# Optimizing Stochastic Supply Chains via Simulation: What is an Appropriate Simulation Run Length?

**Arreola-Risa A[1], Fortuny-Santos J[2], Vintró-Sánchez C[3]**

**Abstract** The most common solution strategy for stochastic supply-chain management problems that are analytically intractable is simulation. But, how can we be sure that the optimal solution obtained by simulation is in fact the true optimal solution? In this paper we try to shed light on this question. We report the results of an extensive simulation study of a base-stock controlled production-inventory system. We tried different values of base-stock levels (R) to determine, via simulation, which was the value that minimized the total inventory holding and backordering costs per period. For 25 different cases (and 100 replications each), we compared the optimal solution obtained from simulation (Rs*) with the true optimal base-stock level (Ra*) obtained from an analytical result, with the goal of obtaining a lower bound of 95% matches. Results show that when the traffic intensity increases, the run length necessary to achieve a minimum of 95% matches increases too, and when the backorder cost increases, the number of matches decreases for each specific run length. In most of the cases simulated, 100,000 demands were enough to achieve reasonably reliable results.

**Keywords:** Stochastic Supply Chains, Optimization, Simulation

[1] Antonio Arreola-Risa
Mays Business School. Texas A&M University, College Station (TX), USA.

[2] Jordi Fortuny-Santos (✉e-mail: jordi.fortuny@upc.edu)
Business Organization Department, Technical University of Catalonia, Manresa, Spain.

[3] Carla Vintró-Sánchez
Business Organization Department, Technical University of Catalonia, Manresa, Spain.

# 1 Introduction

There is a vast literature on inventory control in production-inventory systems. Many models have been created in order to address issues such as lead time demand, advanced information, or optimal policies. Nowadays, these issues have gained relevance, because we now consider not an isolated company but a chain of companies commonly called a supply chain.

Difficulties arise when one tries to solve the above-mentioned models for supply chains. Equations including convolutions and mixtures of random variables soon become intractable and out of the reach of analytic tools. Consequently, these intractable models are usually tested by means of simulation. Researchers develop their own computerized simulators or use commercial applications such as ARENA or PROMODEL. But, how can we be sure that the results yielded by the simulator correspond to the true optimal solution? This question has been studied by researchers in other settings for a long time, especially from the point of view of Statistics, and different tools have been developed; however, those statistical tools are not applicable to our problem setting because the decision variables of interest are discrete. This question is also of interest for practitioners because simulation is widely used by companies in the real world. In fact, this research arises from the needs of a global energy company that required frequent use of simulation for decision making on a supply chain project. Optimization results had to be reliable and accurate but the simulation run length had to be reasonably short. In this paper, we try to shed light on the problem of finding a relationship between simulation run length and the reliability of a simulation optimization results. We have selected a stylized model of a supply chain studied by Arreola-Risa (1996), where demand is a Poisson process, unit manufacturing times are exponentially distributed, inventory is controlled by a base-stock level, and we are interested in the steady-state behaviour of the supply chain. There are two reasons for our selection of such model. First, the studied model is ergodic and hence a steady-state is guaranteed to exist, and to the best of our knowledge, it is one of a few instances where the optimal value of the base-stock level has been determined analytically. Hence we can compare the optimal base-stock level obtained via simulation to the truly optimal base-stock level. Second, given that Poisson demand and exponential unit manufacturing times represent extreme cases of variability in real-world situation (Knott and Sury 1987, and Buzacott and Shanthikumar 1993), our results represent upper bounds for practical settings that researchers may study and practitioners may find in practice.

## 2 Literature Review

The relation between simulation run length and accuracy of the results is qualitatively well known. Broadly speaking, the longer the simulation, the better results we get. It is based on the fact that, in general, as sample size increases, estimates improve in statistical accuracy and, when it comes to queuing problems, they have a transient period and then a steady state which is necessary to reach to collect meaningful statistics from the simulation. The longer the simulation run, the higher the probability that the steady state has been reached.

