MULTI ECHELON SPARE PARTS INVENTORY OPTIMISATION: A SIMULATIVE STUDY

Simone Zanoni, Ivan Ferretti and Lucio Zavanella
Dipartimento di Ingegneria Meccanica
Università degli Studi di Brescia
Via Branze 38, 25123 Brescia, Italy
E-mail: zanoni@ing.unibs.it

KEYWORDS
Inventory, Simulation, Spare parts, Multi Echelon.

ABSTRACT

The research in this paper is motivated by a real life spare parts networks for complex technical systems in oil refining sector.
At the customers the availability of the installed technical systems often is essential for the primary process. Hence, they require a high availability. For improve the management of the spare parts inventory the main company (the customer) has developed a new strategies for pushing down to the supply echelon some inventory so as to increase availability of finished parts or components while the properties is still kept by the supplier until its usage, so as sharing in such a way the inventory costs.
In this paper after reviewing the literature on spare parts management the industrial case is discussed and based on it a model with closed queuing network approach will be shown. Due to its analytically difficultness the optimality of the parameters will be addressed via simulative way.

INTRODUCTION

The importance of service parts management has increased in the past decades. One reason is the fact that system availability and high quality after sales service have become important criteria when selecting suppliers of technically advanced systems. A second reason is the increasing value of service part inventory investment. A survey by Cohen et al. (1997) reports that service parts inventories equal 8.75% of the value of product sales in their sample, being over $23 mln. inventory investment on average.
The spare parts are needed to maintain an installed base of technical systems. Examples are aircrafts, locomotives, frigates and computer systems.
Service parts are often supplied via a multi-echelon distribution network, i.e. a hierarchical network of stocking locations through which service parts are supplied to the installed base at customer's sites. A reason to have a multi-echelon structure is the need for both local stocks close to the customer's sites in order to achieve fast supply and the need for stock centralisation to reduce holding costs. Cohen et al. (1997) report that three-echelon networks are prevalent in their sample followed by two-echelon systems. Four-echelon networks occur in practice as well. There is a trend however to reduce the number of echelons and the number of locations per echelon in order to reduce fixed warehousing costs and service parts obsolescence costs.
All these characteristics cause that service parts management is an increasingly important, yet complex task. A key challenge is to attain high availability of the installed base at low service costs. These service costs include costs for stock holding, warehousing, transportation, service engineers, repair shops and overhead.

Literature review

In the literature, various ways to deal with finite capacity in service part networks have been discussed. One of these methods is to model the network as a closed queuing network (Jackson network, cf. Gross et al., 1978, 1983). This method provides very good estimations of the steady state probabilities in a closed network with fixed parameters, but the numerical algorithms involved make it difficult to find optimal stock levels for each location and each part type. Another approach is based on Markov processes; see Albright and Soni (1988), Gupta and Albright (1992) and Albright and Gupta (1993). A drawback of this approach is the fact that the number of states may become very large and that existing methods to reduce the model size to acceptable dimensions are rather rough.
A similar approach is developed by Avsar and Zijm (2000). They construct an excellent approximation for a two-echelon inventory model, where repair shops can be modelled as open Jackson queuing networks. However, their model considers only item-dedicated repair shops and is difficult to extend to multi-echelon model or model with different types of repair shops, as we consider.
Another possibility is to extend the VARI-METRIC method to deal with finite capacity by replacing the $M/G/\infty$ queuing model for the repair shop by some finite capacity system, cf. Aboud (1996), Diaz and Fu (1997), Kim et al. (2000) and Perlman et al. (2001). They use their method to analyse the impact of finite capacity. Diaz and Fu (1997) show that finite capacity has a serious impact on system performance for a single

Proceedings 19th European Conference on Modelling and Simulation
Yuri Merkuryev, Richard Zobel, Eugène Kerckhoffs © ECMS, 2005
ISBN 1-84233-112-4 (Set) / ISBN 1-84233-113-2 (CD)

indenture, two-echelon system with only one central repair shop. They model the repair shop as a GI/G/k multi-class queuing system, where the part flow of one item type is modelled as one class in the queuing system. Although they discuss formulas for multi-server queues, their numerical results refer to single server queues only. In addition, they discuss an alternative method to plug in throughput times as observed in practice in the M/G/ ∞ model, so that waiting times are included. This approach is also used in the case study for the Caracas Metro subway system that Diaz and Fu (1997) present. Then the impact of finite capacity is less, but still significant, and as we mentioned already this procedure is not suitable for what-if analyses.

