ABSTRACT CODE: 003-0074

A MULTI-ECHELON SYSTEMS' SIMULATION MODEL FOR REPAIRABLE AND CONSUMABLE ITEMS MANAGEMENT: A CASE STUDY

M. Elisa Cunha
Dep. Ciências e Tecnologias
Univ. Autónoma de Lisboa
R. Santa Marta, 56
1169-023 Lisboa, Portugal
E-mail: melisa@ual.pt

A. Paula Barbosa Póvoa
Centro de Estudos de Gestão
CEG-IST
Av. Rovisco Pais
1049-001 Lisboa, Portugal
E-mail: apovoa@ist.utl.pt

A. Assis Lopes
Dep. Economia e Gestão
Univ. Lusíada de Lisboa
R. da Junqueira, 188
1349-001 Lisboa, Portugal
E-mail: alopes@lis.ulusiada.pt

Sixteenth Annual Conference of POMS, Chicago, IL, April 29 – May 2, 2005

KEYWORDS

Inventory management, repairable items, multi-echelon systems, simulation, logistics, stochastic models, decision aid systems.

ABSTRACT

The inventory management of multi-echelon repairable 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. The model not only admits a multi-echelon structure system but also permits the explicit consideration of a hierarchical parts structure for repairable and consumable items.

The model generality is proven trough its application to the stock management of the 1900 Locomotives components. Repair may be required due to components breakdowns or potential limit reached. Depending on the locomotives failure repair can be performed at different locations within the multi-echelon structure considered. Repair time depends on the components availability as well on associated lead times.

As final result an adequate inventory policy is obtained. This guarantees a certain level of service accounting for operations occurrence and repairs conditions.

1. INTRODUCTION

Multi-echelon structures are used in a large number of environments. For example, each production system stage can be considered an echelon; a supply chain can be divided into several echelons such as warehouses and retailers; and distinct repairable centres from a maintenance system can be members of the same or different echelons.

Items flows in these structures are categorized in acyclic or cyclic. In the former (e.g. production and distribution systems) items demand are pulled from the lower echelons and on the later (e.g. maintenance systems) the items flow in any direction along the structure and, despite the interdiction for leaving the system, the items can suffer transformations such as damage or repair. These are the case of repairable items.

The fundamental purpose of a replenishment control system is to answer questions such as “how often the inventory should be determined”, “when an order should be placed” or “how large the demand should be”. If the inventory control is defined for a multi-echelon system the responses should be given for all system’s entities and can result from a decentralized or a centralized approach. In the first method, each entity determines the inventory control parameters weighing out its owned targets and restrictions. In the second procedure all parameters are calculated at the same time complying with the interaction and the hierarchy between system’s entities. Though, in practice, most multi-echelon inventory systems are managed using adaptations of single location methods Muchstadt & Thomas (1980) proved that such methods have performances inferior to methods designed to take advantage of the system structure.

When the material to be managed involves repair, its control is must more complex due to different reasons. Firstly, the items usually belong to multi-echelon structure systems typically with two levels, where on the lower level there are one or more bases each of them are supplied by one or more depots fixed at an upper level. Secondly, it is indispensable to define some probability distribution: one to identify the repairable time, depending on the failures and resources at the repair facility, other to establish the interval between the order’s base to the depot and the instant of time when the item is available at the base, if the depot has on-hand stock, if not, a last probability distribution that characterizes the elapse time to an item is available at the depot. Thirdly, items failure may not be fitted by the recovered items returning from the repair facilities and lastly, the multi-indenture repairable items structure may require a simultaneous optimisation for all components. In addition, to all these aspects the word supply along the structure correspond to different situations: a report a failure item being sent to the repair facility; an immediately upper echelon demand for a reusable item or else an external supplier new item's purchase.

Several multi-echelon inventory models for repairable items have been developed (Guide & Srivastava, 1997). Most of them consider a one-for-one replenishment policy either for ordering's mended parts every time a failure occurs or for damage items being sent to repair. This strategy considerable simplifies the models, regularly very complicated due to the system complexity where it is often the case that the repairable items tend to have a unit size EOQ. Such items are usually characterized by a higher per-unit price, a near to the ground demand, a wider life cycle and a lower repair cost than the acquisition cost.

The first reported model developed to calculate on a system the optimal stock level at each base and at the supporting depot for every first-indenture item was the METRIC (Sherbrooke, 1968). It was applied in a military environment and is sustained by the following key assumptions: (1) items demand is a Poisson process, (2) the decision regarding the repair place depends exclusively on failure complexity, (3) lateral supply is not allowed, (4) unlimited repair facility, (5) the system is conservative, i.e., the condemnation or scrap rate is quite low, (6) recoverable items are the same essentialities, (7) no cannibalisation is allowed. The model's objective is to minimize the expected number of base backorders for a group of recoverable items that is equivalent to maximize the system's availability, if there is no cannibalisation, as Sherbrooke has proved. Over the years, some of the key assumptions imposed on this method have been relaxed. Simon (1971), motivated by the American Air Force supply structure, assumed that an $(s_j - 1, s_j)$ replenishment policy is used at the base j, $j = 1, \ldots, J$, and that an (s_0, S) policy is used at the depot to derive exact expressions for the stationary distributions of stock on hand, stock in repair, and backlogged demand at each facility. He also indicated how these expressions might be utilized for optimisation purposes. Muchstadt (1973) proposed a mathematical model, named MOD-METRIC, to describe the logistics relationship between the components and a particular final assembly and to calculate the base and depot spare stock levels for all items with explicit consideration of this logistics relationship.

