levels are not very sensitive to the lead time's uncertainty and therefore, managers can rely on the deterministic assumption to set the base stock level. Finally, in the particular case where lead times at local warehouses are less than the waiting time threshold, managers should benefit from our proposed approximation solutions, especially when the lead times differ much between local warehouses (i.e. the case of asymmetric inventories at the local warehouses). In fact, we show that in this case the performance can be improved considerably and the benefit can reach 53% for the fraction of demand lost compared to solutions provided in the literature.

The analysis in this paper has been performed under the assumption of Poisson demand. It would be interesting to extend this analysis to the case of compound Poisson demand, since demand for spare parts is often characterized by lumpiness that is better modeled with a compound Poisson process rather than a Poisson process (Lengu et al., 2014; Turrini and Meissner, 2019). The steady-state probability of the number of orders in the pipeline to the central warehouse is still valid under compound Poisson demand. The challenge, however, comes from finding the steady-state probability of the backorders at the central warehouse. This steady-state probability is no longer a hypergeometric distribution because demand arrives in batches. This is therefore an interesting avenue for future research.

7 Appendix

7.1 Appendix 1: Proof of Proposition 1

Proof. An order from local warehouse $j, j \in \{1, ..., J\}$ cannot be admitted in the pipeline if the total number of backorders already outstanding reaches S_j. As the maximum number of outstanding orders in the pipeline of the central warehouse that are associated with local warehouse $j \in \{1, ..., J\}$ cannot exceed $S_0 + S_j$, the number of orders in the pipeline to the central warehouse can be viewed as a queue shared by J customer classes. Each class arrives according to a Poisson process with a rate λ_j, $j \in \{1, \ldots, J\}$. If we let each warehouse correspond to a class, than one way to observe the system is to say that each class j has S_j dedicated servers and shares with all other classes the remaining S_0 servers. Therefore, in Theorem 1 of Kaufman (1981), the steady-state probability of being in state $\mathbf{n} \in \mathcal{S}$ is given by:

$$\lim_{t \to \infty} \mathbb{P}(N_1(t) = n_1, \ldots, N_J(t) = n_J) = C \prod_{i=1}^{J} \frac{(\lambda_i l_0)^{n_i}}{n_i!}, \quad \mathbf{n} \in \mathcal{S}, \tag{22}$$

where C is the normalizing constant. An example of the transition in the state space $\mathcal{S}$ in the case of two local warehouses where L follows an exponential distribution with mean $l_0 = 1/\mu$ is given in Figure 5

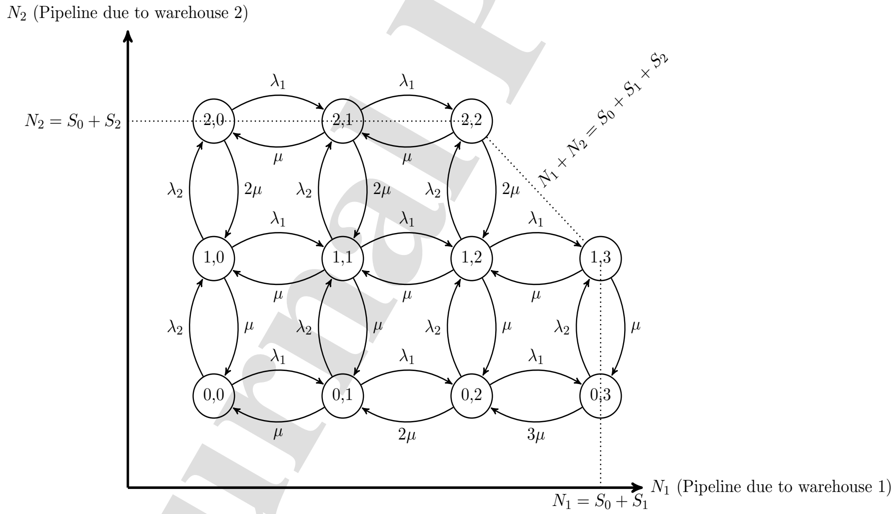

Figure 5: Markov chain transition between states of $\mathcal{S}$ under two local warehouses with $S_0 = 1$; $S_1 = 1$ and $S_2 = 2$