From the literature review it is quite clear how the more convenient way to model such spare parts inventory situation , considering also finite capacity repair shops, is the queuing network approach. There is an extensive literature on queuing analysis, see e.g. Kleinrock (1975) and Gross and Harris (1998). Applications are particularly in computer and telecommunication system analysis and the analysis of manufacturing systems, cf. Hall (1991), Papadopolous (1993) and Buzacott and Shanthikumar (1993).

PROBLEM DEFINITION

The context from which we have taken inspiration for the development of the present research is real practice observed in Italy. In brief the situation concerns the field of the oil refining, and involve three different companies:

- Co1: industry of oil refining and distribution, situated in the gulf of Cagliari (Sardegna, Italy);
- Co2: supplier of special power plant installations;
- Co3: manufacturer of mechanical components, i.e. as particular pump equipment.

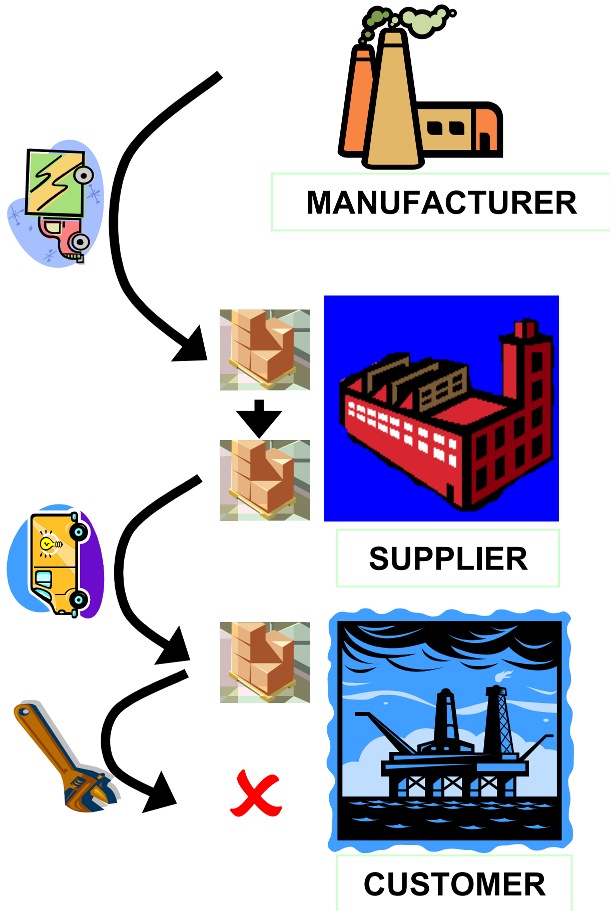
Recently the company Co1 promote some challenging objective, that in some cases involves also its suppliers.

One of this programme is cantered on inventory management optimisation. The company Co1 started to study the possibility of establish a new warehouse for stocking the spare parts that need with certain regularity or with short lead time: the warehouse could be conveniently located in the company Co2, so as it could optimize its production/assembly process located. In the same way company Co2 tried to react with company Co3, establishing there its components warehouse.

This is the problem that we will face in the following part of the paper.

Problem Formulation

A rough representation of the problem described is following reported:



Figures 1: Scheme of the case study

We have reduced the system to three actors: customer (Co1), supplier (Co2) and producer (Co3). The productive structure of the customer is subject to breakdowns that happen at a given failure rate. Such breakdowns are repaired with spare parts supplied from a warehouse managed through a particular agreement with the supplier. The latter guarantees to the customer a given service level. The management of the inventory is entrusted to the supplier who controls the levels of supply and avoids the overcoming of specific limits (defined with contract). The components remains of property of the supplier until when it is used, and paid subsequently. Therefore the cost of maintenance to supply is subdivided in the following way:

- the financial cost is paid by the supplier;
- the physical storage cost is paid by the customer.

Such a situation may be also interpreted as a variant of the Consignment Stock inventory management practice.

It is obvious that this practice may facilitate remarkably the customer, pull down the inventories with less costs. However it is not difficult to understand that this type of agreement assure in such a way the supplier that benefit of an exclusive contract supply. Moreover also the supplier has the component inventory whose level is guaranteed by a similar relationship to that customer-supplier with the manufacturer.

