Performance evaluation of a supply chain design by the integration of production capacity into the guaranteed service approach

Stef Lemmens [1], Catherine J. Decouttere [1], Nico J. Vandaele [1], Mauro Bernuzzi [2]

[1] Research Center for Operations Management of KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium
[2] GlaxoSmithKline Vaccines, Avenue Fleming 20, 1300 Waver, Belgium
{stef.lemmens@kuleuven.be, catherine.decouttere@kuleuven.be, nico.vandaele@kuleuven.be, mauro.bernuzzi@gsk.com}

Abstract. Previous research has integrated multi-echelon inventory management and supply chain responsiveness into the design of a supply chain by the use of the guaranteed service approach. We build further upon this work by integrating the production capacity and product flow to capture the supply chain's responsiveness. The production capacity is modeled with a queuing network to handle both the variability of the production processes and the demand. We discuss that such a responsiveness measure is highly relevant for the vaccine industry. Therefore, this work is a first step to study the trade-offs between supply chain responsiveness, inventories and production capacity for guaranteed service supply chains and aims to be applied to the design of a vaccine supply chain.

Keywords: Supply chain design, Responsiveness, Production capacity, Guaranteed service approach, Vaccine industry

1 Introduction

Emerging pressure from global competitors and stringent agreements with the final customer on the delivery date require a responsive end-to-end supply chain. We define such a responsive supply chain as a supply chain that is able to meet a fluctuating demand with variable, but relatively short lead times. Ideally, this means that a responsive supply chain is still able to satisfy the customer's demand in case of undesirable variability of the process durations (e.g. unexpected maintenance, material issues or quality problems) as well as variability stemming from externalities (e.g. natural disasters leading to sudden demand changes). Such undesirable variability of the process durations may induce a lead time distribution which can take a wide range of values and might expose a fat tail. However, supply chain responsiveness is not only determined by the duration of the different processes and their variability, but also by the position and the volume of the strategic stocks and the load of the installed production capacity in the supply chain.
A literature review on supply chain network design models by Lemmens et al. [14] shows that the majority of the literature imposes an economical performance criterion instead of a responsiveness criterion in the research field on supply chain design whilst literature confirms the importance of a lead time driven supply chain metric. Recent work of de Treville et al. [3] shows for three industrial supply chains (GSK Vaccines, Nissan Europe and Nestlé Switzerland) that managers underestimate the benefits of cutting lead times. This research is in line with the work of Suri [22]: he describes a company wide approach to reduce lead times and emphasizes the practical relevance.
The challenge and importance of the design of a responsive vaccine supply chain is emphasized by Shah [19, 20]. This industry is characterized by complex manufacturing processes and stringent regulatory processes which tend to be slow. The vaccine supply chain involves primary processes (cultivation of the antigen), secondary processes (formulation, filling and packaging) and an intense process of continuous quality control and quality assurance which lead to an overall supply chain lead time of more than 300 days. In this work, our contribution is to integrate quantitative relationships between supply chain responsiveness and production capacity into the guaranteed service approach. A further step of this research is to demonstrate the practical relevance of our approach to evaluate a vaccine supply chain design according to multiple performance measures.