Later on and considering an item that is subject to two types of failures, recoverable and non-recoverable failures, Richard (1976) suggested a model with random demand and random lead times. For a class of simple procurement policies an exact expression for the stationary distribution of the on-hand inventory is derived. Gross et al (1983) relaxed the infinite source and ample service constraints and proposed a closed queueing network theory to represent the stochastic demand process within a two-echelon repairable item provisioning system. An implicit enumeration algorithm is used to solve the resulting optimisation problem. Slay (1984) derived an improvement to the METRIC model, called VARI-METRIC, where the distribution to derive the inventory position at the sites was characterized by the two first central moments of the demand variable. Sherbrooke (1986) based on the real fact that the models tend to understate expected backorders and overstate expected availability of repair items combined multi-indenture and multi-echelon systems applying the VARI-METRIC model and produced estimates of backorders that are very close to the values obtained by simulation. Lee (1987) grouping the bases into *n* disjoint sets such that members of each group are identical, developed a model that allows emergency lateral transhipment between bases belong to the same group. Three prioritised source rules were considered to determine the emergency lateral supply base. Grupta & Albright (1992) provided an approximated method to estimate the steady-state distribution of a limited number of working machines, with a multi-indenture structure, which fail because of modules failures at random time. Diáz & Fu (1997) introduced approximations that can deal with limited repair facilities, under the scenarios of single-class exponentially distributed repair distributions, single-class general repair distribution and multi-class general repair distributions. In addition they demonstrated how these models significantly outperform traditional models in the case of high repair facility utilization. More recently and motivated by an industry case study, Wang et al (2000) relaxed the METRIC assumption related to the independency and the identically distribution of the replenishment lead times.

In spite of the sophisticated models and techniques developed to manage the spares parts inventory in a multi-echelon system some restrictions still exist on its application to real case problems. In fact, all the models have a complex analytical structure and consequently are difficult to understand and to operate and of complicated or even impossible resolution. As stated by Rustenburg et al (2001) the existing type of models, in particular the ones based on the METRIC model, suffer from a series of limiting assumptions. The authors proposed an agenda of points to be further studied. For instance a small population of working items, limited repair facilities, the coexistence of both repairable and consumable items or the occurrence of items' consumption and condemnation.

To surpass some of the limitations previously described we develop in this paper a simulation model to manage both the repairable and consumable items in a multi-echelon structure system where a repairable items' structure is considered explicitly (Cunha et al, 2003). The model aim is to establish the relationship between the average number of operating final items and its components' inventory position. The approach is general and can be adapted to any system, besides it complexity or the items involved or even decision variables characteristics. Moreover, simplifying and disarranging conditions, deviated from the reality, are not imposed, and exact conditions are allowed to use.

The rest of the paper is organized as follows. The system to model and the assumptions considered are described in section 2. The methodology and the model are presented in section 3. In section 4 is illustrated the case study. Finally, section 5 provides the conclusions and the main points for further research.

2. THE SYSTEM TO MODEL AND ITS ASSUMPTIONS

A three-echelon system with the structure depicted in Figure 1 is considered. Each level has a repair and a storage facility. The former has a limited repair capacity imposed by the number of technicians and repair lines. The later stores all components requested by the reparations performed at the echelon or required to satisfy immediate lower echelon demand.

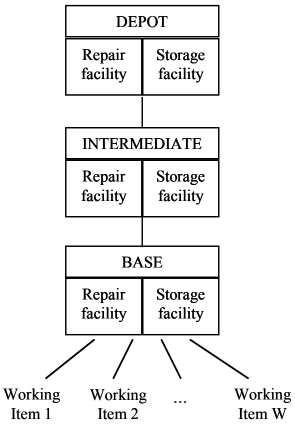

Figure 1: System to model structure

A reduce number of working items are associated with the base and the items hierarchical parts structure can be expressed explicitly as represented in Figure 2. The final item is formed by a set of *N* major assemblies, in the first level they are represented by FLA $n,\ n = 1, \ldots, N$. These on the other hand may consist of several subassemblies, forming the second level assemblies denoted by SLA *nh*, where the first index refer to the father's index and the second one refers to the subassembly number. At this second level again each subassembly can be composed by different third level components, the C *nht* components, which index also inherits the first number from the father, the subassembly *nh*. Subassemblies or components can be accessed without dismounting the item or the assembly, if that describes the reality.

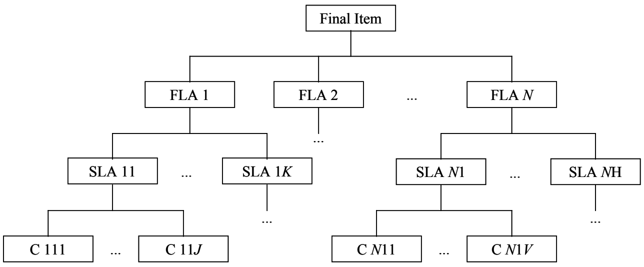

Figure 2: Working Final Item structure

The system behaviour is explained in Figure 3. The numbers between brackets helps to understand the operations performed and the oriented segments represent final items or its components flow, a dashed segment corresponds to a damaged part path while a continuous one denotes a repaired part path.

