Les cahiers Leibniz

A simulation-optimization approach for managing the sales and operations planning in the automotive industry

Lâm Laurent Lim, Gülgün Alpan, Bernard Penz

Laboratoire G-SCOP
46 av. Félix Viallet, 38000 GRENOBLE, France
ISSN : 1298-020X

n° 212

January 2014

Site internet : http://www.g-scop.inpg.fr/CahiersLeibniz/

A simulation-optimization approach for managing the sales and operations planning in the automotive industry

Lâm Laurent LIM, Gülgün ALPAN, Bernard PENZ

January 21, 2014

Abstract

Due to the increasing globalization and the distant sourcing, reconciling industrial constraints and sales requirements becomes very challenging for industries facing an uncertain environment and demanding customers. The sales and operations planning (S&OP) is crucial for efficiently balancing production capacities with the volatile market demand. In this article, we propose an original S&OP model in order to improve the trade-off between the supply chain costs and the customer satisfaction. The problem is formulated as a multi-objective optimization model with ϵ-constraints and is solved by a simulation-optimization approach. Two classes of policies for managing the parts procurement and the flexibility offered to the sales function are presented. The model and the proposed solution are illustrated with the case study of Renault, a French global automobile manufacturer. Several policies and optimization algorithms are compared in terms of system performance and computation time. Managerial insights are derived based on these results.

Keywords

Simulation-optimization ; sales and operations planning ; flexibility ; uncertain demand ; customer impatience ; distant sourcing ; automotive industry

1 Introduction

During the last decades, the automotive industry has largely evolved from mass production to mass customization with more individualized and sophisticated vehicles (MacCarthy et al., 2003; Brabazon et al., 2010). If the product variety represents an important competitive advantage (Ramdas, 2003), it also makes the supply chain more complex and can increase significantly the production costs (Stäblein et al., 2011). To adapt efficiently their production capacities with the volatile and changing market demand, the automobile manufacturers strive to make their production systems more flexible and to implement build-to-order strategies in their supply chains (Miemczyk and Holweg, 2004; Howard et al., 2005; Volling and Spengler, 2011).

More recently, the globalization has significantly increased the procurement lead times. This makes supply chains more vulnerable to various disruptions like uncertain demand (Tang, 2006). The use of a traditional build-to-order strategy may be inefficient to deal with uncertainty and distant sourcing. Indeed, while customers are impatient and would not wait for a long time (Elias, 2002; Holweg et al., 2005), the firm needs to procure parts several weeks beforehand because of distant suppliers. Therefore, part procurements are based on forecasts that are often unreliable, especially in the automotive industry (Elkins et al., 2004; Childerhouse et al., 2008).

A trade-off has to be made between two conflicting objectives: on one hand, the company needs more flexibility and shorter delivery times to satisfy customer demand. On the other hand, the production system asks for a stable production plan and more visibility on the future demand. To make the supply chain more agile and flexible in an uncertain environment, several actions are possible.

In this paper, we focus on the sales and operations planning (S&OP) as a means of flexibility. This process defines the tactical production plan that links strategic perspectives to daily operations (Grimson and Pyke, 2007). The S&OP has to take into account both the sales objectives and the industrial constraints. The S&OP is crucial to improve the trade-off between logistic costs and customer requirements. Recent literature reviews on S&OP are given in Grimson and Pyke (2007) and Thomé et al. (2012). Researchers show that the S&OP can improve significantly firms' performance and this topic becomes increasingly popular in industry. Thomé et al. (2013) analyse a sample of 725 manufacturers around the world and test different hypotheses on the impact of S&OP on the manufacturing performance. The authors argue that the integration of suppliers amplifies the positive impact of internal S&OP on the firm performance.

For the automotive industry, Hahn et al. (2000) describe the new mechanisms of Hyundai to coordinate sales and supply chain. The authors insist on the importance of a better synchronisation and integration of sales and supply chain functions. The study of Tomino et al. (2009) compares the production planning methods of Toyota, Nissan and Mitsubishi and shows how the automobile manufacturers have implemented a market flexible customizing system.

Our research studies a S&OP model, introduced in Lim et al. (2014), for managing the conflicting objectives of sales and supply chain. The originality of this S&OP model lies on sales constraints and flexibility rates for partially controlling the order fulfilment process. Safety stocks are managed according to the flexibility given to the sales function. This planning method is particularly relevant for companies that face distant sourcing and unreliable forecasts. It helps reducing logistic costs while improving the customer satisfaction. A first investigation using a simulation approach of this method is presented in Lim et al. (2014). The authors describe in details all characteristics of the S&OP, compare the proposed model with other companies and models in literature, and show how it can be applied in other situations. The authors also investigate the planning dynamics and highlight its advantages. Here, we extend the research of Lim et al. (2014) by introducing optimal policies for managing parts inventory and sales flexibility in the S&OP. These policies are obtained via a simulation-optimization approach.

Although simulation is a powerful tool to improve operations' efficiency and to incorporate uncertainties in real complex systems (Glover et al., 1999), it can only test what-if scenarios and it is inadequate for solving optimization problems. To remediate this situation, in this paper we use a simulation-optimization approach which consists of a structured method to determine optimal input parameter values, where the objective function is measured by a simulation model (Swisher et al., 2000). Simulation-optimization models have been widely used for solving complex industrial problems (see, among others, Rosen and Harmonosky, 2005; Zeng and Yang, 2009; Li et al., 2009; Keskin et al., 2010). There exists various techniques to search efficiently the best parameter values in simulation-optimization problems: random search (Andradóttir, 2006), metaheuristics (Haddock and Mittenthal, 1992; Rosen and Harmonosky, 2005; Ólafsson, 2006; Alrefaei and Diabat, 2009), gradient-based procedures (Fu, 2006), response surface methodology (Neddermeijer et al., 2000; Kleijnen, 2008) etc. In this research, we consider several random search techniques and a metaheuristic method (simulated annealing). For more information on the vast research area of simulation-optimization and its applications, we refer the readers to the literature reviews of, among others, Fu (1994), Swisher et al. (2000), Fu (2002), Swisher et al. (2004), Fu et al. (2005), Rani and Moreira (2010).

The research objective of this paper is, first of all, to provide an efficient method to compute the optimal S&OP policies. We present a multi-objective optimization model with ϵ-constraints and we investigate two different classes of policies (static and linear). The second objective is to compare different optimization techniques and generated solutions in terms of system performance and computation time. We apply our method on the case study of Renault, a global automanufacturer. Numerical experiments on industrial data show that there are significant benefits of using linear policies instead of static ones. We also provide managerial insights for decision makers and discuss the practical implementation of our solution. To the best of our knowledge, this article is the first that proposes a multi-objective simulation-optimization model for solving a S&OP problem with flexibility, uncertain demand and impatient customers.

The paper is organized as follows. Section 2 presents the problem, the notations and the mathematical model. Section 3 describes the simulation-optimization approach and the different optimization techniques employed in this article. Section 4 details the experimental design and the numerical results based on the Renault's case. Several managerial insights are derived from these results. Finally, the contribution of this paper and research perspectives are discussed in Section 5.

2 Problem description

The problem described in this section is based on an industrial issue detailed in Lim et al. (2014). We refer the reader to this article for a comprehensive problem description, justification of assumptions and for a comparison with other models of the literature.

2.1 S&OP with sales constraints and flexibility rates

We focus on a sales and operations planning with uncertain demand, impatient customers and distant sourcing. We consider a weekly periodic review planning model. We assume that the products are managed independently in the S&OP. Therefore, we only present a single-product model in the following. The objective of the S&OP is to find the best trade-off between the supply chain costs and the customer satisfaction. Measuring the overall system performance is not simple because the S&OP is a cross-functional process that involves business and operations functions with conflicting objectives and performance indicators. In this article, the overall system performance is measured by the logistic costs (inventory and emergency supply), the percentage of delayed orders, the percentage of lost sales and the average delay for a customer order. The supply chain function strives to reduce the logistic costs while the sales function aims to satisfy at best the customer requirements.