Remaining to be found is the conditional probability of having $b_j, j \in \{1, ..., J\}$ in a given pipeline $\mathbf{n} \in \mathcal{S}(n), n > S_0$. To do so, we consider for simplicity two local warehouses and assume that the system state is $(n_1, n_2), n_1 + n_2 > S_0$, in the steady-state, the probability of server 1 processing class 1 is $n_1/(n_1 + n_2)$ and class 2 $n_2/(n_1 + n_2)$. The allocation of server 1 will influence the type of class that the subsequent servers will handle. For example, if server 1 processes class 1, server 2 will process class 1 with probability $(n_1 - 1)/(n_1 + n_2 - 1)$ and class 2 with probability $n_2/(n_1 + n_2 - 1)$, and so forth, up to server $n_1 + n_2 - S_0$. This is a form of sampling without replacement. The probability that exactly b_j backorders of a particular type are in service when there are n_1 and n_2 orders in the central warehouse pipeline is given by the probability mass function of the hypergeometric distribution. Note that this approach is similar to the binomial desegregation introduced in Simon (1971). Furthermore, a hypergeometric desegregation of backorders is also reported in Boucherie et al. (2018) modeling a two-echelon spare parts network with lateral and emergency shipment by a queuing network with overflow by-passes similar to Song and Zipkin (2009). They observe the conditional probability that the number of backorders at the central warehouse when the local warehouse is out of stock follows a hypergeometric distribution. Moreover, a similar result is found in multi-class discrete-time queuing systems under the FCFS service discipline and in $M/M/s$ with time varying arrival and service rates (see, De Clercq et al. (2013) and Ingolfsson (2005)). Finally, Lefèvre (1982) finds that the number of backorders in a two echelon inventory system with a state dependent-arrival rate also forms a hypergeometric distribution.

Using the multivariate hypergeometric distribution, we can write

$$\mathbb{P}\{\mathbf{b}|\mathbf{n}\} = \lim_{t\to\infty} \mathbb{P}\left\{B_0^1(t) = b_1, \ldots, B_0^J(t) = b_J \mid N_1(t) = n_1, \ldots, N_J(t) = n_J\right\} = \frac{\binom{n_1}{b_1}\binom{n_2}{b_2}\cdots\binom{n_J}{b_J}}{\binom{\sum_{j=1}^J n_j}{\sum_{j=1}^J b_j}}, \quad \mathbf{b} \in \Omega, \quad \mathbf{n} \in \mathcal{S}(n) \tag{23}$$

Therefore, the steady-state joint probability is given by

$$\mathbb{P}\{B_0 = \mathbf{b}, N = \mathbf{n}\} := \lim_{t\to\infty} \mathbb{P}\left\{B_0^1(t) = b_1, \ldots, B_0^J(t) = b_J, N_1(t) = n_1, \ldots, N_J(t) = n_J\right\},$$

so that

$$\pi(B_0 = \mathbf{b}, N = \mathbf{n}) = \mathbb{P}\left\{B_0^1 = b_1, \ldots, B_0^J, N_1 = n_1, \ldots, N_J = n_J\right\} \quad = \quad \frac{\binom{n_1}{b_1}\binom{n_2}{b_2}\cdots\binom{n_J}{b_J}}{\binom{\sum_{j=1}^J n_j}{\sum_{j=1}^J b_j}} C \prod_{i=1}^J \frac{(\lambda_i l_0)^{n_i}}{n_i!},$$

which completes the proof. ■

7.2 Appendix 2: Proof of Corollary 1

Proof. Let $n = \sum_{j=1}^J n_j$ and $n - S_0 = \sum_{j=1}^J b_j$, then$\sum_{j=1}^J n_j - b_j = S_0$, so with Proposition 1, we have