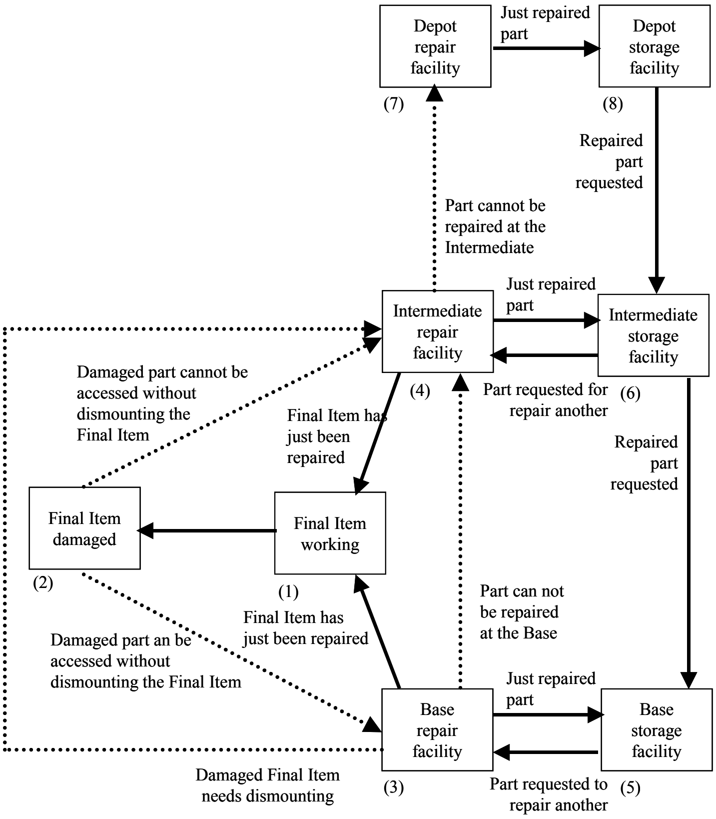

Figure 3: System behaviour; items, assemblies, subassemblies and components movements

A working item (1) is subject to failures caused by consumable or repairable components malfunction. Depending on the failure (2), repair can be done at the base (3) or at the intermediate echelon (4). Transportation time can be equal, different, constant or variable. When arriving at the base (3), an item waits in a FIFO queue until its failures check can be started. As a result a decision on the location for the repair is taken: the base or the intermediate echelon. If at the base (3) the malfunction components are dismounted and dispatched to the repair facility. A request is made for a repairable or new component to the echelon storage facility (5). If at the intermediate echelon (4), the reparation takes a long repair time since it implies item disassembly. Therefore, all first level item assemblies are replaced by repairable ones, independently of being or not damaged. Repairable assemblies are requested from the echelon storage facility (6).

Items repair time depends on the components or assemblies stock on hand or, if no stock is available, it depends on the elapse time for the needed components or assemblies to be at the echelon storage facility. In this later case, they can arrive either from the echelon components or assemblies repair facility or from the immediate upper echelon storage facility.

Depending on the failure, the components and the assemblies can be repaired at the base (3), at the intermediate echelon (4) or at the depot (7). Like items, they wait at the repair centre in a FIFO queue until technicians are released to check the damages and decide if the reparation can proceed at the echelon or if it must take place at the immediate upper echelon. In the later alternative a mended component or assembly is delivered from the upper echelon storage facility, when available.

Components and assemblies repair time depends, also, on the lower-indenture parts stock at the echelon where the repair is performed or, if there are none, depends on the elapse time for all needed parts to be available. Again these parts can arrive from the echelon components or assemblies repair facility, or from the immediate upper echelon storage facility.

When the reparation is finished the item is sent to work (1) and the component or the assembly is sent to the storage repair facility of the echelon where the repair has been done or, if there are backorders, one is satisfied.

Whenever the damage does not justify the repair, items or components must be discarded. This situation imposes an obligation to define an acquisition policy for every parts required. Based on a hierarchical information system, the decision is taken by the depot.

As referred before the model's objective consists on establishing a relationship between the stock inventory position, for the components and assemblies, and the system's operability, estimated from the number of items that have non-work conditions and are at the repair centres.

The following conditions are assumed for the model:

- Failures are independent and identically distributed;
- Failures can follow an empirical or a theoretical distribution, such as Binomial, Negative Binomial or Poisson distribution;
- Damaged final items, assemblies and components transportation time are independent and, if between the same sites, identically distributed;
- Components repair time are independent and identically distributed, if occur at the same location;
- Final items and assemblies repair time depends on the components stock on hand;
- Demand is satisfied on a "first come first served" basis;
- Transportation time for new or repairable components are included in the assembly or item repair time;

- Transportation or repair times can follow an empirical or a theoretical distribution such as: Beta, Erlang, Exponential, Gamma, Geometric HyperGeometric, Uniform, Normal, LogNormal, Triangular, Weibull distribution;
- A $(S - n, S)$ policy, where n represents the number of components dismounted, is used either on demand, of repairable or new components or assemblies, or on the components or assemblies sent to repair;
- Acquisition policies are (s, Q), for consumable parts, and $(S - 1, S)$, for reparable ones.

The model built to simulate the system described. Its functionality is explained in the next section.

3. THE METHODOLOGY AND THE MODEL

To establish the relationship between the parts stock inventory position and the system operability, we suggest a simulation model (Cunha et al, 2004; Cunha, 2003).