Due to distant sourcing, the procurement lead time L may be very large and parts supplies are mainly based on forecasts. We name d_t^f the quantity of forecasted demands for the week t. Due to uncertainty, the quantity of real demands can differ largely from the forecasts. Real demands are known gradually and week after week. Moreover, customer orders are not necessarily asked as early as possible. This means that a customer can order a product for a specific week in advance. To define this progressive arrival of demands, we consider the order arrival rates m_k ($0 \leq m_k \leq 1$) and the total demand D_j asked during a week j. During a week i, the firm receives $N_{i,j} = m_{j-i} D_j$ orders that are asked for week j.

When a new customer order is received, it replaces a forecasted demand. Therefore, for every future week t, the production system knows the quantity of forecasts d_t^f and the quantity of real demands d_t^r

to produce during week t. As usual in production planning problem with uncertainty, there is a frozen horizon of F weeks before product assembly ($F \geq 1$). During this frozen horizon, the future production volume is known with certainty and there is no forecast. Hence, we have $\sum_{k=F}^{H} m_k = 1$. This means that all demands asked for a week k are known, at the latest, during week $k - F$.

The main original aspect of this S&OP model relies on the ability for the supply chain department to create so-called sales constraints to restrict the arrival of customer demands in the production planning. A sales constraint d_t^{max} represents the limits of demands (real and forecast) for a week t that the production system can accept for this week. If a constraint is saturated, then new orders asked for this week are delayed to the next week, with a risk of lost sales. The sales constraints are negotiated during the S&OP between the sales and the supply chain functions. In our model, customer are assumed impatient. If an order is delayed by k weeks, then the probability to lose this order equals p_k.

At the end of each week t, the firm procures x_{t+L} parts from its suppliers. All accepted demands have to be satisfied. Therefore, in case of stockout during a week t, missing parts y_t are supplied by emergency with a fast transportation. The production system is never stopped because of parts shortages.

Figure 1 represents the system dynamics of the problem with the balancing flows of supplies, demands and inventories at the end of week t. There are three noticeable horizons: the frozen horizon during which the production plan is fixed, the flexible horizon during which the production plan is partially controlled by sales constraints and the free horizon during which the sales department can order without constraints.

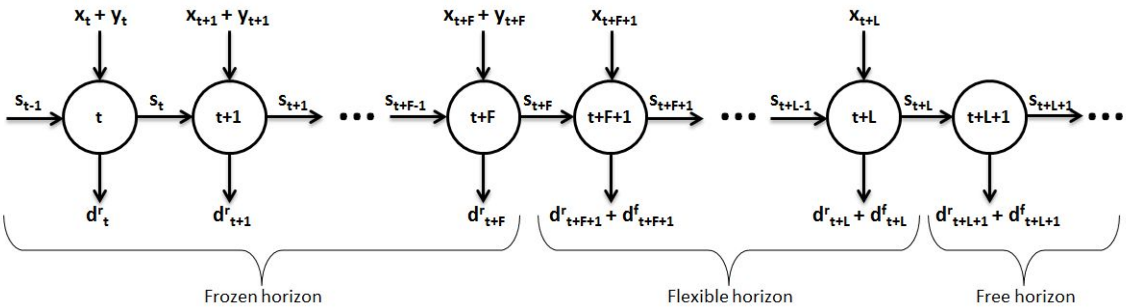

Figure 1: System dynamics: flows of supplies, demands and inventories at the end of week t

In the rest of the paper, we make the following assumptions:

- Procurement lead times are assumed constant, and of length L.
- We consider linear holding and emergency supply costs.
- Fixed cost are neglected.
- There are no capacity restrictions for suppliers or transportation.
- Emergency supplies are always possible, with a negligible lead time (i.e. less than a week).
- There is no inventory limitation.
- The weekly demand follows a uniform distribution.
- The forecast error (difference between real demand and forecast) follows a uniform distribution.

- Customers are impatient. If the order is delayed by a week or more, then there is a probability to lose the order, depending on the delay length.
- Customer demands are independent of the sales constraints.
- Forecasts are never produced. In the frozen horizon, there are only real customer demands.

2.2 Mathematical model formulation

The notations used in this article are summarized in Table 1.

Notation	Description
Input parameters	
H	Number of periods
L	Procurement lead time (normal transportation)
F	Frozen horizon length with $F < L$
c_h	Holding cost per unit per week
c_e	Emergency supply extra cost per unit
s_0	Initial stock level
D_t	Random variable of real demand for week t
F_t	Random variable of demand forecast for week t
$N_{i,j}$	Quantity of new demands asked for week j and received during week i, with $i < j$
m_k	Demand arrival rate, k weeks before real demand, with $1 \leq k \leq H$
p_k	Probability to lose the order if it is delayed by k weeks
Decision variables	
π_t^s	Safety stock margin for week t, with $0 \leq t \leq H$
π_t^f	Flexibility rate for week t, with $0 \leq t \leq H$
System variables	
s_t	Net inventory level of parts at the end of week t, $0 \leq t \leq H$
$\hat{s}_t$	Expected inventory level of parts at the end of week t
x_t	Quantity of parts ordered in week $t - L$ and that will arrive in week t
y_t	Quantity of parts ordered in week t and that will arrive in week t by emergency supply
d_t^f	Quantity of forecasted orders placed in week t
d_t^r	Quantity of real orders placed in week t
d_t^{max}	Sales constraint for week t
b_t	Quantity of delayed orders in week t
l_t	Quantity of lost sales in week t
w	Average delay of a customer order which has been postponed
LC	Average logistic cost per week (inventory and emergency supply)

Table 1: Input parameters and system variables

The sales constraint d_t^{max} for week t is computed based on the expected demand and the flexibility rate π_t^f used during this week, according to Equation (1).

$$d_t^{max} = \left(1 + \pi_t^f\right) \left(d_t^r + d_t^f\right) \qquad \forall t \in \{0, \ldots, H\} \tag{1}$$

The value π_t^f represents the flexibility level offered to the sales department during week t. With a high flexibility, the production system can accept many demands, even if forecasts were low. Conversely, with a low flexibility rate, the number of real demands has to be close to the forecasts otherwise orders are delayed with a risk of lost sales. This flexibility parameter can change every week.

The production system behaves like a traditional materials requirements planning (MRP). To compute the procurement quantity, we need to know the expected inventory level $\hat{s}_{t+L}$ in week $t + L$, that equals the difference between future arrivals of parts (normal replenishments and emergency supplies) and the expected demand (real and forecasted orders), according to Equation (2).

$$\hat{s}_{t+L} = s_0 + \sum_{k=0}^{t+L} \left((x_k + y_k) - \left(d_k^r + d_k^f \right) \right) \qquad \forall t \in \{0, \ldots, H-L\} \tag{2}$$

Then, the procurement quantity x_{t+L} equals the quantity required to satisfy the expected demand plus an additional quantity required to cover demand fluctuation between the end of the frozen horizon and the lead time. This additional quantity is computed based on the expected demand and the safety stock margin π_{t+L}^s used during this week. The value of x_{t+L} is computed according to Equation (3).

$$x_{t+L} = \max \left\{ 0; d_{t+L}^r + d_{t+L}^f + \pi_{t+L}^s \sum_{k=t+F+1}^{t+L} \left(d_k^r + d_k^f \right) - \hat{s}_{t+L-1} \right\} \qquad \forall t \in \{0, \ldots, H-L\} \tag{3}$$

The quantity of emergency supplies for the week t is given in Equation (4).

$$y_t = \max \{0; d_t^r - (s_t + x_t)\} \qquad \forall t \in \{0, \ldots, H\} \tag{4}$$

The average logistic cost is the sum of the average inventory and emergency supply costs. Its expression is given in Equation (5).

$$LC = \frac{1}{H} \sum_{k=0}^{H} (c_h s_k + c_e y_k) \tag{5}$$

The other criteria to optimize are the percentage of delayed orders, the percentage of lost sales and the average delay. There are several ways to formulate the multi-objective optimization models. In this paper, we adopt an ϵ-constraint approach (Chankong and Haimes, 1983) to formulate the optimization problem because this method does not require to normalize the objective functions, that are not easily comparable in our problem. With an ϵ-constraint approach, there is one single objective function and all other criteria to optimize are used to form additional constraints. Therefore, we formulate our problem as follows: the objective is to minimize the logistic costs while satisfying sales requirements in terms of delayed orders, lost sales and average delay. The ϵ-constraint method is a relatively common approach to solve multi-objective optimization problems but other methods (see, for instance, Marler and Arora, 2004) such as weighted sum could be used depending on the context. In our case, the use of weighting coefficients is discouraged because quantifying delays and lost sales to compare them with logistic costs is not relevant in practice.