Consigment Stock

The CS (Consignment Stock) policy is an industrial approach to stock management in the Supply Chain, firstly observed in the automotive field. Its principles and modelling are discussed in Abdel-Malek et al. (2002), Valentini and Zavanella (2003) and Braglia and Zavanella (2003), with reference to deterministic demand and to its extension to a stochastic environment. Let's see more in detail the cost structure under the consignment stock policy. The inventory cost per unit h is driven by two main components: a financial one (h_{fin}) and storage one (h_{stock}). Under a typical supply agreement, these costs are borne as indicated in Table 1.

Table 1: Inventory costs in traditional agreements

		Position of Raw Material	
		Supplier	Company
Relevant costs	Supplier	$h_{s,fin} + h_{s,stock}$	0
	Company	0	$h_{c,fin} + h_{c,stock}$

It should be noted that $h_{c,fin} + h_{c,stock}$ is generally greater than $h_{s,fin} + h_{s,stock}$, mainly because of the financial component, which increases as it goes down the supply chain. The different situation brought about by CS is outlined in table 2.

Table 2: Inventory costs under CS policy

		Position of Raw Material	
		Supplier	Company
Relevant costs	Supplier	$h_{s,fin + hs,stock}$	$h_{s,fin}$
	Company	0	$h_{c,stock}$

As can clearly be seen, the main difference is found in the case where the material has already been delivered to the company. In fact, the company incurs the storage cost, given that the material is located in its warehouse, but it does not yet sustain the financial cost. In fact, given that a good is formally purchased only after its consumption, the supplier is still bearing the financial opportunity cost. Thus, while calculating the total cost for the system, we may reasonably assume that the storage component h_{stock} of the total inventory cost may be considered as more or less identical for the supplier and the company, i.e. $h_{c,stock} = h_{s,stock}$. As a consequence, referring to the same average stock level, the total storage cost of the system is lower in the CS case (as we assume $h_{c,fin} > h_{s,fin}$), even if a part of the cost is "shifted" onto the supplier. However, the supplier perceives some advantages as counterpart: the average quantity of the material stored in his own inventories decreases and, consequently, space is available to allocate other items; finally, the supplier may manage his production plan more flexibly, as it is not constrained by closed-orders. On the other hand, the company "sees" a lower inventory cost per unit, that is, only $h_{c,stock}$ instead of the entire ($h_{c,fin} + h_{c,stock}$).

MODEL DESCRITPION

In this section we introduce our model. Firstly, we present the assumptions and notations used in the model. Secondly, we describe the model for the evaluation of a stocking policy and lastly, we formulate the optimization problem.

Assumptions and notations

We model the situation of the three independent companies. In particular we take into account the situation in which both the supplier and the customer keep spare parts on stock for the customer technical systems. We will consider a single type of spare parts type. These systems installed in the customer facility in number N are subject to failures. Failures occur according to Poisson processes with constant rates λ_f. Each time that a failure occur to one of the systems the customer maintenance staff is devoted to completely replace the system with a new one: if the system is available in the customer MRO inventory it could be replaced with a small lead time (λ_m) otherwise it should be ordered to the supplier and it additionally requires a certain transportation lead times (λ_{t1}). Moreover if the system is not available at the supplier finished product inventory an additionally assembly lead time is needed (λ_p). It has to be noticed that with a stochastic probability the system may be repaired, but this operation it has been performed only by the supplier. So as some percentage of the fault system are sent back to the supplier for its supplier (after that they have been replaced with new one) with a certain lead time (λ_{t2}). Repair operation are always cheaper than new system assembly, so when it is possible all the repairable system are repaired.

Parameters name used in the model, as well theirs symbol, are reported in the following table:

Table 3: Relevant parameters of the model

Parameter	Symbol
Failure rate [day]	λ_f
Average repair time [day]	λ_m
Average shipment lead time (supplier-customer) [day]	λ_{t1}
Average supplier assembly time [day/system]	λ_p
Average shipment lead time (manufacturer-supplier) [day]	λ_{t2}
Average remanufacturing time [day/system]	λ_r
Holding cost for systems at the customer stocking point (physical component) [€/system·day]	*hpc*
Holding cost for systems at the customer stocking point (financial component) [€/system·day]	*hfc*
Holding cost for systems at the supplier stocking point (financial and physical component) [€/system·day]	hs_1
Holding cost for components at the supplier stocking point (financial and physical components) [€/system·day]	hs_2
Fixed cost for the maintenance operation [€/repair]	*cfm*
Set-up cost incurred by the supplier [€/Set-up]	*scs*
Cost for plant unavailability [€/hour of downtime]	*cvm*
Transport cost from customer to supplier and from supplier to manufacturer [€/system]	*ct*