2 Literature Review

A recent comprehensive survey of Eruguz et al. [4] classifies the existing Guaranteed Service Approach (GSA) literature along three dimensions: the considered assumptions, the developed solution methods and the industrial application. Graves and Willems [8] distinguish the GSA and the Stochastic Service Approach (SSA) as two main approaches in the literature on inventory models for multi-echelon supply chains. The GSA optimizes the strategic safety stock placement in supply chains while providing a high customer service level. This work is relevant for managers who face the pressure of reducing inventories in an existing supply chain. The GSA assumes that every supply chain stage quotes an outbound guaranteed service time: this is the time by which the stage under consideration satisfies the (internal or external) demand of the next stage. This quoted guaranteed service time complies with a service level of 100%. The GSA assumes that safety stock can be used to cope with normal demand variability and other countermeasures beyond safety stock must be used when the demand exceeds a normal variability level. In the GSA, such available countermeasures, for example accelerated production and overtime, refer to the operating flexibility assumption of the supply chain stages required to ensure a 100% delivery service. The assumption to be able to cope with normal demand variability is often referred to as *the assumption of bounded demand*. However, the external demand may not be bounded. This implies that one portion of the demand will be buffered with stock, another portion will be handled with operating flexibility measures and the remaining portion will be backlogged or lost. Traditionally, the external demand is assumed to be stationary and is propagated to all the other stages.
Graves and Willems [7, 9] and Moncayo-Martínez and Zhang [17] combine the GSA with supply chain configuration: some stages face the selection of an option of the functionality of the stage. These options may differ in its direct cost and lead time. Moncayo-Martínez and Zhang [17] minimize the total supply chain cost and the supply chain's responsiveness of a supply chain configuration problem simultaneously using a bi-objective MAX-MIN ant system. Funaki [6] and You and Grossmann [27, 28] combine the GSA with facility location decisions: production or distribution facilities have to be located and the appropriate stock levels have to be set to optimize the total supply chain costs and/or supply chain's responsiveness.
We further elaborate on the GSA as we observed that this approach is tractable from both a modeling and computational point of view. Humair et al. [10] confirm the savings for several real-world companies with inventory reductions by implementing the GSA. Billington et al. [1], Farasyn et al. [5] and Wieland et al. [25] elaborate on the GSA as a foundation of a multi-echelon inventory tool to manage inventories at Hewlett-Packard, Procter & Gamble and Intel respectively.
A large body of the publications in the guaranteed service approach literature focuses primarily on inventory optimization or minimizing total supply chain costs and assume infinite production capacity. For this work, supply chain responsiveness is an important key performance indicator and the assumption of infinite production capacity will be relaxed. Sitompul et al. [21] emphasize that locating strategic safety stock becomes a lot more complicated if production capacity constraints are taken into account. The authors derive the necessary excess safety stock based on the production capacity and standard deviation of the demand during the net replenishment lead time for serial supply chains. Production capacity constraints have also been integrated in guaranteed service supply chains by Jung et al. [11]. Their mathematical programming models the production quantities and the excess production capacity as decision variables and minimizes inventory holding costs and backordering costs. Our approach focuses on the integration of quantitative relationships between production capacity (and strategic inventory) and supply chain responsiveness using queuing networks. This is especially relevant for the vaccine industry as this industry faces long and variable lead times. For GSK Vaccines, de Treville et al. [3] explains that these long and variable lead times also resulted from moving production between factories to obtain costs savings from capacity utilization.

3 Methodology

3.1 Guaranteed Service Approach

As in Humair et al. [10], we further build on the distinction between the guaranteed service approach with deterministic lead times (GSA-DET) and the guaranteed service approach with variable lead times (GSA-VAR). The lead time of a stage represents the time from the availability of all the inputs of this

stage until the output is ready to serve the (internal or external) demand of the next stage and may include material handling, machine processing, transportation time and waiting time, but also time to undergo regulatory, quality and release procedures. In GSA-DET, the value for these lead times is deterministic and two easy heuristics exist to choose this value in case of an empirical distribution of the lead times: (1) fixing the lead times to its mean value and (2) fixing the lead times to its maximum value. However, such a reasonable heuristic leads to inaccurate inventory levels and an inaccurate responsiveness performance measure. To the best of our knowledge, the work of Humair et al. [10] and Neale and Willems [18] are the only manuscripts that allow lead time variability into the GSA. These authors demonstrate how inventory levels can be determined in a more accurate way instead of using these two heuristics.
The remainder of this section is organized as follows. First, we introduce the modeling framework of the GSA with deterministic lead times. We refer the interested reader to Graves and Willems [8] for a more detailed description of GSA-DET. Next, we present how Humair et al. [10] allow for lead-time variability. Finally, we show our approach which integrates production capacity and still allows for demand and lead time variability.