Therefore, we formulate the optimization problem as follows.

$$\min \quad LC = \frac{1}{H}\sum_{k=0}^{H}(c_h s_k + c_e y_k) \tag{6}$$

$$\text{subject to:} \quad s_t = s_{t-1} + x_t + y_t - d_t^r \qquad \forall t \in \{1,\dots,H\} \tag{7}$$

$$d_t^f + d_t^r \leq d_t^{max} \qquad \forall t \in \{0,\dots,H\} \tag{8}$$

$$\frac{\sum_{k=0}^{H} b_k}{\sum_{i<j} N_{i,j}} \leq B^{max} \tag{9}$$

$$\frac{\sum_{k=0}^{H} l_k}{\sum_{i<j} N_{i,j}} \leq L^{max} \tag{10}$$

$$w \leq W^{max} \tag{11}$$

$$\pi_t^s \geq 0 \qquad \forall t \in \{0,\dots,H\} \tag{12}$$

$$\pi_t^f \geq 0 \qquad \forall t \in \{0,\dots,H\} \tag{13}$$

The values B^{max}, L^{max} and W^{max} represent, respectively, the maximum percentage of delayed orders, percentage of lost sales and average delay that the sales department can accept. The objective function (6) is to minimize the average holding and emergency supply costs. Constraint (7) is the flow balance equation. Constraint (8) represents the sales constraints that limit the positioning of customer demands in the production planning. Constraints (9), (10) and (11) are the ϵ-constraints related to, respectively, the percentage of delayed orders, the percentage of lost sales and the average delay. Constraints 12 and (13) define the validity of decision variables. Note that $\sum_{i<j} N_{i,j}$ represents the total demand received (not necessarily satisfied) during the H periods.

2.3 Policies for managing inventories and flexibility

In theory, the decision variables (π_t^s, π_t^f) (expressed in percentages) can be continuous. However, in practice, it is unnecessary to use decimal values for percentages and the use of integer values (e.g. using 12% instead of 12.43%) does not significantly impact the numerical results. Indeed, the aim of these variables is to compute procurement quantities and sales constraints. The additional accuracy of using continuous variables appears to provide very slight differences for these values. Therefore, in the rest of the paper, we make the choice of discretizing the variables (π_t^s, π_t^f) by using only integer percentage values. A more accurate discretization is always possible. If the benefits are significant, it is also possible to refine the discretization by using successive optimizations with higher accuracy on variables.

We first introduce the class of static policies, which are the most intuitive and easy to implement in practice. A static policy uses fixed percentages to define the stock margins and the flexibility rates for the whole horizon and independent of other parameters. A policy named $\Pi_{S/x/y}$ refers to a static policy with the fixed values of $x\%$ for the stock margin and $y\%$ for the flexibility rate (see equation (14)).

$$\Pi_{S/x/y}: \quad \pi_t^s = x\% \quad \text{and} \quad \pi_t^f = y\% \qquad \forall t \geq 0 \tag{14}$$

We note that the policy $\Pi_{S/0/+\infty}$ is equivalent to a pure build-to-order policy: the inventory system behaves like a classical MRP without safety stock, and there is no restriction on customer orders (except the frozen horizon). With a policy $\Pi_{S/0/+\infty}$, stockouts are avoided by using emergency supplies.

Optimizing a static policy $\Pi_{S/x/y}$ for a given instance consists in finding the best couple (x, y) that satisfies the problem constraints and leads to the lowest logistic cost. As explained before, we only consider integer values for (x, y).

The main advantage of static policies is their simplicity: this facilitates the communication between the different stakeholders in the S&OP because it is simpler to share only few percentage values instead of numerous parameters. But static policies are limited for the following reasons. The main drawback is that the stock margins and the flexibility rates are fixed and expressed as percentages of the expected demand. Hence, the additional procurement quantity and the maximum limit of accepted orders are proportional to the quantity of forecasted demands. Therefore, the static policies suffer from the "base effect". That is, if the forecasts underestimate the real demand, then only a little additional flexibility is given to sales dealers. Conversely, if the forecasts overestimate the real demand, then the assembly plant orders too many parts which may lead to large and unused inventory. To avoid this inconvenience, we introduce a new class of policies, called the linear policies.

The idea of linear policies is to reduce the impact of the "base effect" described above. To do so, every week, the average historical demand is computed and compared with the quantity of expected demand. In case of underestimating forecasts, we offer higher flexibility for sales dealers and we increase the safety stock margin. Conversely, lower values are used in case of overestimating forecasts (to avoid unused flexibility and inventory).

A linear policy is defined by four parameters: a relative coefficient and an absolute value for the safety stock margin and for the flexibility rate. We note that the class of linear policies is a generalization of the threshold policies presented in Lim et al. (2014), where the authors show that the threshold policies appear to perform well.

We use the following notations:

- α_0: absolute value for the safety stock margin
- α_1: relative coefficient for the safety stock margin
- β_0: absolute value for the flexibility rate
- β_1: relative coefficient for the flexibility rate

In this paper, the average historical demand is measured by the moving average over the last M weeks. To compare the forecast with the historical demand, every week t, we compute η_t, the ratio of the future expected demand over the moving average (Equation (15)).

$$\eta_t = \frac{d^r_{t+L} + d^f_{t+L}}{\sum\limits_{k=t-M+1}^{t} \frac{d^f_k}{M}} \qquad \forall t \in \{M, \ldots, H\} \tag{15}$$

Then, the safety stock margins and the flexibility rates of a linear policy Π_L are computed according to Equation (16).

$$\Pi_L : \quad \forall t \geq m, \qquad \begin{aligned} \pi^s_{t+L} &= \max\{0 \ ; \ \alpha_0 - \alpha_1 \eta_t\} \\ \pi^f_{t+L} &= \max\{0 \ ; \ \beta_0 - \beta_1 \eta_t\} \end{aligned} \tag{16}$$

We note that all coefficients are positive and if η_t is high (this means that forecasts are overestimating), then less flexibility is offered to the sales dealers and less safety stock is used. Conversely, there are more flexibility and safety stocks in case of low η_t (this means that forecasts are underestimating).

From a practical point of view, linear policies are more complicated than the static ones because practitioners have to deal with four parameters for each item managed in the S&OP. Also, the policy parameters (absolute and relative coefficients) of a linear policy may appear less intuitive for decision makers. Despite these difficulties, linear policies remain conceivable in practice especially if there are significant benefits and if they can be optimized within a reasonable time frame.

3 A simulation-optimization solution

In this section, the method used to solve the problem is presented. The complexity and the numerous stochastic parameters (demands, forecasts, impatience) of the system make difficult the use of analytical solutions to evaluate the objective functions. Therefore, we consider a simulation-optimization approach which consists in finding the best values of decision variables for a system where the performance is computed based on the output of a simulation model (Ólafsson and Kim, 2002).