Objective function

With all the above defined assumptions and notations, we can now formulate our optimization problem as follows:

Minimize (TC)

With:

$$TC = \sum_{i=1}^{7} C_i$$

Where:

C_1 = average opportunity loss for the downtime of the system (evaluated on the hypothesis that the part of the plant served by the system fell down couldn't work and the customer face the contribution margin of the production lost)

C_2 = average costs of maintenance operations (payed by the customer)

C_3 = average holding costs for the systems stocked in the customer warehouse (in the standard case fully paid by the customer and in the CS case the financial component paid by the supplier and physical one by the customer)

C_4 = average holding costs for the systems stocked in the supplier finished products warehouse (paid by the supplier)

C_5 = average holding costs for the components stocked in the supplier raw material warehouse (paid by the supplier)

C_6 = average transportation costs of the systems from the supplier to the customer (paid by the supplier)

C_7 = average set-up costs for the systems assembly in the supplier facility (paid by the supplier)

The objective is to find a stocking policy under which the average total cost (for the whole supply chain system) is minimized. In such a way two different approach may be pursued:

- traditional supply agreement
- CS supply agreement

Moreover we will take into account the whole supply chain as well as its different components that pertain to the different actors (i.e. supplier and customer).

We will consider for the three different stocking point the same control policy, i.e an (s,S) one (for additional detail Silver et al., 1998). Therefore we optimize 6 parameters:

- (s1, S1) for the customer;
- (s2, S2) for the supplier;
- (s3, S3) for the manufacturer.

SIMULATION STUDY

As discussed in previous paragraph the inventory management of multi-echelon inventory items structures has been studied largely. However, some restrictions still exist on its applications to real case problems. To surpass some of these limitations a simulation model is developed in this work.

In particular in this study, we use Arena to build our simulation model. The embedded OptQuest is applied to search the optimal decision variables. OptQuest includes sampling techniques and advanced error control to find better answers faster, and incorporates algorithms based on Tabu search, scatter search, integer programming, and neural networks (Kelton et al, 2002). Preliminary experiments have been conducted to validate our simulation model as well as to evaluate the solution quality of OptQuest. It shows that the results between analytical and the simulation models are comparable. However, the disadvantage of using OptQuest is that it still takes long time to find the (near) optimal solutions.

To create uncertainty we use a fixed seed stream in Arena. Each source of uncertainty in this model has its own seed for generating random numbers.

More precisely, in our simulation experiments, we model as stochastic processes the failure process, the production/assembly process and lead times and

In this way the different models are comparable with each other and possible differences due to the use of different random seeds can be excluded.

Simulation results

With the real data provided by the company (here reported slightly changed due to the confidentiality required) it has been possible to perform a complete simulative optimisation for all the parameters and for each of that perform a sensitive analysis. In particular we have taken into account the failure rate (λ_f) as the main parameters for investigate the response of the system.

We will consider the plant of the company consisting of 50 equal pumping groups variously located in the refinery plant. Detailed data for each parameter is reported in Table 4.

Table 4: Parameters value used for the case study

Symbol	Traditional	CS
λ_f	POIS(15÷120)	POIS(15÷120)
λ_m	UNIF(2;4)	UNIF(2;4)
λ_{t1}	NORM(4,3)	NORM(4,3)
λ_p	NORM(2,1)	NORM(2,1)
λ_{t2}	NORM(5,4)	NORM(5,4)
λ_r	UNIF(10;5)	UNIF(10;5)
hpc	9	5
hfc		2
hs_1	6	6
hs_2	4	4
cfm	2	2
scs	200	200
cvm	45	45
ct	3	3

In the following figures it is possible to compare results of the optimised configuration for the two cases under analysis, while varying the failure rate. The simulation length as been set to 3650 period, i.e. considering a planning period of 10 years.