3.2 Guaranteed Service Approach with Deterministic Lead Times (GSA-DET)

According to Graves and Willems [8], a supply chain stage represents a processing resource in the supply chain and might be the procurement of raw materials, production of components or subassemblies, production of assemblies and testing of the finished goods or the transportation from a distribution center to a regional warehouse. The supply chain network can also be represented by a graph whereby the nodes correspond to the supply chain stages and the arcs denote the precedence relationships between the nodes. A strategic safety stock point can be located at each stage.
For GSA-DET, we assume that each stage j has a *deterministic production lead time*, $\mathrm{ew_j}$. This means, for example, if the inputs for stage j are available at time t, then the considered stage has completed its processing request at time $t + \mathrm{ew_j}$. Graves and Willems [8] mention the important assumption that the production lead time is not influenced by the capacity utilization. We will relax this assumption in our model.
The GSA assumes that each supply chain stage quotes a *Guaranteed Service Time* (GST). The GST is the time by which the stage can guarantee a 100 % delivery service. The stage must hold sufficient inventory such that it is able to ensure the strict service guarantee. For each stage j, an inbound guaranteed service time ($\mathrm{S_j^{IN}}$) and an outbound guaranteed service time ($\mathrm{S_j^{OUT}}$) is quoted. The inbound GST is the time when all inputs of the stage are available and processing can start. Remark that the inbound GST is equal to the maximum outbound GST of the preceding stages in case of a non-serial supply chain (e.g. assembly type supply chain topology). According to Graves and Willems [8], this can be expressed as

$$\mathrm{S_j^{IN}} = \max_{i:(i,j)\in A} \{\mathrm{S_i^{OUT}}\}. \tag{1}$$

where A denotes the set of arcs in the supply chain network. For GSA-DET the replenishment time is deterministic as both the service time and production lead time are deterministic. [27, 28] emphasize the replenishment lead time as an important difference between single-echelon and multi-echelon inventory systems. In a single stage inventory system, replenishment lead time is often exogenous and treated as parameter whilst in a multi-echelon inventory system it depends on the inventory level of the predecessor(s). We denote the replenishment time as the *Net (replenishment) Lead Time* (NLT). The GSA computes the NLT as the sum of the stage's inbound service time and the production lead time minus the outbound service time, or formally:

$$\mathrm{NLT_j} = \mathrm{S_j^{IN}} + \mathrm{ew_j} - \mathrm{S_j^{OUT}}. \tag{2}$$

You and Grossmann [28] show that the supply chain's responsiveness of the entire supply chain can be measured as the outbound GST of the last stage (e.g. customer markets) of the supply chain. This responsiveness measure quantifies the maximum time within which the (external) demand is satisfied. In case of multiple customer markets, the supply chain's responsiveness may be computed as (1) the maximum outbound GST of the markets or (2) a weighted average value of the outbound GSTs of the customer markets. In such a case we will opt for the first method.

Graves and Willems [7] derive the strategic Safety Stock ($\mathrm{SS_j}$) level held at the stage under consideration as a function of its net replenishment lead time:

$$\mathrm{SS_j} = \mathrm{k_j} \sigma_\mathrm{j}^\mathrm{D} \sqrt{\mathrm{NLT_j}}. \tag{3}$$

In these equations, $\mathrm{k_j}$ represents the *safety stock factor* at stage j. This parameter determines the stage's demand that can be covered with the stage's safety stock. A common assumption is that the demand of stage j is normally distributed with mean $\mu_\mathrm{j}^\mathrm{D}$ and standard deviation $\sigma_\mathrm{j}^\mathrm{D}$. The expected *Work-In-Process* inventories (WIP) or *pipeline stock* at stage j is determined using Little's law [15]:

$$\mathrm{WIP_j} = \mathrm{ew_j} \mu_\mathrm{j}^\mathrm{D}. \tag{4}$$

where $\mathrm{ew_j}$ is the stage's average lead time. We formulate the complete deterministic mathematical program of the guaranteed service framework:

$$\min \sum_{j \in \mathrm{J}} \mathrm{c_j^{SS}} \mathrm{k_j} \sigma_\mathrm{j}^\mathrm{D} \sqrt{\mathrm{S_j^{IN}} + \mathrm{ew_j} - \mathrm{S_j^{OUT}}} + \sum_{j \in \mathrm{J}} \mathrm{c_j^{WIP}} \mathrm{ew_j} \mu_\mathrm{j}^\mathrm{D} \tag{5}$$

$$\min \mathrm{R} = \mathrm{S_j^{OUT}} \quad \forall j = |\mathrm{J}| \tag{6}$$

subject to

$$\mathrm{S_j^{OUT}} - \mathrm{S_j^{IN}} \leq \mathrm{ew_j} \quad \forall j \in \mathrm{J} \tag{7}$$

$$\mathrm{S_j^{IN}} - \mathrm{S_i^{OUT}} \geq 0 \quad \forall (i,j) \in \mathrm{A} \tag{8}$$

$$\mathrm{S_j^{IN}}, \mathrm{S_j^{OUT}} \geq 0 \quad \forall j \in \mathrm{J} \tag{9}$$