3.1 The simulation model

The stochastic objective functions of our problem are evaluated by a simulation model implemented in Java programming language. For a given policy and set of parameters, its aim is to compute logistic costs, the percentage of delayed orders, the percentage of lost sales and the average delay. Each simulation run starts with a warm-up period of W weeks during which all demands are known with certainty. The model simulates the forecast generation, the arrivals of customer orders, the dynamics of sales constraints and delayed orders, the positioning of customer demands in the production planning, the part procurements, the emergency supplies and the creation of sales constraints. Figure 2 illustrates the main steps of the simulation module.

Our model can handle different types of probability distribution for the demand and the forecast accuracy. The classes of policies described in the previous section are implemented in the simulation module. Different types of customer impatience and demand arrivals can also be included in this model. The settings we used in this article are completely detailed in Section 4.

Figure 3 summarizes the different input and ouput parameters.

3.2 The optimization procedure

The aim of the optimization module is to explore efficiently the state-space of solutions for finding the best values for decision variables. A first difficulty is that there is no information about the properties of the different objective functions (the one to optimize and the others which are ϵ-constrained) evaluated by the simulation model. Therefore, we cannot guarantee the optimality of a solution except if we use a complete enumeration but this method is very time consuming.

In this article, we consider several optimization techniques. The aim of our research is to propose an applicable solution for a practical problem. Most of the techniques used in simulation-optimization models rely on local search. To perform a local search, the algorithm needs to compare different solutions. Hence, we first define the quantity γ as follows (Equation (17)).

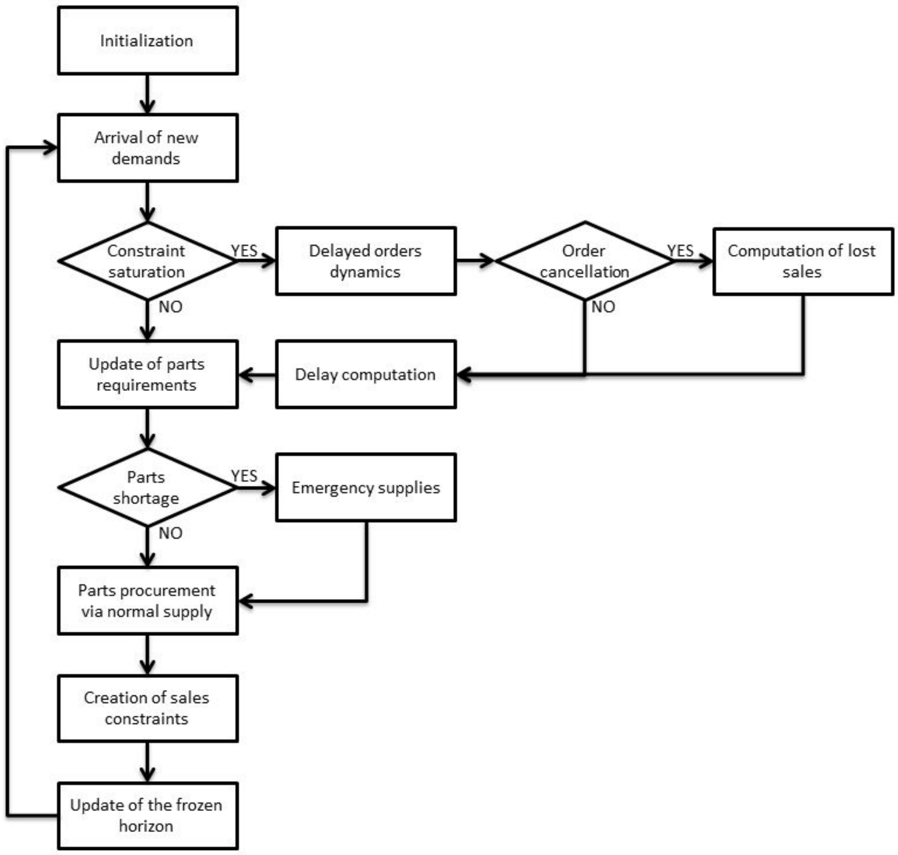

Figure 2: Framework of the simulation model

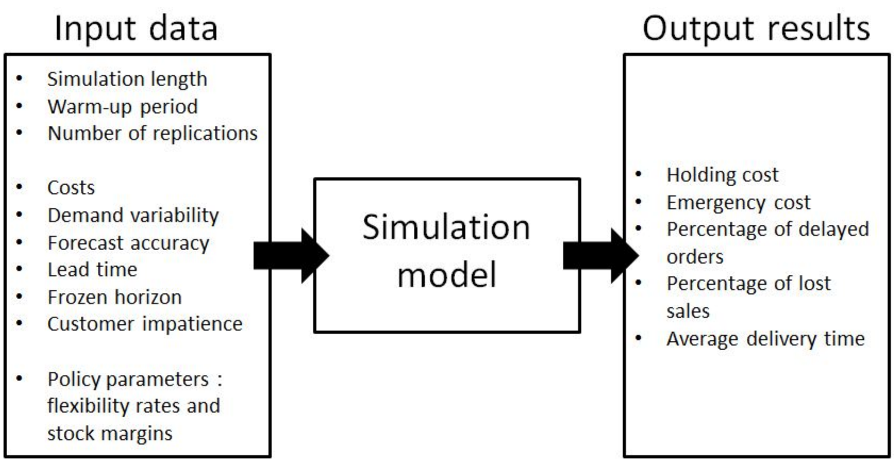

Figure 3: Simulation model: input and ouput parameters

$$\gamma = \max\left\{0\ ;\ \frac{\sum_{k=0}^{H} b_k}{\sum_{i<j} N_{i,j}} - B^{max}\right\} + \max\left\{0\ ;\ \frac{\sum_{k=0}^{H} l_k}{\sum_{i<j} N_{i,j}} - L^{max}\right\} + \max\{0\ ;\ w - W^{max}\} \tag{17}$$

The value of γ is an overall indicator of ϵ-constraint saturation (these constraints are related to the customer satisfaction). If $\gamma = 0$, then all ϵ-constraints are satisfied. If $\gamma > 0$, then at least one ϵ-constraint is not satisfied. This variable γ is used to compare different solutions in the optimization procedure. Indeed, we consider that a solution A is better than a solution B if Statement (18) is verified.

$$\begin{cases} LC(A) \leq LC(B) \\ \gamma(A) = \gamma(B) = 0 \end{cases} \quad \text{or} \quad \begin{cases} \gamma(B) > 0 \\ \gamma(A) \leq \gamma(B) \end{cases} \tag{18}$$

The statement on the left ensures that if two solutions satisfy the ϵ-constraints, then the best is the one with the lowest logistic cost. And the statement on the right says if at least one solution is not feasible, then the best is the one with the lowest γ (the closest for satisfying the constraints).

We consider five different optimization techniques, that are presented below.

- Complete local search (CLS): from a given starting point, the algorithm explores all nearby directions and selects the best one (that reduces logistic costs if sales requirements are satisfied, or that improves sales performances). The process is iterated until finding a local optimum.

- Random local search (RLS): from a starting point, the algorithm explores randomly different nearby directions and moves to the first direction that improves the solution. Contrary to the complete local search, all nearby states are not necessarily explored. This can highly reduce the computation time if there are many variables to optimize. The process is iterated until finding a local optimum.

- Multi-start random local search (MS): the algorithm is identical to the random local search except that it starts from several initial states. This enables the examination of many more solutions.