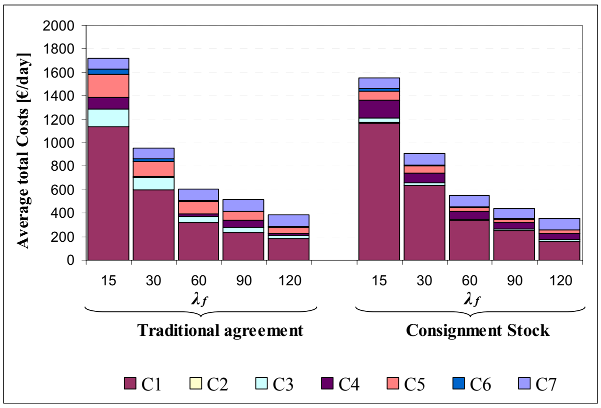


Figure 2: Average costs distribution while varying the failure rate within the two agreement option

With the CS agreement the average total cost is always lower than in the traditional agreement case. In the following table it is reported the percentage improvement in the objective function.

Table 5: Average cost improvement associated with the CS policy adoption

λ_f	15	30	60	90	120
Average cost improvement	-9.5%	-4.4%	-8.4%	-15.4%	-7.2%

It is also interesting to investigate how the system responds in terms of service level, measured as the effective available time of the plant on the requested time of its availability, while varying the failure rate with the two different policy.

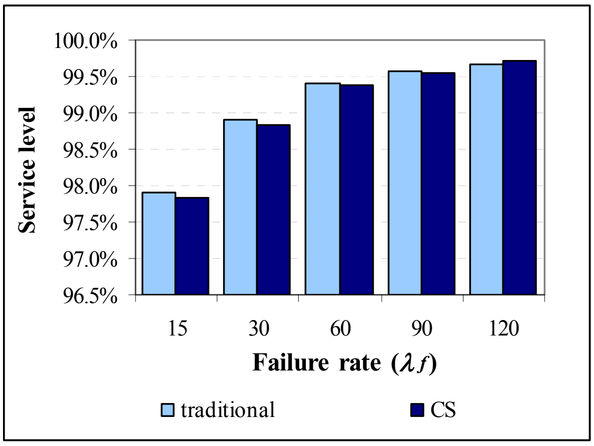


Figure 3: Service level while varying the failure rate

Moreover with the simulation optimisation performed it has been possible to design a trade-off curve between the service level and supply chain costs: curve is reported in figure 4.

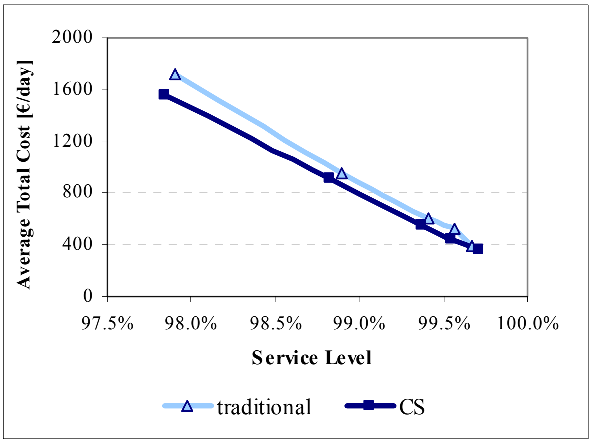


Figure 4: Trade-off curve between average total costs and service level

Also in this case the CS policy agreement perform better than the traditional one.

CONCLUSIONS

Various process and manufacturing plants using complex machines often require large quantities of spare parts to guarantee high system availability which in turn results in excessive holding cost.

On the one hand, companies can find themselves carrying an excessive number of spare parts. On the other hand, if they were not available when needed, companies will face severe downtime consequences. As many parts are very expensive, critically important and their failure rates are so low that they are difficult to forecast, spare parts inventory management within these industries is one of the hardest problems to deal with.

In this paper we have presented a simulative approach for planning the spare parts inventory in a refinery plant, where a large amount of components have a failure rate that significantly affect the availability of the system. Moreover we have also shown how under a particular supply agreement, named Consignment Stock, the whole system perform better, both in terms of total costs and service level with a given failure rate.

The same approach and methodology may be applied to other supply-chain cases, thus offering the proper information to undertake the best course of action in the multi echelon spare parts inventory management.

REFERENCES

Abdel-Malek; L., Valentini G. and L.E.; Zavanella, 2002, Managing stocks in Supply Chains: modelling and issues, in Seuring S.; Goldbach M., Cost Management in Supply Chains, Heildeberg: Physica Verlag, pp.325-335.

Aboud, N.E., 1996, "The Markovian two-echelon repairable item provisioning problem", Journal of the Operational Research Society 47, 284-296.