The first objective function (5) minimizes the strategic safety stock and WIP cost across the supply chain. The second objective function (6) optimizes the $\mathrm{S_j^{OUT}}$ of the last stage, which is the supply chain's responsiveness and denoted as R. As mentioned earlier, the majority of the current literature that applies the guaranteed service model, proposes an objective function based on (5). Remark that the strategic safety stock levels are nonlinearly related to the NLTs.
Constraint set (7) implies that the NLTs are positive. Constraints (8) are a linearization of (1): the stage's inbound GST is the time when the outputs of all the preceding stages are available. The nonnegativity of the guaranteed service times is assured by constraint set (9).

3.3 Guaranteed Service Approach with Variable Lead Times (GSA-VAR)

Humair et al. [10] emphasize the importance of integrating variable lead times into the GSA as every supply chain is confronted with variable lead times at certain stages. The authors develop a closed-form expression to compute the safety stock for the GSA with variable lead times such that the total inventory cost can be minimized. They provide a numerical comparison of safety stock and WIP levels of GSA-VAR and two heuristics for fixing the lead times for GSA-DET: (1) fixing the lead times to its mean value and (2) fixing the lead times to its maximum value. Compared to GSA-VAR, the safety stock levels may be significantly underestimated for both heuristics and the WIP level may be overestimated for fixing the lead times to its maximum values by GSA-DET.
Remark that in case of variable lead times, processing at a stage might also be finished early because of shorter lead time realizations. In this case, it is impossible to pass stock to the next stage as the GSA assumes that this downstream stage can only start processing at its inbound GST. This leads to *Early-Arrival Stock* (EAS) in the supply chain. Such a phenomenon only occurs if the lead time realization is smaller than the difference between the outbound and inbound service time of the stage. In this case, the NLT is negative and this leads to a *negative shortfall*. For completeness, we note that there is a *positive shortfall* when the NLT is positive and no shortfall when the NLT equals zero.
Humair et al. [10] determine a closed-form expression for the computation of the safety stock (for GSA-DET, see (3)) and the average early-arrival stock in case of variable lead times. In this work, we show the

resulting closed-form expressions. We refer the interested reader to the appendix of Humair et al. [10] for the full derivation of these equations:

$$\mathrm{SS_j} = \mathrm{k_j}\sqrt{\mathrm{Q}([\mathrm{S_j^{OUT}} - \mathrm{S_j^{IN}}]^+)(\sigma_j^D)^2 + (\mu_j^D)^2\mathrm{R}([\mathrm{S_j^{OUT}} - \mathrm{S_j^{IN}}]^+)}. \tag{10}$$

$$\mathrm{EAS_j} = \mu_j^D(\mathrm{Q}([\mathrm{S_j^{OUT}} - \mathrm{S_j^{IN}}]^+) - \mathrm{ew_j} + [\mathrm{S_j^{OUT}} - \mathrm{S_j^{IN}}]^+). \tag{11}$$

The expected value and the variance of the positive part of the NLT are computed by $\mathrm{Q}([\mathrm{S_j^{OUT}} - \mathrm{S_j^{IN}}]^+)$ and $\mathrm{R}([\mathrm{S_j^{OUT}} - \mathrm{S_j^{IN}}]^+)$ respectively. In equation (10), the square root term denotes the standard deviation of the positive shortfall. In equation (11), $\mathrm{ew_j}$ represents the mean value of the lead time distribution of stage j. Remark that the closed-form expressions hold for both discrete and continuous distributions of the lead times.