- Simulated annealing (SA): this well-known metaheuristic is detailed in Metropolis et al. (1953) and in the book of van Laarhoven and Aarts (1987). SA is a common method to solve simulation-optimization problems (Ahmed and Alkhamis, 2002; Rosen and Harmonosky, 2005). The main principle of SA is to explore nearby solutions. If the neighbor is better, then it becomes the current state. Otherwise, the neighbor is accepted with a certain probability that decreases step by step during the optimization.

- "Sequential search" (SEQ): this algorithm is not classical and is specific for the static policies. This algorithm consists in, first, searching the static flexibility rate that meets sales requirements (if it exists) while keeping a fixed safety stock margin. Then, the flexibility rate is fixed to the obtained value and the algorithm searches the first safety stock margin that minimizes logistic costs (if it exists). Otherwise, this algorithm does not give a feasible solution.

The detailed settings of the different optimization algorithms are given in Section 4. The optimization module has been implemented in Java and is directly coupled with the simulation model. Figure 4 summarizes the simulation-optimization framework with the related input and output parameters.

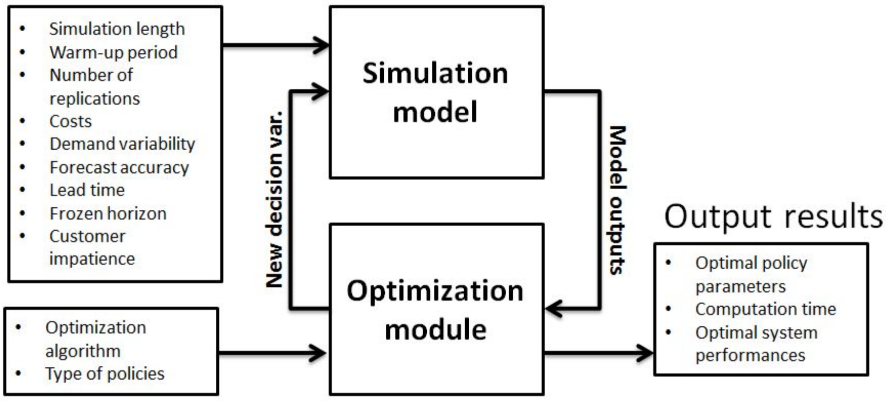

Figure 4: Simulation-optimization framework

4 Numerical results: application to the case study of Renault

We apply the simulation-optimization model to the industrial case of Renault, a French global automobile manufacturer. In this section, we present the experimental design and the problem instances based on industrial data. Then, we provide numerical results to compare the system performance and the computation time of the different optimization algorithms, the static and the linear policies. Practical recommendations are discussed in the last subsection.

4.1 Experimental design

For confidentiality reasons, we cannot provide a test bed based on the exact situation of a vehicle-assembly plant. However, the parameter settings and the 135 instances considered in this research are generated based on real parts data characteristics. In practice, a vehicle-assembly plant may have more than two thousands different parts. A problem instance is characterized by:

- the sales requirements,
- the costs (inventory and emergency supply),
- the demand variability,
- and the forecast accuracy.

For the sales requirements, we consider three levels for the maximum percentage of delayed orders B^{max}: 15% (highly demanding), 20% and 25% (undemanding). The maximal percentage of lost sales L^{max} and the maximal average delay W^{max} are set respectively to 10% and 2 weeks.

For simplicity, we create three price categories and three weight categories to define the extra cost due to emergency supplies. We associate for each category a different extra cost value. Without loss of generality, we set c_h to 1€ per unit per week for each instance. Then, depending on the part's characteristics, the extra cost for emergency supply c_e equals 5, 10, 20, 30 or 50€ per unit, according to Table 2.

The weekly demand is assumed uniformly distributed between D_{min} and D_{max}. The demand variability depends on the part: for instance, very common components (e.g. wipers) have low variability

Price \ Weight	Low	Medium	High
Low	5	10	20
Medium	10	20	30
High	20	30	50

Table 2: Cost of one emergency supply c_e in € per unit depending of the part's price and weight

contrary to specialized parts (e.g. cables, batteries). The values used in the numerical study are given in Table 3.

Demand variability	D_{min}	D_{max}
Low	200	400
Medium	100	500
High	50	550

Table 3: Demand variability

The relative difference between real demand and forecast is assumed uniformly distributed between ±60%, ±80%, ±100% depending on the forecast accuracy. In practice, the forecast accuracy depends on the procurement lead time and the part (common or specific).

In total, the test bed is composed of 135 different instances (3 sets of values for the sales requirements, 5 for the costs, 3 for the demand and 3 for the forecast).

4.2 Simulation settings

We use the following default simulation parameters (Table 4) for the numerical experiment. These values are also based on industrial data.

Parameter	Value	Definition
H	2000	Simulation length
R	50	Number of replications
W	15	Warm-up period
L	10	Procurement lead time
F	4	Frozen horizon
s_0	0	Initial stock
M	6	Moving average for linear policies
m_k	see A	Demand arrival rates
p_k	see B	Customer impatience

Table 4: Default simulation parameters

With a personal laptop (HP ProBook 6450b Intel® Celeron® 2×2 Ghz, 4 Go RAM memory), it takes less than one second to compute a simulation run of 2000 weeks. To obtain reliable simulated results, we calibrate the simulation length to 2000 weeks and the number of replications to 50 runs. By using these settings, the relative error of the logistic cost is less than 0.5% compared to a limit-situation of 5000 replications of 4000 weeks (see Lim et al., 2014, for detailed simulation results).

4.3 Optimization settings

Since the linear policies have more variables to optimize than the static ones, the use of CLS algorithms is too time consuming for linear policies. Therefore, we did not implement this technique for the class of linear policies. We also do not implement MS and SA algorithms for the static policies because it appears that we can already obtain very good results with the other algorithms, which are simpler and faster. The SEQ algorithm cannot be implemented for linear policies because it has been created specifically for static policies.

Before exploring the state-space, we define two boundaries, σ and τ, for the variables used to compute, respectively, the safety stock margins and the flexibility rates. These boundaries should be sufficiently high to explore good solutions. Based on initial numerical experiments, we set $\sigma = 50\%$ and $\tau = 200\%$. Therefore, we explore the state-space of solutions with values lower or equal to σ (respectively τ) for x, α_0 and α_1 (respectively y, β_0 and β_1). Table 5 summarizes the settings (starting points and stop criteria) used for each optimization technique.

Algorithm	Starting point	Stop criteria	Policies
CLS (complete local search)	$(x, y) = (\sigma/2, \tau/2)$	local optimum	static
RLS (random local search)	$(x, y) = (\sigma/2, \tau/2)$ or $(\alpha_0, \alpha_1, \beta_0, \beta_1) = (\sigma/2, \sigma/2, \tau/2, \tau/2)$	local optimum	static and linear
MS (multi-start)	$(\alpha_0, \alpha_1, \beta_0, \beta_1) = (\sigma/2, \sigma/2, \tau/2, \tau/2)$ $(\alpha_0, \alpha_1, \beta_0, \beta_1) = (0, 0, 0, 0)$ $(\alpha_0, \alpha_1, \beta_0, \beta_1) = (\sigma, \sigma, \tau, \tau)$ $(\alpha_0, \alpha_1, \beta_0, \beta_1) = (0, \sigma/2, 0, \tau/2)$	local optimum	linear
SA (simulated annealing)	$(\alpha_0, \alpha_1, \beta_0, \beta_1) = (\sigma/2, \sigma/2, \tau/2, \tau/2)$	computation time	linear
SEQ (sequential search)	$(x, y) = (0, 0)$	local optimum	static

Table 5: Default settings for the optimization algorithms

For the SA, the algorithm stops if the computation time exceeds a certain value. To determine this value, we first compute the average time required by the RLS algorithm for solving the simulation-optimization problem (results are given in the following subsections). We stop the SA algorithm when the computation time exceeds three times the average of the RLS algorithm. Indeed, we give more time to the SA because the algorithm can escape from a local optimum but it requires more time for exploring deteriorated solutions. Moreover, the choice of the temperature is important for the performance of the SA. After several trials, we decided to use a temperature that linearly decreases because it is relatively simple and seems to perform well for our case study. More research can be conducted on the configuration of algorithms to improve their performance but this is out of the scope of this article.