$$Q(n) = \sum_{\sum_{j=1}^J b_j = n - S_0, \sum_{j=1}^J n_j - b_j = S_0} \mathbb{P}\left\{B_0^1 = b_1, \ldots, B_0^J = b_J, N_1 = n_1, \ldots, N_J = n_j\right\}, \tag{24}$$

$$= C\frac{S_0!n!}{(S_0+n)!} \sum_{\sum_{j=1}^J b_j = n - S_0, \sum_{j=1}^J n_j - b_j = S_0} \prod_{i=1}^J \frac{(\lambda_i l_0)^{b_i}}{b_i!} \prod_{i=1}^J \frac{(\lambda_i l_0)^{n_i - b_i}}{(n_i - b_i)!}, \tag{25}$$

$$= C\frac{S_0!n!}{(S_0+n)!} \sum_{\sum_{j=1}^J b_j = n - S_0,} \prod_{i=1}^J \frac{(\lambda_i l_0)^{b_i}}{b_i!} \sum_{\sum_{j=1}^J n_j - b_j = S_0} \prod_{i=1}^J \frac{(\lambda_i l_0)^{n_i - b_i}}{(n_i - b_i)!}.$$

But $\sum_{\sum_{j=1}^{J} n_j - b_j = S_0} \prod_{i=1}^{J} \frac{(\lambda_i l_0)^{n_i - b_i}}{(n_i - b_i)!} = \frac{(\lambda_0 l_0)^{S_0}}{S_0!}$, therefore,

$$Q(n) = C \frac{S_0! n!}{(S_0 + n)!} \frac{(\lambda_0 l_0)^{S_0}}{S_0!} \sum_{\sum_{j=1}^{J} b_j = n - S_0,} \prod_{i=1}^{J} \frac{(\lambda_i l_0)^{b_i}}{b_i!}, \quad (b_1, b_2, \ldots, b_J) \in \Omega(n). \tag{26}$$

for $(b_1, \ldots, b_J)$ such that $\sum_{j=1}^{J} b_j \leq \sum_{j=1}^{J} S_j$ and $b_j \leq S_j$.

Note that Equations 24 and 26 link the number of backorders, $\mathbf{B_0}$, to the number of orders in the pipeline to the central warehouse $\mathbf{n}$. In addition to the steady-state probability $Q(n), n \geq 0$, these equations also provide the number of backorders B_0^j in the pipeline to the central warehouse. The convolution of all $B_0^j, j = 1, \ldots, J$ is sufficient to find $Q(n), n \geq 0$, independently of the number of orders in the pipeline $\mathbf{n}$. ■

7.3 Appendix 3: Proof of Corollary 2

Proof. We wish to compute the steady-state probability of having n backorders at the central warehouse using the method of Alvarez and van der Heijden (2014). To do so, we first need to compute the steady-state probability of the Markov chain plotted in their Figure 2. We denote with $B_0^j,\ j = 1, \ldots, J$, the random variables that represent the number of backorders at the central warehouse that are associated with local warehouse j in steady-state. Alvarez and van der Heijden (2014) assume that the vector $(B_0^1, \ldots, B_0^J)$ has a truncated multinomial distribution with mass parameter n and probability vector $\left(\frac{\lambda_1}{\lambda_0}, \ldots, \frac{\lambda_J}{\lambda_0}\right)$. Let the vector $(S_1, \ldots, S_J)$ be a set of integers representing the maximum number of backorders associated with local warehouse $j, j = 1, ..., J$.