3.4 Incorporation of Production Capacity Constraints into the Guaranteed Service Approach with Variable Lead Times

Figure 1 shows the supply chain structure of Chain 01 of Willems [26] consisting of three echelons and eight stages: three part suppliers (Part1, Part2 and Part3), two manufacturers (Manu1 and Manu2) and three retailers (Ret1, Ret2 and Ret3). The triangles represent candidate strategic safety stock locations. For the ease of explanation, we refer to the stages as nodes, denote the set of nodes as J and partition the set of nodes according to the three echelons: $\mathrm{J_P}$, $\mathrm{J_M}$ and $\mathrm{J_R}$ where $\mathrm{J_P}$ is the set of part supplier nodes, $\mathrm{J_M}$ is the set of manufacturer nodes and $\mathrm{J_R}$ is the set of retailer nodes. At each node, an activity with a certain duration has to be performed. The arcs represent the precedence relationships between the nodes and we assume an assembly type supply chain: the activities at Manu1, Manu2 and Ret2 can only start when all the predecessors are ready. Remark that such an activity may include machine processing time, but also waiting, material handling, transportation time or time to undergo regulatory and quality procedures. We refer to the node's total (average) activity duration as the node's average lead time.

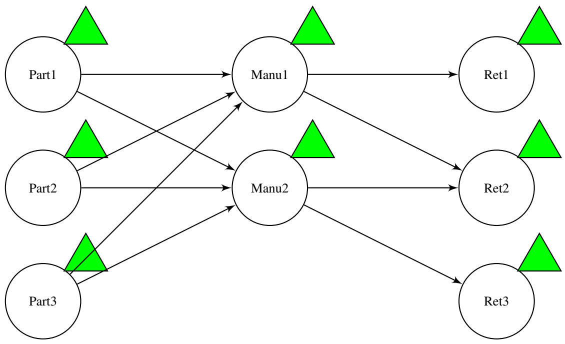

Figure 1: Structure of supply chain Chain 01 of Willems [26]

Furthermore we assume that the average lead time and its variability are available for the part supplier and the retailer echelon. The manufacturing echelon consists of intense processing nodes. Remark that such a node may represent a complex production system and could be disaggregated into a subnetwork of nodes. For each manufacturing node we model only one machine and assume that its processing time and variability are known and the lead time and its variability are to be determined. One of our contributions is to study these lead times as a function of the installed production capacity for guaranteed service supply chains. In case of heavy traffic, the waiting time for such a processing activity may increase drastically. Such a long waiting time has a negative impact on the corresponding node's outbound service time and subsequently damages the supply chain's responsiveness if no additional production capacity and/or inventory buffers can be inserted. We distinguish two phases for integrating these quantitative relationships into guaranteed service supply chains. The first phase calculates the lead time and lead time variability for the manufacturing nodes. The second phase calculates the supply chain's responsiveness and total stock with GSA-VAR using the lead times and lead time variabilities obtained in the first phase.

3.4.1 Phase 1: Lead Time and Lead Time Variability Calculation Phase

We assume that the average demand per time unit and its standard deviation, denoted as μ_j^D and σ_j^D respectively, is available for $j \in J_R$ and follows a stationary process. For an assembly type supply chain, the internal demand of the manufacturer echelon J_M and the part supplier echelon J_P can be computed sequentially by:

$$\mu_j^D = \sum_{i:(j,i)\in A} \theta_{j,i}\mu_i^D. \tag{12}$$

where $\theta_{j,i}$ represents the number of items of upstream node j required for downstream node i according to the bill of material.

In our work, we also propagate supply variability by using Squared Coefficients of Variation (SCVs). We denote the SCVs of the arrival processes as $(C_j^A)^2$ and assume that these are available for nodes $j \in J_P$. For this echelon, we assume that the SCVs of the departing processes, $(C_j^D)^2$, are equal to $(C_j^A)^2$ as this echelon does not require intense machine processing operations.

For the manufacturing echelon, we estimate each node's lead time and corresponding variability. According to (10) and (11), this information influences the necessary safety stock levels and emerging early-arrival stock. We compute the expected lead time and its variability by taking production capacity constraints into account using a G/G/1 queueing model. Furthermore, we allow the lead time and its variability to be dependent of the batch size and assume general interarrival and processing time distributions. For nodes $j \in J_M$, we compute the expected lead time ew_j based on Lambrecht and Vandaele [13]:

$$ew_j = ew_j^C + ew_j^Q + ew_j^P. \tag{13}$$

where ew_j^C is the batch collection time, ew_j^Q is the batch waiting time and ew_j^P is the batch processing time. As such batch collection time is less relevant for our industrial application, we omit the batch collection time. For ew_j^P, we assume that the processing times are independent and identically distributed and a setup time τ_j which is independent from the processing times. The expected batch processing time can be computed similarly to Lambrecht and Vandaele [13]:

$$ew_j^P = \tau_j + \frac{1}{2}(Q_j + 1)t_j. \tag{14}$$

where Q_j and t_j represent the batch size and the unit processing time respectively. The expected waiting time is derived by the use of the bulk arrival-bulk service approach for a G/G/1 queue:

$$ew_j^Q = \frac{\rho_j^2((C_j^{BA})^2 + (C_j^{BS})^2}{2\lambda_j^{BA}(1-\rho_j)} g(\rho_j, (C_j^{BA})^2, (C_j^{BS})^2). \tag{15}$$

where $(C_j^{BA})^2$ and $(C_j^{BS})^2$ refer to the SCV of the Batch Arrival (BA) and Batch Service (BS) processes respectively. The utilization rate ρ_j and $g(\rho_j, (C_j^{BA})^2, (C_j^{BS})^2)$ are used as defined in Whitt [24] and Lambrecht and Vandaele [13]. For nodes $j \in J_M$, the lead time variability can be obtained by:

$$(\sigma_j^{LT})^2 = (\sigma_j^{ew^Q})^2 + (\sigma_j^t)^2 + 2cov(ew_j^Q, ew_j^P). \tag{16}$$

which includes the variance of the expected waiting time, the variance of the batch processing time and their covariance as they are not independent. Based on simulation results, Lambrecht and Vandaele [13] postulate that such covariance is hard to obtain in case of a lot sizing model and drop this computation as they expect it to have a low contribution to $(\sigma_j^{LT})^2$. In this work, we also consider the computation of this covariance term as out of scope. We use the approximation of Whitt [24] to approximate the variance of the batch waiting time.

All the previous information allows us to compute the first and second order moment of the lead time distribution. However, for the intermediary computations for (10) and (11) the lead time probability distribution should be known. Even when the probability distribution of the waiting time and the machine processing time are known, the probability distribution of the lead time might be hard to obtain. Experiments of Lambrecht et al. [12], Lambrecht and Vandaele [13] and Vandaele [23] indicate that the (right-) skewed lognormal distribution provides a good fit for the lead time distribution function. For that reason we will also use a lognormal distribution and determine its scale parameter and shape parameter in the same way as the authors above.

For nodes $j \in J_M$ and a given pair of inbound and outbound service times, the safety stock and early-arrival stock can now be computed in the second phase. Note that both safety stock and early-arrival stock are now related to the production capacity and its utilization.

3.4.2 Phase 2: Responsiveness and Stock Calculation Phase with GSA-VAR

We are interested in exploring the trade-offs between supply chain responsiveness, supply chain stock and production capacity for guaranteed service supply chains. However, the supply chain's total inventories is a nonlinear function of the decision variables (see equations (10) and (11)). Therefore we integrate a piecewise linear approximation formulation into GSA-VAR. This formulation is also referred to as the Multiple Choice Model and is based on the work of Magnanti et al. [16] and Croxton et al. [2]. We prefer the use of such a formulation as is it does not require a specific supply chain topology.

4 Results for Phase 1

We now illustrate how the first phase can be used to evaluate a supply chain design according to multiple performance criteria. For this section, we further elaborate on Chain 01 of Willems [26] and assume that the demand level of Ret3 varies under different scenarios. For every scenario, we want to evaluate the total safety stock, early-arrival stock, the capacity utilization and the supply chain's responsiveness. However, we show the performance of manufacturing node Manu2 under different scenarios. Studying this node is particularly interesting as its performance is subject to production capacity constraints as well as the demand of Ret3. Table 1 gives a summary of the generated supply chain input data. The table entries filled with a cross indicate data which are computed by our approach. Table 2 shows the queuing results of the first phase for node Manu2 under the different scenarios. Remark that the utilization rate, the average lead time and the lead time variability increase as the node's load increases. For each scenario, the second phase is able to evaluate the safety stock and the accumulated early-arrival stock for a pair of inbound and outbound guaranteed service times.