4.4 Preliminary results

First, as preliminary results, we show that the proportion of good solutions in the state-space may be highly variable from one instance to another in our case study. We present two instances where linear policies are used, to illustrate this phenomenon. We compute all possible solutions over a given state-space by using a simple enumeration procedure. Table 6 shows the proportion of solutions that performs relatively well compared to the optimal solution.

In instance A, there are 2% of feasible solutions that give an average logistic cost very close (less than 1%) to the optimal one. And about 26% of feasible solutions provide an average logistic cost less than 5% to the optimal one. This means that many solutions appear to be efficient in this instance and therefore,

	Instance A		Instance B	
	Number	%	Number	%
Feasible solutions	4959	100%	1678	100%
Logistic cost ≤ optimal+1%	85	2%	3	0%
Logistic cost ≤ optimal+2%	370	7%	6	0%
Logistic cost ≤ optimal+5%	1277	26%	36	2%
Logistic cost ≤ optimal+10%	2023	41%	135	8%
Logistic cost ≤ optimal+15%	2585	52%	284	17%
Logistic cost ≤ optimal+20%	2893	58%	544	32%

Table 6: Proportion of good solutions: examples with two instances

it is relatively easy to find good policy parameters. However, this result does not hold for the instance B. As we can see in Table 6, about 2% of feasible solutions give an average logistic cost relatively close to the optimal. In this example, it might be difficult to find a good solution, especially if there are many local optimums that are distant from the global one. These two examples illustrate that the proportion of good solutions in the state-space may be very variable from one instance to another. Moreover, we do not have information on the structure of the objective functions (evaluated by simulation). These reasons justify the test of different methods to search for the best solution.

4.5 Algorithms performance for static policies

In this subsection, the different methods to optimize static policies are compared. The average computation time and the obtained logistic costs are measured. Results are given in Table 7. We also compare the obtained results with the best one, for both computation time and logistic cost. Results of the worst and the best instances are also given in the Min-Max column.

Algorithm	Time (sec.)	vs. best	Min - Max	Logistic cost	vs. best	Min - Max
CLS	280.22	146.2%	77.43 - 621.94	516.31	0.0%	260.75 - 1042.02
RLS	201.99	77.4%	75.43 - 424.65	518.35	0.4%	267.01 - 1046.04
SEQ	113.83	0.0%	17.04 - 326.67	522.55	1.2%	281.77 - 1044.74

Table 7: Algorithms performance over the 135 instances: static policies

In average over the 135 instances, the algorithms take between 114 and 280 seconds to optimize one instance with static policies. In the worst case, it takes about 10 minutes for the CLS algorithm. In practice at Renault, this optimization process is done, at most, every week for each component that is managed by the S&OP with flexibility. Depending on the assembly-plant and the situation, the S&OP may require between 10 and 30 instance calculations per week. For practitioners, it is largely sufficient if the method takes less than one hour to compute one instance with a personal laptop. Indeed, calculations can be performed on dedicated mainframes. Therefore, for static policies, the computation time is largely acceptable for the implementation of this simulation-optimization approach in practice.

The results show that the fastest algorithm to optimize static policies is the SEQ (sequential search). The RLS (random local search) requires about 77% more time, but in absolute values it only represents 90 seconds more which remains reasonable for our industrial application. As we could expect, the worst algorithm in terms of computation time is the CLS (complete local search). However, the CLS leads to the best logistic cost. There are slight differences with RLS and SEQ in terms of logistic cost (respectively +0.4% and +1.2%). The RLS and SEQ methods perform relatively well and improve significantly the calculation time.

4.6 Algorithms performance for linear policies

Similar to static policies, we compare, in this subsection, the different optimization techniques for the class of linear policies. Numerical results are given in Table 8.

Algorithm	Time (sec.)	vs. best	Min - Max	Logistic cost	vs. best	Min - Max
RLS	1115.31	0.0%	613.24 - 2504.63	497.40	2.8%	260.34 - 1011.09
MS	4226.11	278.9%	2413.81 - 6221.18	483.71	0.0%	263.77 - 989.76
SA	3401.95	205.0%	3240.48 - 3618.12	490.76	1.5%	264.06 - 1008.91

Table 8: Algorithms performance over the 135 instances: linear policies

The algorithms take between 1115 and 4226 seconds (or 19 and 70 minutes) in average to optimize one instance with linear policies. As we could expect, the fastest technique is the RLS (random local search) algorithm and it could be used in practice if there are not many components to manage and not many changes in sales requirements per week. The MS (multi-start) algorithm is far more slower than RLS because it uses multiple starting points to explore the state-space. If it provides the best performance in terms of cost, it is too slow to be used in practice for our industrial case. The SA (simulated annealing) method could be a good trade-off between computation time and the cost performance. Indeed, the cost difference is only 1.5% with the best one (MS) but it takes on the average 14 minutes less. It may be possible to improve these performances by using different settings for the temperature schedule, the stop criteria and the starting point.

4.7 Comparison of static and linear policies

In this subsection, we compare the performances of static and linear policies. We start with an overall comparison and then, we detail the differences between these policies, as a function of system parameters.

The table 9 compares the best (in terms of cost) and the fastest (in terms of time) policies. For the static (respectively linear) policies, the lowest cost is obtained by the CLS (respectively MS) algorithm and the fastest method is the SEQ (respectively RLS) algorithm.

Policy	Time	vs. best	Min - Max	Logistic cost	vs. best	Min - Max
Best static	280.22	146.2%	77.43 - 621.94	516.31	6.7%	260.75 - 1042.02
Fastest static	113.83	0.0%	17.04 - 326.67	522.55	8.0%	281.77 - 1044.74
Best linear	4226.11	3612.7%	2413.81 - 6221.18	483.71	0.0%	263.77 - 989.76
Fastest linear	1115.31	879.8%	613.24 - 2504.63	497.40	2.8%	260.34 - 1011.09

Table 9: Global comparison of static and linear policies over the 135 instances

We expect that the static policies are faster to optimize than the linear ones because they have twice less variables to optimize. The results show that, on the average, the difference is very high and it may take up to 37 times longer to optimize a linear policy with a slow algorithm compared to a static policy. Regarding the logistic cost, the linear policies are significantly better than the static ones while satisfying the same sales requirements. For instance, using a static policy with a fast algorithm leads to a cost increase of 8% compared to the best linear policy. In real life, this can represent very significant logistic costs for a vehicle-assembly plant.

In the following results, we compare the best static policy with the best linear one (computed with, respectively, the algorithms CLS and MS) in terms of cost performance. We show that the outperformance of linear policies mainly depends on the system parameters. Table 10 details the policies performances depending on the structure of the demand (variability and forecast accuracy). Table 11 details the results depending on the sales requirements and the emergency cost.

Forecast accuracy	Demand variability	Static policy	Linear policy	Relative difference	Number of instances
High (±60%)		363.55	349.05	-4.0%	45
	Low (200-400)	380.43	361.93	-4.9%	15
	Medium (100-500)	364.94	350.24	-4.0%	15
	High (50-550)	345.28	334.98	-3.0%	15
Medium (±80%)		513.23	479.95	-6.5%	45
	Low (200-400)	536.58	505.86	-5.7%	15
	Medium (100-500)	516.65	483.26	-6.5%	15
	High (50-550)	486.45	450.73	-7.3%	15
Low (±100%)		672.16	622.12	-7.4%	45
	Low (200-400)	695.85	645.68	-7.2%	15
	Medium (100-500)	672.19	620.32	-7.7%	15
	High (50-550)	648.43	600.37	-7.4%	15

Table 10: Static vs linear: average logistic cost depending on the demand

As we can see in Table 10, the cost reduction by using linear policies is higher when the forecasts are less accurate. This result is not surprising because the aim of linear policies is to reduce the inconvenience of the "base effect" previously described, which is stronger when the difference between forecast and real demand is high. These results also show that the demand variability does not seem to have a strong impact on the outperformance of linear policies.