According to Alvarez and van der Heijden (2014), the steady-state probability $Q(n)$ of backorders is

$$Q(n) = \begin{cases} \frac{(\lambda_0 L_0)^n}{n!}, n \leq S_0, \\ \frac{(L_0)^n}{(n)!} \prod_{y=0}^{n-1} M(y),\ n > S_0, \end{cases} \tag{27}$$

where

$$M(n) = \begin{cases} \lambda_0, & n \leq S_0 \\ \sum_{j=1}^{J} \lambda_j \left(1 - \mathbb{P}\left\{B_0^j = S_j | n - S_0\right\}\right), & n > S_0 \end{cases}. \tag{28}$$

To compute $Q(n), n > S_0$, we need to express $\mathbb{P}\left\{B_0^j = S_j | n - S_0, n > S_0\right\}$ in a closed form. Henceforth, we express this probability in a simpler form. Since $(B_0^1, \ldots, B_0^J)$ has the multinomial distribution with mass parameter n and probability vector $\left(\frac{\lambda_1}{\lambda_0}, \ldots, \frac{\lambda_J}{\lambda_0}\right)$, Alvarez and van der Heijden (2014) show that

$$\mathbb{P}\left\{B_0^j = x_j | n - S_0\right\} = \frac{\mathbb{P}\left\{B_0^1 \leq x_1, \ldots, B_0^j \leq x_j, \ldots, B_0^J \leq x_J | n - S_0\right\}}{\mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | n - S_0\right\}} - \frac{\mathbb{P}\left\{B_0^1 \leq x_1, \ldots, B_0^j \leq x_j - 1, \ldots, B_0^J \leq x_J | n - S_0\right\}}{\mathbb{P}\left\{B_0^1 \leq S_1, , \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | n - S_0\right\}} \quad (29)$$

and we can thus write

$$M(n) = \sum_{j=1}^{J} \lambda_j \left(1 - \mathbb{P}\left\{B_0^j = S_j \mid n - S_0\right\}\right) = \frac{\sum_{j=1}^{J} \lambda_j \mathbb{P}\left\{B_0^1 \leq x_1, \ldots, B_0^j \leq S_j - 1, \ldots, B_0^J \leq x_J | n - S_0\right\}}{\mathbb{P}\left\{B_0^1 \leq S_1, , \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | n - S_0\right\}}. \quad (30)$$

It is however complex to use this formula to compute $Q(n)$ for very high J. We next simplify the expression $Q(n)$. From Frey (2009), we can write:

$$\mathbb{P}\left\{B_0^1 \leq x_1, B_0^2 \leq x_2, \ldots, B_0^J \leq x_J | n - S_0\right\} = (n - S_0)! \sum_{i_1=0}^{x_1} \sum_{i_2=0}^{x_1} \cdots \sum_{i_J=0}^{x_J} \prod_{j=1}^{J} \frac{1}{i_j!} \left(\frac{\lambda_j}{\lambda 0}\right)^{i_j} 1_{\left(\sum_{j=1}^{J} i_j = n - S_0\right)}, \quad (31)$$

where

$$1_{\left(\sum_{j=1}^{J} i_j = n - S_0\right)} = \begin{cases} 1, & \sum_{j=1}^{J} i_j = n - S_0 \\ 0, & \text{otherwise} \end{cases} \quad (32)$$

is the indicator function. Since $n - S_0 = \sum_{j=1}^{J} i_j$ and $\lambda_0 = \sum_{i=1}^{J} \lambda_i$, we can write

$$\lambda_0 \mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | n - S_0\right\} = (n - S_0)! \sum_{i_1=0}^{S_1} \sum_{i_2=0}^{S_1} \cdots \sum_{i_J=0}^{S_J} \prod_{j=1}^{J} \frac{1}{i_j!} \left(\frac{\lambda_j}{\lambda 0}\right)^{i_j} 1_{\left(\sum_{j=1}^{J} i_j = n - S_0\right)} \quad (33)$$

and $\lambda_0 \mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | n - S_0\right\}$ can be simplified to