Table 1: Summary of the input data for Chain 01

Node name	j	Q_j (units)	t_j (minutes)	τ_j (minutes)	μ_j^{LT} (days)	σ_j^{LT} (daily)	μ_j^D (daily)	σ_j^D (daily)
Part1	0	NA	NA	NA	28	11.22	x	x
Part2	1	NA	NA	NA	15	0	x	x
Part3	2	NA	NA	NA	10	0	x	x
Manu1	3	146	3.15	134.99	x	x	x	x
Manu2	4	27	6.14	35.56	x	x	x	x
Ret1	5	NA	NA	NA	0	0	253	36.62
Ret2	6	NA	NA	NA	0	0	45	1
Ret3	7	NA	NA	NA	0	0	75	2

5 Conclusions and Future Steps

This work extends the guaranteed service approach with variable lead times by the integration of production capacity. Therefore we decompose a manufacturing node into a single machine batch processing activity. This extension allows a supply chain design to be evaluated according to multiple performance measures: supply chain inventories, supply chain responsiveness and capacity utilization. Our preliminary results show that the integration of production capacity is relevant. We consider load-dependent lead times and variabilities in our approach. Such load-dependent lead times and variabilities may influence the strategic stock levels in case of different demand scenario realizations.
We are currently applying the model to the design of a vaccine supply chain. One further step of this research is to decompose such a manufacturing node into a (sub)network of nodes. In this way, the production capacity will be integrated in a more elaborate way. Such an elaborated approach will also increase the relevance of this methodology for our industrial application.

Table 2: Results for node Manu2 (j = 4) under different scenarios

Scenario name	Scenario number	μ_j^D (daily)	ρ_j	ew_j^Q (days)	$\sigma_j^{ew^Q}$ (daily)	μ_j^{LT} (days)	σ_j^{LT} (daily)
Ret3_Dem = 75	1	120	0.6211	0.1221	0.4677	0.2619	0.4695
Ret3_Dem = 80	2	125	0.6424	0.1328	0.4780	0.2715	0.4797
Ret3_Dem = 85	3	130	0.6637	0.1448	0.4896	0.2827	0.4912
Ret3_Dem = 90	4	135	0.6850	0.1585	0.5027	0.2955	0.5042
Ret3_Dem = 95	5	140	0.7063	0.1742	0.5174	0.3105	0.5189
Ret3_Dem = 100	6	145	0.7276	0.1924	0.5341	0.3279	0.5355
Ret3_Dem = 105	7	150	0.7490	0.2137	0.5532	0.3486	0.5545
Ret3_Dem = 110	8	155	0.7703	0.2390	0.5751	0.3732	0.5764
Ret3_Dem = 115	9	160	0.7916	0.2696	0.6006	0.4031	0.6018
Ret3_Dem = 120	10	165	0.8129	0.3070	0.6307	0.4400	0.6318

References

[1] Billington, C., Callioni, G., Crane, B., Ruark, J.D., Rapp, J.U., White, T., Willems, S.P.: Accelerating the Profitability of Hewlett Packard's Supply Chains. Interfaces 34(1), 59–72 (2004)

[2] Croxton, K.L, Gendron, B., Magnanti, T.L.: A Comparison of Mixed-Integer Programming Models for Nonconvex Piecewise Linear Cost Minimization Problems. Management Science 49(9), 1268–1273 (2003)

[3] de Treville, S., Bicer, I., Chavez-Demoulin, V., Hagspiel, V., Shürhoff, N., Tasserit, C., Wager, S.: Valuing lead time. Journal of Operations Management 32(6), 337–346 (2014)

[4] Eruguz, A.S., Sahin, E., Jemaai, Z., Dallery, Y.: A comprehensive survey of guaranteed-service models for multi-echelon inventory optimization.: International Journal of Production Economics 172, 110–125 (2016)

[5] Farasyn, I., Humair, S., Kahn, J.I., Neale, J.J., Rosen, O., Ruark, J., Tarlton, W., Van de Velde, W., Wegryn, G. Willems, S.P.: Inventory Optimization at Procter & Gamble: Achieving Real Benefits Through User Adoption of Inventory Tools. Interfaces 41(1), 66–78 (2011)

[6] Funaki, K.: Strategic safety stock placement in supply chain design with due-date based demand. International Journal of Production Economics 135(1), 4–13 (2012)

[7] Graves, S.C., Willems, S.P.: Optimizing Strategic Safety Stock Placement in Supply Chains. Manufacturing & Service Operations Management 2(1), 66–78 (2000)