The problem of convergence for an M/M/1 system is analyzed by Fishman (1978). Fishman (1971) had already set forth an autoregressive approach for estimating the sample size of a process with a specified level of precision in simulation experiments. That procedure was intended to determine when enough data had been collected in ongoing experiments.

Whitt (1989) finds heuristics to be used to estimate run lengths before the simulation starts. For the particular case of an M/M/1 model, results show that, with a 5% error and a 95% confidence, a length of 27,500 periods is enough for a traffic intensity ($\rho$) of 0.8 while 592,000 periods are necessary for a level of $\rho$ equal to 0.95.

As stated above, in our research, we simulate a base-stock controlled production-inventory system. Under this approach, we have a warehouse (or an inventory supermarket in lean manufacturing) that has an amount, called base-stock level, of inventory ($R_s$) of a certain item. Demand for that item, stochastic in nature, is immediately served -if enough units are in stock, and a replenishment order is immediately placed to the production system (i.e. by means of kanban cards, or the equivalent, in lean companies where production is adjusted to demand).

The behavior of a base-stock policy has been studied by Clark and Scarf (1960), Zipkin (1991) and Veatch and Wein (1994) amongst others. And besides the analytical approach, base-stock level has also been used in simulation. Liberopoulos and Koukoumialos (2005) determine the optimal base-stock level using simulation. In each experiment the authors used a simulation run length of 60 million time units. That yielded 95% confidence intervals on the estimated values of cost and base-stock level with half width values of less than 0.5% or 4% of their respective estimated values. Snyder and Schen (2006) test the supply and demand uncertainty using base-stock level. They made 10 independent replications with 10,000 periods each and 100 warm-up periods.

## 3 Research Approach: a Simulation Experiment

In order to determine the necessary simulation run, we have carried out a simulation experiment. We have designed our own computer simulator in C++ language. Parameter input and data output are captured in EXCEL spreadsheets.

We have simulated a base-stock controlled production-inventory system where a warehouse has an initial amount of inventory ($R$). Demand is stochastic and each arriving demand is for one item. When a demand arrives, it is immediately served -if possible- following the first-come first-served (FCFS) rule, and an order is placed to the production system. After production is completed, the item is instantaneously delivered to the warehouse. Demands that cannot be fulfilled because the warehouse is temporarily out of stock are backordered.

In this model, we change the value of $R$ in order to find which one minimizes the sum of inventory holding and demand backordering costs. That value is the optimal base-stock level ($R_s^*$). We have assumed that demand is a Poisson process and unit manufacturing time is exponentially distributed. There are two main reasons for these assumptions. The first one is that this problem has been solved analytically and therefore we know which is the optimal solution that should be obtained via simulation. The second one is that the Poisson and exponential distributions represent extreme practical cases of randomness and therefore if a simulation run length is valid for such disperse results, it will be an upper bound for other models with less variability.

We have considered different levels of backorder cost per unit at the warehouse ($p$): 2, 4, 8, 16 and 32 times the holding cost per unit per unit time ($h$). The relationship between supply and demand is modeled in terms of the traffic intensity ($\rho$) defined as the average arrival rate ($\lambda$) divided by the average service rate ($\mu$). Five traffic intensity values were used: 0.1, 0.3, 0.5, 0.7 and 0.9. The combination of cost ratios ($p/h$) and traffic intensities ($\rho$) gives 25 different cases to be simulated (Table 1). For example, case 13 represents a product being manufactured in a facility that operates at capacity utilization of 50% and with a backorder cost rate that is 8 times the holding cost rate.

**Table 1** Simulation cases

| | Traffic intensity ($\rho$) | | | | |
|---|---|---|---|---|---|
| Cost ratio ($p/h$) | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 |
| 2 | 1 | 6 | 11 | 16 | 21 |
| 4 | 2 | 7 | 12 | 17 | 22 |
| 8 | 3 | 8 | 13 | 18 | 23 |
| 16 | 4 | 9 | 14 | 19 | 24 |
| 32 | 5 | 10 | 15 | 20 | 25 |

The simulation experiment for each of the 25 cases included 100 independent replications. The warm up period was set equal to 1,000 demand arrivals. Then, cost data was collected for the following $n$ demand arrivals, where $n$ is the proxy for simulation length and hence the control variable in our simulation experiment.