$$\begin{aligned}
&\lambda_0 \mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | n - S_0\right\} = (n - S_0)! \sum_{i_1=0}^{S_1} \sum_{i_2=0}^{S_1} \cdots \sum_{i_J=0}^{S_J} \prod_{j=1}^{J} \frac{1}{i_j!} \left(\frac{\lambda_j}{\lambda 0}\right)^{i_j} 1_{\left(\sum_{j=1}^{J} i_j = n - S_0\right)} \quad (34) \\
&= \lambda_0 (n - S_0 - 1)! \sum_{i_1=0}^{S_1} \sum_{i_2=0}^{S_2} \cdots \sum_{i_J=0}^{S_J} \prod_{j=1}^{J} \frac{n - S_0}{i_j!} \left(\frac{\lambda_1}{\lambda_0}\right)^{i_j} 1_{\left(\sum_{j=1}^{J} i_j = n - S_0\right)} \\
&= \lambda_0 \, (n - S_0 - 1)! \sum_{i_1=0}^{S_1} \sum_{i_2=0}^{S_2} \cdots \sum_{i_J=0}^{S_J} \prod_{j=1}^{J} \frac{\sum_{k=1}^{J} i_k}{i_j!} \left(\frac{\lambda_j}{\lambda_0}\right)^{i_j} 1_{\left(\sum_{j=1}^{J} i_j = n - S_0\right)} \\
&= \lambda_0 (n - S_0 - 1)! \sum_{k=1}^{J} \left[\sum_{i_k=0}^{S_k} \frac{1}{(i_k - 1)!} \frac{\lambda_k}{\lambda_0} \left(\frac{\lambda_k}{\lambda_0}\right)^{i_k - 1} \left\{\sum_{i_1=0}^{S_1} \cdots \sum_{i_J=0}^{S_J} \prod_{j=1, j \neq k}^{J} \frac{1}{i_j!} \left(\frac{\lambda_j}{\lambda_0}\right) 1_{\left(\sum_{j=1}^{J} i_j = n - S_0\right)}\right\}\right] \\
&= (n - S_0 - 1)! \sum_{k=1}^{J} \left[\sum_{u_k=0}^{S_k - 1} \frac{\lambda_k}{(u_k)!} \left(\frac{\lambda_k}{\lambda_0}\right)^{u_k} \left\{\sum_{i_1=0}^{S_1} \cdots \sum_{i_J=0}^{S_J} \prod_{j=1, j \neq k}^{J} \frac{1}{i_j!} \left(\frac{\lambda_j}{\lambda_0}\right) 1_{\left(\sum_{j=1}^{J} i_j = n - S_0 - 1\right)}\right\}\right] \\
&= \sum_{j=1}^{J} \lambda_j \mathbb{P}\left\{B_0^1 \leq x_1, \ldots, B_0^j \leq S_j - 1, \ldots, B_0^J \leq x_J | n - S_0 - 1\right\}.
\end{aligned}$$

Therefore, for $n > S_0$, we have

$$Q(n) = \frac{(L_0)^n}{(n)!} \prod_{y=0}^{n-1} M(y) = \frac{(L_0)^n}{(n)!} \prod_{y=0}^{S_0 - 1} M(y) \prod_{y=S_0}^{n-1} M\,(y) = \frac{(\lambda_0)^{S_0} (L_0)^n}{(n)!} \prod_{y=S_0}^{n-1} M\,(y)$$

But $\prod_{y=S_0}^{n-1} M(y) = \lambda_0 \prod_{y=S_0+1}^{n-1} M(y)$, allowing us to simplify $M(n), n > S_0$ to

$$M(n) = \frac{\lambda_0 \mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | n - S_0 + 1\right\}}{\mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | n - S_0\right\}}. \tag{35}$$

Then

$$\begin{aligned}\lambda_0 \prod_{y=S_0+1}^{n-1} M(y) &= \lambda_0 \left[\frac{\lambda_0 \mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | 2\right\}}{\mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | 1\right\}} \times \frac{\lambda_0 \mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | 3\right\}}{\mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | 2\right\}}\right. \\ &\vdots \\ &\times \left.\frac{\lambda_0 \mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | n - S_0\right\}}{\mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | n - S_0 - 1\right\}}\right] = \frac{\lambda_0^{n-S_0} \mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | n - S_0\right\}}{\mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | 1\right\}}.\end{aligned} \tag{36}$$