[8] Graves, S.C., Willems, S.P.: Supply Chain Design: Safety Stock Placement and Supply Chain Configuration. In: Graves, S.C., de Kok, A.G., Supply Chain Management: Design, Coordination and Operation, volume 11 of Handbooks in Operations Research and Management Science, pp. 95–132, Elsevier (2003)

[9] Graves, S.C., Willems, S.P.: Optimizing the Supply Chain Configuration for New Products. Management Science 51(8), 1165–1180 (2005)

[10] Humair, S., Ruark, J., Tomlin, B., Willems, S.P.: Incorporating Stochastic Lead Times Into the Guaranteed Service Model of Safety Stock Optimization. Interfaces 43(5), 421–434 (2013)

[11] Jung, J.Y., Blau, G., Pekny, J.F., Reklaitis, G.V., Everysdyk, D.: Integrated safety stock management for multi-stage supply chains under production capacity constraints. Computers & Chemical Engineering 32(11), 2570–2581 (2008)

[12] Lambrecht, M.R., Ivens, P.L., Vandaele, N.J.: ACLIPS: A Capacity and Lead Time Integrated Procedure for Scheduling. Management Science 44(11), 1548–1561 (1998)

[13] Lambrecht, M.R., Vandaele, N.J.: A general approximation for the single product lot sizing model with queueing delays. European Journal of Operational Research 95(1) 73–88 (1996)

[14] Lemmens, S., Decouttere, C.J., Vandaele, N.J., Bernuzzi, M.: A review of integrated supply chain network design models: key issues for vaccine supply chains. Chemical Engineering Research and Design 109C, 366–384 (2016)

[15] Little, J.D.C.: A Proof for the Queuing formula: $L = \lambda W$. Operations Research 9(3), 383–387 (1961)

[16] Magnanti, T.L., Shen, Z.-J.M., Shu, J., Simchi-Levi, D., Teo, C.-P.: Inventory placement in acyclic supply chain networks. Operations Research Letters 34(2), 228–238 (2006)

[17] Moncayo-Martínez, L.A., Zhang, D.Z.: Optimising safety stock placement and lead time in an assembly supply chain using bi-objective MAX-MIN ant system. International Journal of Production Economics 145(1), 18–28 (2013)

[18] Neale, J.J., Willems, S.P.: Managing Inventory in Supply Chains with Nonstationary Demand. Interfaces 39(5), 288–399 (2009)

[19] Shah, N.: Pharmaceutical supply chains: key issues and strategies for optimisation. Computers & Chemical Engineering 28(6-7), 929–941 (2004)

[20] Shah, N.: Process industry supply chains: Advances and challenges. Computers & Chemical Engineering 29(6), 1225–1236 (2005)

[21] Sitompul, C., Aghezzaf, E.-H., Dullaert, W., Van Landeghem, H.: Safety stock placement problem in capacitated supply chains. International Journal of Production Research 46(17), 4709–4727 (2008)

[22] Suri, R.: Quick Response Manufacturing: A Companywide Approach to Reducing Lead Times. Chicago: American Technical Society (1998)

[23] Vandaele, N.J.: The Impact of Lot Sizing on Queueing Delays: Multi Product, Multi Machine Models. PhD thesis, KU Leuven (1996)

[24] Whitt, W.: The Queueing Network Analyzer. The Bell System Technical Journal 62(9), 2779–2815 (1983)

[25] Wieland, B., Mastrantonio, P., Willems, S.P., Kempf, K.G.: Optimizing Inventory Levels Within Intel's Channel Supply Demand Operations. Interfaces 42(6), 517–527 (2012)

[26] Willems, S.P.: Data Set – Real-World Multiechelon Supply Chains Used for Inventory Optimization. Manufacturing & Service Operations Management 10(1), 19–23 (2008)

[27] You, F., Grossmann, I.E.: Integrated multi-echelon supply chain design with inventories under uncertainty: MINLP models, computational strategies. American Institute of Chemical Engineers Journal 56(2), 419–440 (2010)

[28] You, F., Grossmann, I.E.: Balancing responsiveness and economics in process supply chain design with multi-echelon stochastic inventory. American Institute of Chemical Engineers Journal 57(1), 178–192 (2011)