The model was built using Extend^{TM} (1998) package that has inherent an object-oriented simulation, with predefined object frames that describe various simulation requirements. Those objects are called blocks and are composed of an icon, connectors and a programme. The icon is the graphical representation. The connectors are represented as small squares localized around the icon and through which block are linked and the information flows. There are two types of information involved; one concerning the moving objects along the simulation model that can have associated attributes and priorities, the other about values that can change during the simulation run. The programme defines how the information is used and transformed by the block. Also, there are two blocks' types, labelled simple and hierarchical blocks. The formers have associated a dialog box and a programme and specify a system's action or procedure. The hierarchical blocks refer to a logical set of sequential operations carried out in the system and are compounded of others simple or hierarchical blocks.

The model developed is a discrete simulation model and the simulation time increment is based on the following events: a working item's failure; an damaged item beginning or ending repair at the base or at the intermediate echelon; a damaged component or assembly starting or concluding reparation at any echelon; an item, component or assembly arriving at the repair centre.

The model structure is identical to the modelled system's organization as shown in Figure 4. Six main blocks are considered: the "Executive" block, the "External Supplier" block, the "Depot" block, the "Intermediate" block, the "Base" block and the "Working Items" block. Items, assemblies or components' movements are allowed from different blocks and are represented in Figure 4 by oriented segments.

The "Executive" block is a compulsive block in all Extend™ discrete simulation model, being responsive for the time increment that is processed in the variable time slice depending on the events' occurrence.

The "External Supplier" block represents the system external suppliers to whom the depot acquires new items, assemblies or components. As it is an outside system's entity, its tasks are not considered and its presence in Figure 4 is justified since it represents a link between the system and its external environment.

The "Working Items" block symbolizes items in their activity at the working place. It is a hierarchical block so it is composed by other blocks, that are related with the actions involved in the working items. The item can breakdown due to wearing and this possibility is checked based on a failures' items probability distribution function. The identification of the item damaged assembly or component is also considered. Depending on the failed part, the item's repair is performed at the base or at the intermediate echelon. On both cases the transport is generated from a probability distribution function.

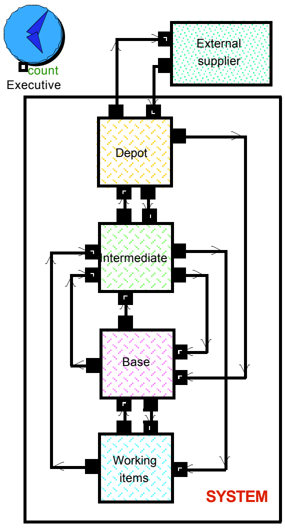

Figure 4: Simulation model structure

The "Base" block in Figure 4 is also a hierarchical block and can be decomposed in a set of blocks representing the base tasks that can be summarised the following way. When arriving at the base the damaged item waits in a FIFO queue until necessary conditions are verified, such as an empty repair line and available technicians. At that time the item enters the repair line. Firstly, failures are checked and processing time is generated by a probability distribution function. Depending on the malfunction element the item is sent to the appropriate place. If possible, reparation proceeds at the base, if not the item is sent to the intermediate echelon. If the item remains in the base then the damaged item assemblies or components are retired from the item and sent to the parts repair place. Simultaneously an order, for all required parts, is posed to the echelon storage facility. If there is stock the parts are sent immediately; if not the demand is backordered until there is an available part. This may arrive from the base repair facility or from the intermediate echelon storage facility. When the repair is finished the item is sent to the working place. Consequently the technicians involved in the repair become available and the repair line becomes free.

The tasks related with assemblies and components reparation are similar to the item reparation. When the assembly or component arrives at the base, they wait in a FIFO queue until technicians and a part repair line are available. Failures are checked and the repair place is decided, depending on the failure's gravity. Check time is generated from a probability distribution. If the failure cannot be repaired at the base the part is sent to the intermediate echelon. If the assembly or the component can be repaired at the base, spare parts are ordered to the base storage facility. If all spares are in stock the storage facility sends them immediately, otherwise the demand is backordered until a spare arrives from the base repair facility or from the intermediate echelon storage facility. When the repair is finished all assemblies or components are sent to the echelon storage facility, or if the are backorders one is satisfied.

The item or parts transportation time to the intermediate echelon is generated from a probability distribution function.

The "Intermediate" hierarchical block, from Figure 4, models the intermediate echelon and the associated tasks that can be explained like this. When an item arrives at the intermediate echelon from the working place or from the base it waits in a FIFO queue until a repair line becomes free and technicians are available. When the conditions are validated the item goes into the repair line. To begin, the item is dismounted to get off the first level assemblies; dismounted time is generated from a probability distribution function. Behind, two simultaneously tasks occur: (1) an order for the first level assemblies is placed to the echelon storage facility; (2) assemblies retired are sent to the assemblies and components repair facility. As a consequence of task (1) and if the echelon storage facility has stock on hand, demand is immediately satisfied; if not the demand is backordered until there is an available assembly arriving from the echelon repair facility or from the depot storage facility. When arriving, the required assemblies are mounted in the item. Once more a probability distribution function is used to generate the mounted time. After repair, the item is sent to the work place, the repair line becomes free and technicians are enabled to start another reparation.

Parts reparation at the intermediate echelon are performed similarly as described for the item repair.