In order to gather information on the relationship between $n$ and the accuracy of the simulation results, for each of the 25 cases and each of the 100 runs per case, the authors exhaustively tried different values of $R$ in order to determine the optimal base-stock level ($R_s^*$) and compared it with the true optimal base-stock level ($R_a^*$) obtained from the analytical result, to know how many times (out of 100) $R_s^*$ is exactly equal to $R_a^*$ for each value of $n$. Values of $n$ in the range 100 to 10,000,000 were tested with the aim of reaching $R_a^*$ in, at least, 95 replications.

Equation 1.1 is the analytical formula for the optimal base-stock level ($R_a^*$), where $\lfloor x \rfloor$ denotes the greatest integer less than or equal to $x$.

$$R_a^* = \left\lfloor \frac{\ln\left(\dfrac{h}{\rho h + (1-\rho)\lambda p}\right)}{\ln \rho} \right\rfloor \tag{1.1}$$

Table 2 summarizes the following two analytical results for the 25 simulated cases: The optimal base-stock level ($R_a^*$) and its associated expected total cost $E(TC)[h, p, R_a^*]$, where $E(\cdot)$ is the expected value operator. The system decision variable in the simulation experiment is the base-stock level $R$. Let $E(OH)$ and $E(BO)$ respectively be the expected on-hand inventory and the expected number of backorders per unit time (both functions of $R$). Equation 1.2 describes the system's objective function as the sum of the expected inventory holding cost and demand backordering costs.

$$E(TC)[h, p, R] = hE(OH) + pE(BO) \tag{1.2}$$

**Table 2** Analytical values of optimal base-stock level ($R_a^*$) and the expected total cost ($E(TC)[h, p, R_a^*]$) for each simulation case. The expected total cost is in parenthesis

| | Traffic intensity ($\rho$) | | | | |
|---|---|---|---|---|---|
| Cost ratio ($p/h$) | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 |
| 2 | 1 (1.100) | 1 (1.300) | 1 (1.500) | 1 (1.700) | 1 (1.900) |
| 4 | 1 (1.300) | 1 (1.900) | 2 (2.250) | 2 (2.770) | 3 (3.477) |
| 8 | 1 (1.700) | 2 (2.330) | 3 (3.125) | 4 (4.148) | 6 (6.035) |
| 16 | 2 (2.050) | 3 (3.015) | 4 (4.063) | 5 (5.748) | 9 (9.686) |
| 32 | 2 (2.210) | 3 (3.447) | 5 (5.031) | 7 (7.494) | 14 (14.380) |

Next, by means of case 13 ($\rho$ = 0.5; $p/h$ = 8), we describe the procedure to determine the rung length. The same procedure was applied to each one of the 25 cases.

Table 2 shows that when $\rho$ = 0.5 and $p/h$ = 8, $R_a^*$ is 3 units of the item, with an expected total cost of 3.125 monetary units. We set the simulation parameters, including the simulation length ($n$ = 100) and run the simulation (created the queue). We computed the costs for different values of $R$ (i.e. between 0 and 10).

For each run (one hundred independent runs were made), $R_s^*$ was selected as the value of $R$ that minimized the expected total inventory costs (results were between 2 and 5 units). $R_s^*$ matched $R_a^*$ in 63% of the runs. That result did not reach the minimum bound of 95% hits. The experiment was repeated one hundred more times for $n$ = 1,000 periods .The number of hits improved to 93% but it was still unsatisfactory. Finally, after one hundred more replications with $n$ = 10,000 periods, $R_s^*$ matched $R_a^*$ in 100% of the runs. Figure 1 shows that the number of hits ($R_s^* = R_a^*$) increases with the length of the simulation $n$.