But $\mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | 1\right\} = 1$ and

$$\lambda_0^{n-S_0} \mathbb{P}\left\{B_0^1 \leq S_1, \ldots, B_0^j \leq S_j, \ldots, B_0^J \leq S_J | n - S_0\right\} = (n - S_0)! \sum_{i_1=0}^{S_1} \sum_{i_2=0}^{S_2} \cdots \sum_{i_J=0}^{S_J} \prod_{j=1}^{J} \frac{(\lambda_j)^{i_j}}{i_j!} 1_{\left(\sum_{j=1}^J i_j = n - S_0\right)}. \tag{37}$$

Therefore,

$$\lambda_0 \prod_{y=S_0+1}^{n-1} M(y) = (n - S_0)! \sum_{i_1=0}^{S_1} \sum_{i_2=0}^{S_2} \cdots \sum_{i_J=0}^{S_J} \prod_{j=1}^{J} \frac{(\lambda_j)^{i_j}}{i_j!} 1_{\left(\sum_{j=1}^J i_j = n - S_0\right)}, \tag{38}$$

and

$$Q(n) = \frac{(\lambda_0)^{S_0} (L_0)^n}{(n)!} \prod_{y=S_0}^{n-1} M(y) = \frac{(n - S_0)! \, (\lambda_0 l_0)^{S0}}{(n)!} n! \sum_{i_1=0}^{S_1} \sum_{i_2=0}^{S_2} \cdots \sum_{i_J=0}^{S_J} \prod_{j=1}^{J} \frac{(\lambda_j l_0)^{i_j}}{i_j!} 1_{\left(\sum_{j=1}^J i_j = n - S_0\right)}, n > S_0. \tag{39}$$

■

7.4 Appendix 2: Comparison under symmetric instances for $w_j \geq 0.75$

References

Alvarez, E. and van der Heijden, M. (2014). On two-echelon inventory systems with poisson demand and lost sales. *European Journal of Operational Research*, 235(1):334 – 338.

Andersson, J. and Melchiors, P. (2001). A two-echelon inventory model with lost sales. *International Journal of Production Economics*, 69(3):307–315.

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

Babai, M. Z., Dai, Y., Li, Q., Syntetos, A., and Wang, X. (2022). Forecasting of lead-time demand variance: implications for safety stock calculations. *European Journal of Operational Research*, 296(3):846–861.