Finally, the "Depot" block represents the tasks executed at the depot. When a damaged part comes in, it also waits in a FIFO queue until necessary conditions are verified. At that moment technicians disband the part, if it is needed, and check the failure extension to decide if the components can or cannot be repaired. Disband time and check time are generated from probability distribution functions. If the response is positive the part is sent to another FIFO queue until the reparation can start. When reparation starts all constituents are dismounted and mended, in case they are recoverable. The time for doing the jobs is generated from probability distribution functions. After that, an order is posed to the echelon storage facility on the spares required to end the reparation. If there are spars in stock, demand is satisfied, if not it is backordered until there is a spare accessible arriving from the repair centre or from an external supplier, if a replenishment demand had been done. When the repair is finished all components are sent to the storage facility, the technicians become free and the repair line stays in a vacant stage.

Whenever the damage cannot be repaired the element must be sent to the scrap. Simultaneously, it must be verified if a new replenishment is required. This verification is based on a hierarchical information "wire" system where the element's code is the key to provide all indispensable information. Firstly, one unit is subtracted from the component stock inventory position to the whole system. Secondly, if the component is repairable an unit size replenishment is ordered to the external supplier; otherwise, a Q unit size replenishment is ordered if the inventory position has dropt to the reorder point or below that. Whichever situation occurs, the data is written in a file that must be read from an external software to place the order automatically.

The above system’s structure, item’s structure or even the capacity of the repair centres are general. To adapt the model to another configuration it is sufficient to copy and paste the suitable model’s blocks or to change the appropriated block’s parameters.

To prove the generality of the model it was applied to analyse the stock management problem of the 1900 Locomotives components described in the next section.

4. THE CASE STUDY

The management problem of the 1900 Locomotives components is studied.

The inoperative time of the 1900 Locomotives due to components breakdowns or potential limit reached is very widely as a result of some repairable components run out of stock. This fact increases greatly the number of locomotives waiting in the repair centre and, consequently, the system operability decreases. So the Railway Executives desired to understand the situation and to establish a relationship between the operational 1900 Locomotives’ average number and the stock inventory position for the Turbo Compressor, the Diesel Engine, the Traction Engine and the Wheel System, the mainly responsible components for the delay verifies in the locomotive’s reparation.

To achieve the target we have adjusted the model, described in the previous section, to fit the case study characteristics that are expressed below.

4.1. Description

The 1900 Locomotives' operation and maintenance system has a three-echelon structure as depicted in Figure 5. In the lower level there are two Maintenance Regions, one identified by Centre (CMR) other named South (SMR), where thirty locomotives are affected; in the middle level there is the Barreiro Repair Shop (BRS) and at the top there is the Entrocamento Repair Shop (ERS). Depending on the failures, the two inferior levels can repair locomotives and the two superior levels are able to recover locomotives components. Each level stores exclusively the parts required to the repair executed.

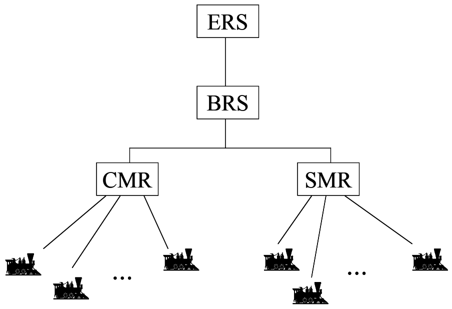

Figure 5: Case study system's structure

The operational locomotives may need repair due to equipment random breakdowns or due to potential limit.

In the case study four gears and respective components are considered: the Turbo Compressor (TC), the Diesel Engine (DE), the Traction Engine (TE) and the Wheels System (WS).

When the potential limit of at least one of the items is reached the locomotive goes to the BRS to do a V1 or a R preventive maintenance. The former (V1) is chosen when the item that has reached the potential limit is the TC, the TE or the WS and involves the simultaneous replacement of the three items mentioned before. The later (R) is selected when DE reaches the potential limit and involves the DE substitution in addition to the TC, the TE and the WS changing.

In the case of equipment breakdowns, if the malfunction is caused either by the TC or its components the broken TC must be dismounted and replaced by a recoverable one. Depending on the potential of the non-failure components the locomotive is sent to the Maintenance Region, to which it is affected, or to the BRS, if any of the non-failure components has reached at least 75% of the potential limit. In this case the V1 or the R maintenance reparation is anticipated. If the locomotive stop is originated in the DE, the TE, the WS or their components breakdowns the locomotive is sent to the BRS to replace the damaged element by a recovered one. A V1 or a R can be forestalled if any of the non-failure components have reached 75%, or more, of the potential limit.

When arriving at the CMR, at the SMR or at the BRS the locomotive waits in a FIFO queue until necessary conditions are verified.

If the reparation occurs at one of the Maintenance Regions (CMR or SMR) an order for a reparable TC is placed to the BRS. If there is stock on hand the item is sent immediately, if not the demand is backordered until a TC is available. Meanwhile the broken TC is taken off from the locomotive and sent to the BRS repair facility. When the TC arrives the locomotive is repaired and sent to operation.

If the repair is done at the BRS the locomotive is washed and dismounted. After that, according to the procedures required, the specific items are taken off and an order is place to the echelon storage facility for identical elements but in recoverable condition. If all spares are in stock they are sent immediately or else the demand is backordered while the required spares are not accessible. When the repair is finished the locomotive is sent to operation.