Sales requirements	Emergency cost c_e	Static policy	Linear policy	Relative difference	Number of instances
High ($P_d \leq 15\%$)		569.04	541.00	-4.9%	45
	5	353.21	342.27	-3.1%	9
	10	456.82	441.38	-3.4%	9
	20	596.64	565.92	-5.1%	9
	30	675.16	634.83	-6.0%	9
	50	763.40	720.62	-5.6%	9
Medium ($P_d \leq 20\%$)		514.16	480.69	-6.5%	45
	5	349.72	344.06	-1.6%	9
	10	437.09	419.38	-4.1%	9
	20	537.04	501.58	-6.6%	9
	30	594.53	544.30	-8.4%	9
	50	652.41	597.12	-8.9%	9
Low ($P_d \leq 25\%$)		465.73	429.43	-7.8%	45
	5	347.43	342.55	-1.4%	9
	10	420.58	409.09	-2.7%	9
	20	481.57	439.35	-8.8%	9
	30	521.27	465.51	-10.7%	9
	50	557.81	490.66	-12.0%	9

Table 11: Static vs linear: average logistic cost depending on the sales requirements and emergency cost

Table 11 highlights interesting results. The outperformance of linear policies is higher when the sales requirements are low (4.9% cost reduction when requirements are very demanding and 7.8% when they are not). This means that if the company wants to have a high customer satisfaction, the advantage of linear policies is less important. Indeed, with strong required levels on the percentage of delayed orders, the optimization problem is very constrained and feasible solutions are very costly because emergency supplies are often required. This leads to an important increase of logistic costs, and the benefits of using linear policies are less significant. Moreover, as we can see in this table, the benefits obtained with linear policies depends heavily on the emergency supply cost. For instance, if the sales require less than 25% delayed orders, then the outperformance of linear policies vary between 1.4% and 12.0% depending on the emergency costs. The more expensive the emergency supply, the larger the gain is of using linear policies.

4.8 Managerial insights

The previous results suggest that decision makers can expect significant gains by using linear policies instead of fixed values for the safety stock margins and the flexbility rates. However, the computation time required to optimize linear policies could be a major obstacle to overcome, if the practitioners need to compute and update the parameter values frequently. In the automotive industry, this situation may exist, especially for new models when the demand is highly volatile and depends on many external factors, or for highly customized vehicles that require numerous different components. However, for models with less variety and less changes in bill-of-materials, linear policies are manageable and should be considered by decision makers because they provide significant gains. The linear policies are a bit more complicated than the static ones and this may complicate the coordination of sales and operations. Therefore, if linear policies have clear advantages from a theoretical point of view, in practice they may lead to difficult managerial issues, depending on the industrial context. To remediate this, it could be possible to start with static policies to facilitate the adoption of these new S&OP processes and then switch to linear policies for critical products or components, or once practitioners are at ease with sales constraints and flexibility.

The choice of the type of policies and also the optimization technique depends on the situation of the company, especially the emphasis on cost performance and computation time. We note that it is completely possible to use different types of policies or algorithms depending on the component or the product. In this case, our study suggests that one should better implement static policies for parts with low emergency cost and good forecast accuracy. Otherwise linear policies should be preferred. The use of different types of policies depending on the parts seems to be a good (but more sophisticated and difficult to manage) solution for dealing with numerous and various components. To do so, the company should categorize their parts and associate for each category a type of policy and an optimization technique.

The different algorithms tested in this paper appear to perform relatively well in terms of cost performance. For the static policies, the choice of the optimization method is not crucial because, for all algorithms, the computation times are reasonable and the cost performances are relatively close. In this case, we suggest to use the CLS algorithm that provides the best solution in terms of logistic costs while satisfying the same levels of customer satisfaction. However, for the linear policies, the choice of the method should be carefully examined because the calculation time can vary greatly from one to another. The cost difference between the different algorithms is also larger for linear policies than static policies. Our study suggests that SA is the best algorithm if the computation time is not a priority for practitioners. Otherwise the RLS algorithm can be used as it represents a good trade-off between the logistic cost and the computation time performances.

We also note that the aim of the proposed simulation-optimization solution is not to be a black-box tool that automatically defines the best parameter values. The objective is mainly to provide insights and recommendations for decision makers for helping them to set up several strategic parameters on flexibility and safety stocks. Our solution should be considered as a decision-aid tool. And it can also be used for benchmark studies and to try different scenarios to evaluate potential benefits.

5 Conclusion and research perspectives

The sales and operations planning is a challenging issue for improving customer satisfaction and controlling production costs. Globalization and the uncertain environment incitate the companies to adapt and improve the coordination between sales and supply chain functions. In this article, we present an original S&OP model with flexibility rates for managing the compromise between sales requirements and industrial constraints. We propose a multi-objective and ϵ-constrained model solved by a simulation-optimization approach to compute the best policy values for controlling parts inventory and sales flexibility. To the best of our knowledge, this research is the first that proposes a multi-objective simulation-optimization model for solving a S&OP problem with sales flexibility, uncertain demand and impatient customers. Our research is particularly relevant for industries that face a highly uncertain environment and demanding customers. Moreover, our ϵ-constraint formulation for the optimization problem is also relevant when the objectives of sales function are not easily comparable with the supply chain objectives, or when it is difficult to prioritize them, which is mostly the case in practice.

The contribution of this paper is threefold. First, we provide a new multi-objective model and an efficient simulation-optimization solution which is easy to implement in practice. Second, two classes of policies (static and linear) for managing parts procurement and flexibility are presented. Third, we apply our model and solution on the case study of Renault, a global automanufacturer. We provide a numerical study based on industrial data. Several policies and optimization techniques (local search, simulated annealing etc.) are compared in terms of computation time and system performance. Results show that significant gains can be obtained with the linear policies. Moreover, computation times and system performance can significantly differ from one algorithm to another. Managerial insights are derived from this numerical experiment. We also give practical recommendations to implement our solution in real life.

Our research also shows that the outperformance of linear policies depends heavily on the system parameters. We highlight that all the optimization techniques used in this article perform relatively well for the static policies. Some algorithms require a high computation time for linear policies. Depending on the industrial context, these calculation times can be acceptable or not. More research can be conducted on the setup of the optimization algorithms to improve the time performance. Furthermore, as often with simulation-optimization models, the computation time is mostly due to the simulation module used to evaluate the objective functions. New methods could also be investigated to reduce the time spent in simulation.

Many other research directions are possible from this study. The use of another optimization formulation than our ϵ-constraint model can lead to different results. This extension would be valuable for companies where the ϵ-constraint approach is not relevant. Furthermore, in this research, we only investigate static and linear policies to manage sales and inventory but more extensive comparison with other inventory strategies may be conducted.

Acknowledgements

The authors thank Alain Benichou for his support and guidance. They also thank the ANRT and the automobile manufacturer of this research paper for their financial support.

References

M. A. Ahmed and T. M. Alkhamis. Simulation-based optimization using simulated annealing with ranking and selection. *Computers & Operations Research*, 29(4):387–402, 2002.

M. H. Alrefaei and A. H. Diabat. A simulated annealing technique for multi-objective simulation optimization. *Applied mathematics and computation*, 215(8):3029–3035, 2009.

S. Andradóttir. An overview of simulation optimization via random search. *Handbooks in operations research and management science*, 13:617–631, 2006.