Table 8: Comparison under symmetric instances for $w_j \geq 0.75$

Input parameters						Simulation Model		Our Model		*JO* Model	
l_0	l_1	S_0	S_1	λ_1	w	$\alpha_1(\%)$	$\delta_1(\%)$	$\alpha_1(\%)$	$\delta_1(\%)$	$\alpha_1(\%)$	$\delta_1(\%)$
2	5	1	2	0.1	1.25	90.80	5.51	90.69	5.63	90.70	5.62
20	5	1	2	0.1	1.25	60.96	34.21	60.88	34.32	62.07	33.18
2	5	5	2	0.1	1.25	91.59	4.84	91.59	4.86	91.59	4.86
20	5	5	2	0.1	1.25	87.83	8.25	87.56	8.56	87.46	8.66
2	5	1	2	0.1	2.5	90.25	3.06	90.18	3.11	90.19	3.11
20	5	1	2	0.1	2.5	58.21	32.34	58.22	32.38	59.47	31.24
2	5	5	2	0.1	2.5	91.17	2.44	91.18	2.44	91.18	2.44
20	5	5	2	0.1	2.5	86.67	6.07	86.51	6.26	86.42	6.36
2	1	1	1	0.1	0.75	87.94	5.16	87.86	5.29	87.90	5.26
2	5	1	2	0.1	3.75	89.92	1.16	89.93	1.19	89.94	1.19
20	1	1	1	0.1	0.75	45.13	51.36	45.13	51.36	**48.21**	**48.04**
20	5	1	2	0.1	3.75	55.74	30.32	55.81	30.37	57.11	29.26
2	1	5	1	0.1	0.75	90.50	2.42	90.51	2.44	90.51	2.44
2	5	5	2	0.1	3.75	90.99	0.70	91.01	0.69	91.01	0.69
20	1	5	1	0.1	0.75	83.14	10.36	82.98	10.56	82.68	10.88
20	5	5	2	0.1	3.75	85.82	4.33	85.73	4.38	85.66	4.47
2	5	1	2	0.2	1.25	74.07	16.18	73.82	16.53	73.85	16.50
20	5	1	2	0.2	1.25	34.41	58.26	34.39	58.28	35.93	56.66
2	5	5	2	0.2	1.25	76.66	13.85	76.68	13.85	76.68	13.85
20	5	5	2	0.2	1.25	59.86	30.75	59.60	31.09	60.30	30.47
2	5	1	2	0.2	2.5	71.40	10.24	71.23	10.54	71.26	10.51
20	5	1	2	0.2	2.5	29.61	56.56	29.61	56.54	31.16	54.81
2	5	5	2	0.2	2.5	74.65	7.66	74.65	7.69	74.65	7.69
20	5	5	2	0.2	2.5	54.43	27.78	54.30	27.97	55.33	27.15
2	1	1	1	0.2	0.75	75.02	12.84	74.83	13.06	75.04	12.81
2	5	1	2	0.2	3.75	69.82	4.90	69.74	5.01	69.78	4.98
20	1	1	1	0.2	0.75	25.44	70.40	25.45	70.43	**28.74**	**66.60**
20	5	1	2	0.2	3.75	25.60	54.65	25.59	54.66	27.15	52.86
2	1	5	1	0.2	0.75	81.97	4.75	81.97	4.77	81.97	4.77
2	5	5	2	0.2	3.75	73.74	2.44	73.73	2.44	73.73	2.44
20	1	5	1	0.2	0.75	55.82	35.15	55.70	35.28	57.12	33.63
20	5	5	2	0.2	3.75	49.86	24.90	49.81	24.93	51.23	23.97

Bijvank, M., Huh, W., Janakiraman, G., and Kang, W. (2014). Robustness of order-up-to policies in lost-sales inventory systems. *Operations Research*, 62:5.

Bijvank, M. and Vis, I. (2011). Lost-sales inventory theory: A review. *European Journal of Operational Research*, 215:1–13.

Boucherie, R. J. and van Dijk, N. M. (2011). *Queueing Networks: A Fundamental Approach.* International Series in Operations Research and Management Science. Springer.

Boucherie, R. J., van Houtum, G.-J., Timmer, J., and van Ommeren, J.-K. (2018). A two-echelon spare parts network with lateral and emergency shipments: A product-form approximation. *Probability in the Engineering and Informational Sciences*, 32(4):536–555.

Boute, R. N., Lambrecht, M. R., and Van Houdt, B. (2007). Performance evaluation of a production/inventory system with periodic review and endogenous lead times. *Naval Research Logistics (NRL)*, 54(4):462–473.

Costantino, F., Gravio, G. D., Patriarca, R., and Petrella, L. (2018). Spare parts management for irregular demand items. *Omega*, 81:57 – 66.

De Clercq, S., Laevens, K., Steyaert, B., and Bruneel, H. (2013). A multi-class discrete-time queueing system under the fcfs service discipline. *Annals of Operations Research*, 202(1):59–73.

Drent, M. and Arts, J. (2021). Expediting in two-echelon spare parts inventory systems. *Manufacturing & Service Operations Management*, 23(6):1431–1448.

Dreyfuss, M. and Giat, Y. (2017). Optimal spares allocation to an exchangeable-item repair system with tolerable wait. *European Journal of Operational Research*, 261(2):584 – 594.

Dreyfuss, M. and Giat, Y. (2018). Optimal allocation of spares to maximize the window fill rate in a two-echelon exchangeable-item repair system. *European Journal of Operational Research*, 270(3):1053 – 1062.

Dreyfuss, M. and Giat, Y. (2019). Allocating spares to maximize the window fill rate in a periodic review inventory system. *International Journal of Production Economics*, 214:151 – 162.