Whilst the BRS has two repair lines, either the CMR or the SMR has only one affected to the 1900 Locomotives.

The TC, DE and TE, with lower level of damage gravity, are recovered at the BRS; the highly damage TE and the WS are repaired at the ERS. Because of items' different characteristics, each of the echelon has one section for each item type to recover and every section can repair two items at the same time. When arriving at the repair facility the item waits in a FIFO queue until necessary conditions are satisfied. Then the item is disassembled; if the damaged justifies the repair, the repairable components are repaired, if not the components are discarded as well as the consumable components.

The locomotives repair time depends on the components' stock or, if there is none, on the repair time of the missing items. The repair time of the fifty components considered is generated from different probability distribution functions.

4.2. Data treatment

The model described previously was adapted to fit the system objective and an input data models was developed bearing in mind the four steps suggested by Banks et al (2001). First of all, data related to (1) the locomotive failures, (2) the locomotive failures caused by the TC, the DE, the TE or the WS, (3) the TC, the DE, the TE, the WS and their components repair time was colleted, resumed and validated; information concerning to the reparable or consumable replenishment time was not accessible. At this stage, two graphical techniques were used, the scatter diagram and the correlation plot, for informally assessing the data independence. We have concluded that the TE and the WS repair time were not independent because of the transportation process done in groups of three or six, respectively. The second step consisted in identifying the probability distribution families to represent the independent input processes. The families' hypothesises were based in the data characteristics as summary statistics, histograms and line graphs. In third step, the data was used to estimate the maximum-likelihood estimators for the distribution families' parameters. Lastly, it was determined how representative the fitted distributions were. At this phase the histograms of the empirical and the theoretical distribution were compared and goodness-of-fit tests, such as the χ^2 test and the Kolmogorov-Smirnov test, were employed. We have concluded that the probability distribution functions to be used are the following: the TC, the DE, the TE or the locomotives' failures are generated from Poisson distributions with a different parameter, the TC and the DE repair time are generated from a LogNormal and from an Exponencial distribution, the components repair time is generated from empirical distribution.

The dependent data was studied appealing to the Box Jenkins general methodology for developing an appropriate ARIMA time series model. Their approach consists of a three-step iterative procedure. To begin a tentative model of the ARIMA class was identified through analysis of the TE and the WS repair time collected data. In particular using the autocorrelation function and the partial autocorrelation function. Then the unknown parameters of the model were estimated by means of the least-squares estimates. Finally, diagnostic checks were performed to determine the adequacy of the model. If the fitted model was adequate it should transform the observation to a white noise process so the residuals samples were calculated. Their autocorrelation functions, which values did not differ significantly from zero for all lags greater than one, lead to the conclusion that the repair time of the lower level damage TE is generated from an ARMA model and the repair time of the highly damage TE and of the WS is generated from two different parameters AR models.

All statistical analysis were done within the SPSS (2001) software.

The information about the input data models developed was introduced in the simulation model. Adversely, acquisition policies for none of the components could be incorporated because of the lack of information concerning the replenishment time.

4.3. Simulation's run length

Within this case study we are interested in estimate a steady-state or long-run system's property that is not influenced by the initials conditions of the model at the beginning of the simulation run.

Based on this the simulation run length was chosen with three conditions in mind: the estimator's bias to arbitrary initial conditions, the desired precision of the point estimator as a measurement of the estimator variability and the constraints on computer resources.

Above all estimator's bias due to arbitrary initial conditions can be severe if the run length is too short but it usually decreases as run length increases. To reduce the point-estimator bias we have combined two methods. In the first we colleted data on the real system to specify the initial conditions; in the other method each simulation run was divided into two phases: initialisation phase or warm-up period from time 0 to time T_0, followed by a data-collection phase from time T_0 until time T_0+T_E in which data collection, on the response variables of interest, is done. Namely the number of locomotives in queuing or being repair at the different repair facilities was considered.

To delimit the simulation interval length an iterative process was used comprising the calculation of a variable representing the relative number of the locomotives in no operation. While the series of the calculated values has not reached the steadiness the simulation interval was increased by 365 days. The steadiness was measured by comparing the calculated value with a threshold. The iterative process has pointed the value of 36500 days to affect to $\mathrm{T_0+T_E}$.

A graphical procedure duo to Welch (1983) was applied to define the warm up interval extension. Ten replications were done. The simulation interval was divided into twenty-five adjacent subintervals and the operational locomotives' average number was calculated for each one. The results are in Table 1.

The last column in Table 1 refers to the average of all averages calculated for each interval. If the two first values are removed the other are very smooth distributed around the value of 20,3, therefore the warm-up period was considered until the 4380th day occur.

After T_0 was obtained the value of T_0+T_E had been adjusted to 48180 days as suggested by Banks et al (2001).