P. G. Brabazon, B. L. MacCarthy, A. Woodcock, and R. W. Hawkins. Mass customization in the automotive industry: comparing interdealer trading and reconfiguration flexibilities in order fulfillment. *Production and Operations Management*, 19(5):489–502, 2010.

V. Chankong and Y. Y. Haimes. *Multiobjective decision making: theory and methodology.* Number 8. North-Holland, 1983.

P. Childerhouse, S. M. Disney, and D. R. Towill. On the impact of order volatility in the European automotive sector. *International Journal of Production Economics*, 114(1):2–13, 2008.

S. Elias. New car buyer behaviour. *3DayCar Research Report, Cardiff Business School*, 2002.

D. A. Elkins, N. Huang, and J. M. Alden. Agile manufacturing systems in the automotive industry. *International Journal of Production Economics*, 91(3):201–214, 2004.

M. C. Fu. Optimization via simulation: a review. *Annals of Operations Research*, 53(1):199–248, 1994.

M. C. Fu. Optimization for simulation: theory vs. practice. *Journal on Computing*, 14(3):192–215, 2002.

M. C. Fu. Gradient estimation. *Handbooks in operations research and management science*, 13:575–616, 2006.

M. C. Fu, F. Glover, and J. April. Simulation optimization: a review, new developments, and applications. In *Proceedings of the 32nd conference on Winter simulation*, pages 83–85. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey, USA, 2005.

F. Glover, J. P. Kelly, and M. Laguna. New advances for wedding optimization and simulation. In *Proceedings of the 1999 Winter Simulation Conference*, volume 1, pages 255–260. IEEE, 1999.

J. A. Grimson and D. F. Pyke. Sales and operations planning: an exploratory study and framework. *The International Journal of Logistics Management*, 18(3):322–346, 2007.

J. Haddock and J. Mittenthal. Simulation optimization using simulated annealing. *Computers & Industrial Engineering*, 22(4):387–395, 1992.

C. K. Hahn, E. A. Duplaga, and J. L. Hartley. Supply-chain synchronization: lessons from Hyundai Motor Company. *Interfaces*, 30(4):32–45, 2000.

M. Holweg, S. Disney, P. Hines, and M. Naim. Towards responsive vehicle supply: a simulation-based investigation into automotive scheduling systems. *Journal of Operations Management*, 23(5):507–530, 2005.

M. Howard, P. Powell, and R. Vidgen. Automotive industry information systems: from a mass production to build-to-order. *Journal of Cases on Information Technology*, 7(2):16–30, 2005.

B. Keskin, S. Melouk, and I. Meyer. A simulation-optimization approach for integrated sourcing and inventory decisions. *Computers & Operations Research*, 37(9):1648–1661, 2010.

J. P. Kleijnen. Response surface methodology for constrained simulation optimization: An overview. *Simulation Modelling Practice and Theory*, 16(1):50–64, 2008.

J. Li, M. Gonzalez, and Y. Zhu. A hybrid simulation optimization method for production planning of dedicated remanufacturing. *International Journal of Production Economics*, 117(2):286–301, 2009.

L. L. Lim, G. Alpan, and B. Penz. Reconciling sales and operations management with distant suppliers in the automotive industry: a simulation approach. *To appear in International Journal of Production Economics*, 2014.

B. MacCarthy, P. G. Brabazon, and J. Bramham. Fundamental modes of operation for mass customization. *International Journal of Production Economics*, 85(3):289–308, 2003.

R. T. Marler and J. S. Arora. Survey of multi-objective optimization methods for engineering. *Structural and multidisciplinary optimization*, 26(6):369–395, 2004.

N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21:1087, 1953.

J. Miemczyk and M. Holweg. Building cars to customer order - what does it mean for inbound logistics operations? *Journal of Business Logistics*, 25(2):171–197, 2004.

H. G. Neddermeijer, G. J. van Oortmarssen, N. Piersma, and R. Dekker. A framework for response surface methodology for simulation optimization. In *Proceedings of the 32nd conference on Winter simulation*, pages 129–136. Society for Computer Simulation International, 2000.

S. Ólafsson. Metaheuristics. *Handbooks in operations research and management science*, 13:633–654, 2006.

S. Ólafsson and J. Kim. Simulation optimization. In *Proceedings of the 2002 Winter Simulation Conference*, volume 1, pages 79–84. IEEE, 2002.

K. Ramdas. Managing product variety: an integrative review and research directions. *Production and Operations Management*, 12(1):79–101, 2003.

D. Rani and M. M. Moreira. Simulation–optimization modeling: a survey and potential application in reservoir systems operation. *Water resources management*, 24(6):1107–1138, 2010.

S. L. Rosen and C. M. Harmonosky. An improved simulated annealing simulation optimization method for discrete parameter stochastic systems. *Computers & Operations Research*, 32(2):343–358, 2005.

T. Stäblein, M. Holweg, and J. Miemczyk. Theoretical versus actual product variety: how much customization do customers really demand? *International Journal of Operations & Production Management*, 31(3):350–370, 2011.

J. R. Swisher, P. D. Hyden, S. H. Jacobson, and L. W. Schruben. A survey of simulation optimization techniques and procedures. In *Proceedings of the 2000 Winter Simulation Conference*, volume 1, pages 119–128. IEEE, 2000.

J. R. Swisher, P. D. Hyden, S. H. Jacobson, and L. W. Schruben. A survey of recent advances in discrete input parameter discrete-event simulation optimization. *IIE Transactions*, 36(6):591–600, 2004.

C. S. Tang. Perspectives in supply chain risk management. *International Journal of Production Economics*, 103(2):451–488, 2006.

A. M. T. Thomé, L. F. Scavarda, N. S. Fernandez, and A. J. Scavarda. Sales and operations planning: A research synthesis. *International Journal of Production Economics*, 138(1):1–13, 2012.

A. M. T. Thomé, R. S. Sousa, and L. F. R. R. Scavarda do Carmo. The impact of sales and operations planning practices on manufacturing operational performance. *International Journal of Production Research*, (ahead-of-print):1–14, 2013.

T. Tomino, Y. W. Park, P. Hong, and J. J. Roh. Market flexible customizing system (MFCS) of Japanese vehicle manufacturers: an analysis of Toyota, Nissan and Mitsubishi. *International Journal of Production Economics*, 118(2):375–386, 2009.

P. J. van Laarhoven and E. H. Aarts. *Simulated Annealing: Theory and Applications*, volume 37. Springer, 1987.

T. Volling and T. S. Spengler. Modeling and simulation of order-driven planning policies in build-to-order automobile production. *International Journal of Production Economics*, 131(1):183–193, 2011.

Q. Zeng and Z. Yang. Integrating simulation and optimization to schedule loading operations in container terminals. *Computers & Operations Research*, 36(6):1935–1944, 2009.

A Demand arrival rates

We use the following parameter values for the demand arrival rates m_k (see Table 12). These values are estimated based on industrial data.

k	m_k
$k < F$	0
$k = F$	0.40
$k = F + 1$	0.30
$k = F + 2$	0.15
$k = F + 3$	0.10
$k = F + 4$	0.05
$k > F + 4$	0

Table 12: Parameter values for the demand arrival rates for the case study of Renault

B Customer impatience

We use the following parameter values for the probabilities to lose a customer order after a certain delay (see Table 13). These estimations are based on marketing studies carried out by Renault. We note that for any delay higher than seven weeks, the customer order is lost.

k	p_k
0	0
1	0.05
2	0.10
3	0.20
4	0.30
5	0.50
6	0.70
7	0.90
$k > 7$	1.00

Table 13: Parameter values for the customer impatience for the case study of Renault

Les cahiers Leibniz ont pour vocation la diffusion des rapports de recherche, des séminaires ou des projets de publication sur des problèmes liés au mathématiques discrètes.

Responsables de la publication : Nadia Brauner et András Sebő
ISSN : 1298-020X -