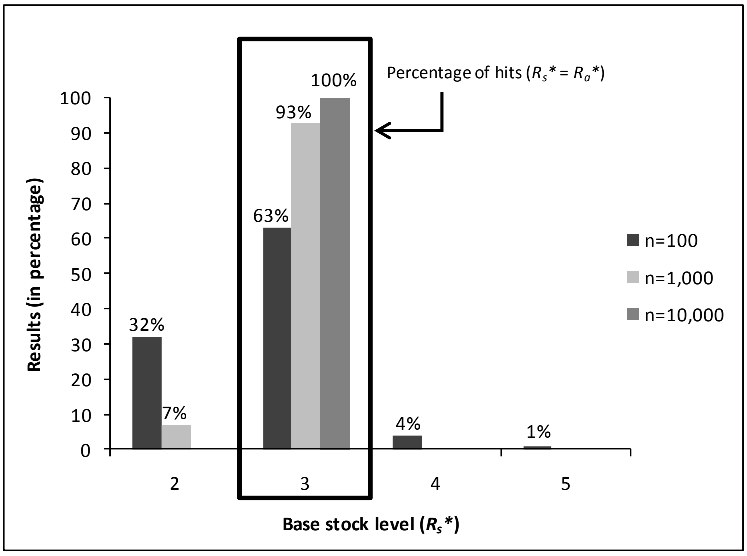


**Fig 1** Summary of the simulation experiment results for case where $\rho$ = 0.5 and $p/h$ = 8

## 4 Results

The simulation experiment required more than 7,500 runs. The percentages of runs in which $Rs^* = Ra^*$ are presented in Table 3.

Note that we increased exponentially the number of periods to be simulated ($n$) in order to reach at least 95% hits ($R_s^* = R_a^*$) in the simulation experiment. We first defined the experiment with $n$ = 100; 1,000 and 10,000 periods. Since the pre-defined values of $n$ were not enough for higher values of traffic intensity ($\rho$), we considered values of $n$ up to 100,000 periods. In most of the cases simulated, we reached a 100% of hits.

We do not include the values of $R_s^*$ nor the values of $E(TC)[h, p, R]$ obtained in the simulation since the aim of the experiment was to determine the minimum number of runs needed to ensure that the simulation of the supply chain has achieved steady state and consequently, that the optimal base-stock level obtained in the simulation ($R_s^*$) is in fact the optimal solution of the inventory system under study ($R_a^*$).

**Table 3** Summary of the simulation experiment results (number of hits: $R_s^* = R_a^*$) for the 25 cases considered

| Case | $\rho$ | $p/h$ | Periods ($n$) | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | $10^2$ | $10^3$ | $10^4$ | $10^5$ | $10^6$ | $10^7$ |
| 1 | 0.1 | 2 | 100 | 100 | 100 | | | |
| 2 | 0.1 | 4 | 100 | 100 | 100 | | | |
| 3 | 0.1 | 8 | 99 | 100 | 100 | | | |
| 4 | 0.1 | 16 | 100 | 100 | 100 | | | |
| 5 | 0.1 | 32 | 100 | 100 | 100 | | | |
| 6 | 0.3 | 2 | 100 | 100 | 100 | | | |
| 7 | 0.3 | 4 | 76 | 100 | 100 | | | |
| 8 | 0.3 | 8 | 99 | 100 | 100 | | | |
| 9 | 0.3 | 16 | 54 | 66 | 81 | 100 | | |
| 10 | 0.3 | 32 | 82 | 99 | 100 | 100 | | |
| 11 | 0.5 | 2 | 99 | 100 | 100 | | | |
| 12 | 0.5 | 4 | 99 | 100 | 100 | | | |
| 13 | 0.5 | 8 | 63 | 93 | 100 | | | |
| 14 | 0.5 | 16 | 41 | 75 | 94 | 100 | 100 | |
| 15 | 0.5 | 32 | 35 | 64 | 69 | 96 | 100 | |
| 16 | 0.7 | 2 | | 100 | 100 | 100 | | |
| 17 | 0.7 | 4 | | 92 | 100 | 100 | | |
| 18 | 0.7 | 8 | | 70 | 99 | 100 | | |
| 19 | 0.7 | 16 | | 66 | 94 | 100 | | |
| 20 | 0.7 | 32 | | 51 | 93 | 100 | | |
| 21 | 0.9 | 2 | | | | 100 | 100 | 100 |
| 22 | 0.9 | 4 | | | | 100 | 100 | 100 |
| 23 | 0.9 | 8 | | | | 60 | 90 | 100 |
| 24 | 0.9 | 16 | | | | 95 | 99 | 99 |
| 25 | 0.9 | 32 | | | | 85 | 100 | 100 |