Interval number	1º Run	2º Run	3º Run	4º Run	5º Run	6º Run	7º Run	8º Run	9º Run	10º Run	Average
1	17,238	18,033	19,317	16,256	17,926	16,953	15,448	16,637	18,630	16,699	17,314
2	18,534	20,844	17,468	19,536	19,293	16,910	17,239	18,735	21,839	19,147	18,955
3	21,442	21,720	21,133	20,986	20,027	20,855	20,182	19,553	20,956	20,611	20,746
4	20,722	19,144	19,910	20,193	21,378	21,314	20,156	20,039	21,347	21,236	20,544
5	18,888	20,226	19,494	20,584	20,135	19,449	20,141	19,942	20,651	18,857	19,837
6	21,029	20,488	20,471	19,933	19,845	19,758	20,599	19,418	18,785	19,941	20,027
7	19,412	20,473	19,130	22,249	21,431	20,954	20,520	20,656	20,110	20,093	20,503
8	19,472	21,705	21,095	19,683	19,333	20,857	18,305	20,340	20,618	20,627	20,203
9	19,604	20,579	20,557	20,145	20,889	20,143	22,395	19,478	19,819	20,617	20,423
10	21,016	21,651	19,720	19,834	19,823	20,816	20,192	20,229	21,398	20,331	20,501
11	19,700	20,394	21,238	20,958	19,564	19,763	21,146	20,871	21,284	21,147	20,607
12	20,628	21,482	20,001	19,532	20,148	20,423	20,634	20,516	19,015	20,508	20,289
13	20,864	20,332	19,879	21,397	19,589	21,366	20,643	19,808	20,747	20,226	20,485
14	18,685	21,802	20,457	20,502	20,735	18,930	20,530	19,950	19,615	20,779	20,199
15	19,995	19,707	21,001	20,012	19,938	21,059	20,103	21,949	20,422	20,645	20,483
16	18,903	20,387	20,778	20,374	20,739	21,568	20,439	19,696	21,514	21,084	20,548
17	22,108	20,972	19,696	20,203	18,811	19,425	21,004	21,487	19,623	21,171	20,450
18	19,859	20,589	20,864	19,284	20,655	20,238	19,924	20,226	20,326	20,044	20,201
19	20,373	19,796	19,044	21,767	20,843	19,604	18,796	19,639	19,241	19,885	19,899
20	20,481	20,068	20,752	21,193	20,476	20,378	21,163	20,775	18,722	20,356	20,436
21	20,923	19,877	20,655	19,199	19,136	19,542	21,403	20,211	21,365	19,509	20,182
22	19,860	20,775	19,525	20,255	20,547	20,123	20,325	20,780	20,643	19,186	20,202
23	20,932	17,376	20,642	19,937	20,492	21,516	18,004	20,371	19,503	19,877	19,865
24	20,474	20,556	20,958	20,436	20,032	19,919	21,185	20,425	19,538	21,113	20,464
25	20,587	22,193	20,260	20,951	19,627	20,228	19,338	21,509	20,433	21,024	20,615

Table 1: Welch procedure to define T_0

4.4. Simulation results

Approaches to estimate the number of operating locomotives

To calculate the estimator for the average number of operating locomotives we tested two methods: the replication / deletion approach and the batch means for interval estimation in steady-state simulation. The former was selected due to the reasonably good statistical performance given and due to the different system configuration comparison made possible, additionally to its easily understanding and implementation (Law & Kelton, 2000). The later was elected since different authors pointed out that data deleted in the replication / deletion method are wasted data, wasted computing time and suggested that there may be merit in using an experiment design based on a single long replication (Banks et al, 2001).

In the replication / deletion approach and following Banks et al (2001) proposal twenty five simulations with a 48180 days run length were performed. The interval was divided on thirty-three adjacent subintervals with a 1460 days length. The three first subintervals were affected by the warm-up period. For each of the remaining intervals the batch means within each replication was averaged to obtain a replication average. The values are express in the Table 2.

Run Nr.	1°	2°	3°	4°	5°	6°	7°	8°	9°
$\overline{Y}_{r\bullet}$	20,335	13,420	20,472	20,424	20,422	20,452	20,376	20,308	20,334
Run Nr.	10°	11°	12°	13°	14°	15°	16°	17°	18°
$\overline{Y}_{r\bullet}$	20,310	20,276	20,422	20,365	20,135	20,240	13,371	20,243	20,514
Run Nr.	19°	20°	21°	22°	23°	24°	25°		
$\overline{Y}_{r\bullet}$	20,500	20,508	20,412	20,589	20,249	19,408	20,459		

Table 2: Replication average to calculate the point estimator and its variability

Finally, the point estimator ($\overline{Y}_{\bullet\bullet}(n,d)$) for the average number of operative locomotive and its variability were computed using respectively the following expressions

$$\overline{Y}_{\bullet\bullet}(n,d)=\frac{1}{R}\sum_{r=1}^{R}\overline{Y}_{r\bullet}(n,d) \tag{1}$$

$$Var\left[\overline{Y}_{\bullet\bullet}(n,d)\right]=\frac{1}{R(R-1)}\left(\sum_{r=1}^{R}\overline{Y}_{r\bullet}^{2}-R\overline{Y}_{\bullet\bullet}^{2}\right) \tag{2}$$

in which n denotes the number of subinterval considers in the simulation run, d is the number of intervals associated to the warm-up period, R is the runs' number and $\overline{Y}_{r\bullet}$ represent the batch means average within a simulation run. The values obtained were 19,782 for the point estimator and 0,037 for its variability with a 95% confidence interval $[19{,}388; 20{,}176]$.