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

Ingolfsson, A. (2005). Modeling the M(t)/M/s(t) queue with an exhaustive discipline. Technical report. https://sites.ualberta.ca/~aingolfs/documents/mmsnote.pdf.

Jakšič, M., Fransoo, J., Tan, T., De Kok, A., and Rusjan, B. (2011). Inventory management with advance capacity information. *Naval Research Logistics (NRL)*, 58(4):355–369.

Johansen, S. G. (2005). Base-stock policies for the lost sales inventory system with poisson demand and erlangian lead times. *International Journal of Production Economics*, 93-94:429–437. Proceedings of the Twelfth International Symposium on Inventories.

Johansson, L. and Olsson, F. (2018). Age-based inventory control in a multi-echelon system with emergency replenishments. *European Journal of Operational Research*, 265(3):951 – 961.

Kaufman, J. (1981). Blocking in a shared resource environment. *IEEE Transactions on Communications*, 29(10):1474–1481.

Kouki, C., Babai, M. Z., Jemai, Z., and Minner, S. (2019). Solution procedures for lost sales base-stock inventory systems with compound poisson demand. *International Journal of Production Economics*, 209:172 – 182. The Proceedings of the 19th International Symposium on Inventories.

Kouki, C., Legros, B., Zied Babai, M., and Jouini, O. (2020). Analysis of base-stock perishable inventory systems with general lifetime and lead-time. *European Journal of Operational Research*, 287(3):901–915.

Lefèvre, C. (1982). Some probability distributions related to the arrival of demands in a two-echelon inventory system. *International Journal Of Systems Science*, 13(3):317–321.

Lengu, D., Syntetos, A. A., and Babai, M. Z. (2014). Spare parts management: Linking distributional assumptions to demand classification. *European Journal of Operational Research*, 235(3):624–635.

Muckstadt, J. (2005). *Analysis and Algorithms for Service Part Supply Chains.* Springer: Berlin.

Nahmias, S. and Smith, S. A. (1994). Optimizing inventory levels in a two-echelon retailer system with partial lost sales. *Management Science*, 40(5):582–596.

Noorwali, A., Babai, M. Z., and Ducq, Y. (2023). Impacts of additive manufacturing on supply chains:

an empirical investigation. *Supply Chain Forum: An International Journal*, 24(2):182–193.

Özkan, E., van Houtum, G.-J., and Serin, Y. (2015). A new approximate evaluation method for two-echelon inventory systems with emergency shipments. *Annals of operations research*, 224(1):147–169.

Sherbrooke, C. (1968). METRIC: A multi-echelon technique for recoverable item control. *Operations Research*, 16(1):122–141.

Sherbrooke, C. (2004). *Optimal inventory modeling of systems: Multi-echelon techniques.* Wiley, 2 edition.

Simon, R. M. (1971). Stationary properties of a two-echelon inventory model for low demand items. *Operations Research*, 19(3):761–773.

Song, J. and Zipkin, P. (2009). Inventories with multiple supply sources and networks of queues with overflow bypasses. *Management Science*, 55(3):362–372.

Topan, E., Bayındır, Z. P., and Tan, T. (2017). Heuristics for multi-item two-echelon spare parts inventory control subject to aggregate and individual service measures. *European Journal of Operational Research*, 256(1):126 – 138.

Turrini, L. and Meissner, J. (2019). Spare parts inventory management: New evidence from distribution fitting. *European Journal of Operational Research*, 273(1):118–130.

Van Houtum, G.-J. and Kranenburg, B. (2015). *Spare parts inventory control under system availability constraints.* Springer.

Wolff, R. W. (1982). Poisson arrivals see time averages. *Operations Research*, 30(2):223–231.

Highlights

- We consider a two-echelon inventory system facing a Poisson demand
- We analyze the inventory system in the case of network lost sales
- We model the system using a queuing network representation
- We generalize the findings of earlier research for general lead times
- We propose more accurate solutions than those proposed in earlier research