## 5 Findings, Relevance and Concluding Remarks

This paper fills a void in the literature by providing supply-chain management researchers and practitioners with a practical understanding of the necessary length of a simulation run to achieve reasonably reliable results. We know of cases where researchers had no choice but to perform simulations with more than one billion demands in order to be sure that they had reached the steady state and hence reliable results. Our simulation experiment suggests that in most practical situations, simulating 100,000 demands should be enough, given that Poisson demands and exponential unit manufacturing times represent upper bounds for most supply chains that would be found in the real world.

An analysis of the results in Table 4 reveals the behavior of the system:

- For each of the 25 cases, as expected, the number of hits ($R_s^* = R_a^*$) increases with the length of the simulation $n$.
- For each cost ratio $p/h$, when the traffic intensity ($\rho$) increases, the run length necessary to achieve a 100% of correct results increases too. For example, for a cost of $p/h = 32$, when $\rho = 0.1$, a simulation of 100 demands is enough, and when $\rho = 0.9$, a simulation of 1,000,000 is required. This is because the system requires more time to achieve the steady state due to a higher congestion level. In consequence, the results from a small $n$, do not correspond to the steady state of the system and in consequence they do not correspond to the expected value of $R_a^*$, which is derived for the steady state.
- For each level of traffic intensity ($\rho$), when the backorder cost ($p/h$) increases, the number of hits (the reliability in predicting the optimal value of the base-stock level) decreases for each specific run length. For example, if we take cases 16 to 20 in Table 4 ($\rho = 0.7$) and look at the number of hits for one thousand order arrivals ($n = 1{,}000$), they fall from 100 times (if $p/h = 2$) to 51 times (if $p/h = 32$). Therefore, the number of demands in the simulation must go up to increase reliability.

The main limitation of our research is that it considers only one production-inventory system configuration. It can be extended to other configurations, as long as we have optimal analytical results to use.

## 6 References

Arreola-Risa A (1996) Integrated multi-item production-inventory systems. European Journal of Operational Research 89(2):326-340

Buzacott JA, Shantikumar JG (1993) Stochastic Models of Manufacturing Systems. Prentice-Hall, Englewood Cliffs, New Jersey.

Clark AJ, Scarf H (1960) Optimal policies for multi-echelon inventory problem. Management Science 6(4):475-490

Fishman GS (1971) Estimating sample size in computing simulation experiments. Management Science 18(1):21-38

Fishman GS (1978) Principles of Discrete Event Simulation. John Wiley & Sons, New York: 219-237

Knott K, Sury RJ (1987) A study of work-time distributions of unpaced tasks. IIE Transactions 19:50-55

Liberopoulos G, Koukoumialos S (2005) Trade-offs between base stock levels, numbers of kanbans and planned supply lead times in production/inventory systems with advance demand information. International Journal of Production Economics 96(2):213–232

Snyder LV, Shen ZJM (2006) Supply and Demand Uncertainty in Multi-Echelon Supply Chains. Working Paper. Online. http://www.ieor.berkeley.edu/~shen/papers/paper38.pdf. Cited 12 February 2013

Veatch MH, Wein LM (1994) Optimal control of a two-station tandem production/inventory. Operations Research 42:337-350

Whitt W (1989) Planning Queuing Simulations. Management Science 35(11):1341-1366

Zipkin P (1991) Evaluation of base-stock policies in multiechelon inventory systems with compound-Poisson demands. Naval Research Logistics 38(3):397-412