Albright, S.C. and A. Gupta, 1993, “Steady-state approximation of a multi-echelon multiindentured

repairable-item inventory system with a single repair facility", Naval Research Logistics 40(4), 479-493.
Albright, S.C. and A. Soni, 1988, "Markovian multi-echelon repairable inventory system", Naval Research Logistics Quarterly 35, 49-61.
Avsar Z.M. and W.H.M. Zijm, 2000, "Resource-constrained two-echelon inventory models for repairable item systems", working paper, University of Twente, Enschede, The Netherlands.
Axsäter, S. 1990. "Modelling emergency lateral transhipments in inventory systems". Management Science, Vol. 36, 1329-1338.
Braglia, M. and L:E. Zavanella, 2003, An industrial strategy for stock management in Supply Chains: modelling and performance evaluation, International Journal of Production research, vol. 41, no. 16, 3793–3808.
Buzacott, J.A. and J.G. Shanthikumar, 1993, Stochastic models of manufacturing systems, Prentice Hall, Englewood Cliffs, NJ
Cohen, M.A., Y-S. Zheng and V. Agrawal, 1997, "Service parts logistics: a benchmark analysis", IIE Transactions 29, 627-639.
Diaz, A. and M.C. Fu, 1997, "Models for multi-echelon repairable item inventory systems with limited repair capacity", European Journal of Operational Research 97, 480-492.
Gross, D. and C.M. Harris, 1998, Fundamentals of queuing theory, Wiley, New York
Gross, D. and J.F. Ince, 1978, "Spares provisioning for repairable items: Cyclic queues in light traffic", AIIE Transactions 10(3), 1978, 307-314.
Gross, D., D.R. Miller and R.M. Soland, 1983, "A closed queueing network model for multi-echelon repairable item provisioning", IIE Transactions 15(4), 344-352.
Gupta, A. and S.C. Albright, 1992, "Steady-state approximations for multi-echelon multi indentured repairable-item inventory system", European Journal of Operational Research 97(3), 340-353.
Hall, R.W., 1991, Queueing methods for services and manufacturing, Prentice Hall, Englewood Cliffs, NJ
Kelton, W. D., Sadowski, R. P. and Sadowski, D. A. (2002), Simulation with Arena, McGraw-Hill, Boston.
Kim, J-S, K-C Shin and S-K Park, 2000, "An optimal algorithm for repairable-item inventory systems with depot spares", Journal of the Operational Research Society 51, 350-357.
Kleinrock, L., 1975, Queuing systems, Wiley, New York
Papadopolous H.T., C. Heavey and J. Browne, 1993, Queuing theory in manufacturing systems analysis and design, Chapman & Hall, London
Perlman, Y., A. Mehrez and M. Kaspi, 2001, "Setting expediting repair policy in a multiechelon repairable-item inventory system with limited repair capacity", Journal of the Operational Research Society 52, 198-209.
Silver, E.A., Pyke, D.F. and R. Peterson, 1998, *Inventory Management and Production Planning and Scheduling 3rd Ed.*, John Wiley & Sons, New York.
Valentini, G. and L.E. Zavanella, 2003, The Consignment Stock of Inventories: Industrial Case and Performance Analysis, International Journal of Productions Economics, Vol. 81-82, pp.215-224.

AUTHORS BIOGRAPHIES

SIMONE ZANONI, Ph.D is currently an assistant professor in the Mechanical Engineering Department at the University of Brescia (Italy). He graduated (with distinction) in 2001 in Mechanical Engineering and took his a Ph.D at the University of Brescia (Italy) in 2005. His main research interests are focused on Layout Design, Inventory Management and Closed-loop Supply Chain. His e-mail address is : zanoni@ing.unibs.it.

IVAN FERRETTI, graduated in 2003 in Industrial Engineering at the University of Brescia (Italy). Since 2004, he is a Ph.D student at the same university. His main research interests are focused on Simulation, e-business, Supply Chain Management and Reverse Logistics. His e-mail address is : ivan.ferretti@ing.unibs.it.

LUCIO ZAVANELLA graduated in 1982 in mechanical engineering at the Politecnico di Milano. Since 1986, he has been a researcher in the field of production systems at the Università degli Studi di Brescia. Presently, he is associate professor of Industrial Plants. His research activities mainly concern manufacturing systems, also with reference to environmental and safety aspects. His e-mail address is: lucio.zavanella@ing.unibs.it.