In the batch means for interval estimation in steady-state simulation method data was collected from the 4381st day until the 48180th day. This was grouped in a hundred and twenty subintervals for which the operational locomotive average number was calculated. Although the batch means are typically auto-correlated at all lags, the lag-1 autocorrelation is usually studied to assess their independence. The following expression

$$\hat{\rho}_1 = \frac{\sum_{j=1}^{k-1}\left(\overline{Y}_j - \overline{Y}\right)\left(\overline{Y}_{j+1} - \overline{Y}\right)}{\sum_{j=1}^{k}\left(\overline{Y}_j - \overline{Y}\right)^2} \qquad (3)$$

was used and the value of $(-0{,}204)$ was obtained. The test proposed by Banks et al (2001) was used to conclude that the correlation is sufficiently small. As an additional check, for batch means independency, the von Newman test revised by Alexopoulos & Seila (1998) was applied. Lastly the expression (1) was applied and the value 20,460 was reached for the average number of operational locomotives.

Considering a 95% confidence coefficient the confidence interval limits are 18,836 and 22,085.

The calculation performed was based on the no stock on hand assumption at the BRS or at any of the Maintenance Regions. This fact is justified by the situation verified in the real system.

A careful model analysis, in particular the information about the repair lines and its waiting queues reveals that few locomotives go to the Maintenance Regions. This is explained by the low TC failure frequency that is still decreased by the V1 or R preventive maintenance forestalled policy. In contrast the BRS is an overcrowded point. On average there are eight locomotives waiting during a hundred and three days for some kind of reparation at BRS and the average repair time is twenty-seven days. The largely waiting time is explained by the philosophy used whenever a locomotive suffers a breakdown or reaches the potential limit of any component. As told before the locomotive is stopped and leaded to the repair place. On contrary in the real system the locomotive continues working until it breaks down. This situation explains the lower simulated average number of operational locomotives, twenty locomotives, comparing to the practice average number, twenty-two locomotive. On the other hand, the reparation time in the real system can vary between twenty-five and fifty days values that allows to declare that some thing goes wrong within the repair centre.

Based on the low percentage of operational locomotives, only 67% if no stock was considered, we have decided to do a sensibility analysis' study for the TC, the DE, the TE and the WS inventory position. The batch means for the interval estimation in steady-state simulation method was used.

Sensibility analysis

To establish the relationship between the average number of operational 1900 Locomotives and the stock inventory position for the TC, the DE, the TE and the WS components a sensitivity analysis was performed in two phases.

In the first phase, the components on hand stock were changed individually, at the beginning, and jointly, later on. Founded on the TC, the DE, the TE and the WS failures frequency, the available storage area, the per unit price and the cost of carrying stock technicians had recommended that the values to consider for each first level component had to be different and should be between the intervals point out in Table 3.

First level component	Interval
Turbo Compressor (TC)	[0, 5]
Diesel Engine (DE)	[0, 3]
Traction Engine (TE)	[0, 30]
Wheel System (WS)	[0, 12]

Table 3: Stock on hand possible values

As result of this variation we concluded that the number of TC, DE and WS repaired in the BRS storage facility does not change the average number of operating locomotives; on the opposite, if there are six or more TE components the average number of operational locomotives is increase by one. This points out to the TE primacy. In addition the simulation resulted in the overcrowded characteristics of the BRS due to the constantly unavailable TE required for the locomotive repair. Although, a labour restructuring is not desired at the real system the simulation's results evoke the importance of a study to be carry out at the TE repair section so as to regard as a possible section reorganization as a form of increasing the average number of operational locomotives.

In the second phase and choosing the cheaper stock combination, the one that assumes only six TE in stock, the previously assumed value for the percentage of potential to anticipate a V1 or a R preventive maintenance, if a failure occur, was alter to the following 25%, 50% and 90%. In these scenarios the operational locomotives' average number was computed. The values are indicated in the Table 4. It can be concluded that on average there is one more operational locomotive. Then 70% of the total locomotives will become operational.

Potential % that anticipate a preventive reparation	Operating locomotives' average number
25%	18
50%	19
75%	21
90%	21

Table 4: Operating locomotives' average number depending on potential percentage that anticipate a preventive reparation

Finally, an increase of 20% was assigned to the potential limit considered before. This increment is justified since security motivation and material's characteristics force the value taken for granted to be much lower than the accurate limit value. In the real system this increase is frequently used to delay the V1 and the R preventive maintenance and to reduce the number of locomotives waiting for reparation at the BRS. Also, again the simulation was run considering six TE stocked and the percentages of 25%, 50% and 90% for the potential percentage that anticipate a V1 or a R preventive maintenance when a breakdown occur. The news values are depicted in Table 5.

Potential % that anticipate a preventive reparation	Operating locomotives' average number
25%	20
50%	22
75%	23
90%	24

Table 5: Operating locomotives' average number depending on potential percentage that anticipate a preventive reparation

As a matter of fact, comparing the homologous values from Table 4 it can be concluded that, in mean, the operational locomotives' average number are augmented by two. This improvement can be considered significantly as it represents an increase of 7% over all the operational locomotives.

So in conclusion we can state that if there are no stock and if the 75% of the potential limit is assumed to anticipate a V1 or a R preventive maintenance, if a failure occur, there are on average twenty operational locomotives and ten locomotives at the BRS waiting for repair. If we assume six TE in stock there are twenty-one operational locomotives. The number of TC, DE and WS repaired in the BRS storage facility does not change the average number of operating locomotives as well as if there are more than six TE in stock. If the percentage of the potential limit assumed to anticipate a V1 or a R preventive maintenance, if a failure occur, is altered to 90% and considering six TE in stock then on average twenty four of the thirty locomotives are operational.

5. CONCLUSION AND FUTURE RESEARCH

A large number of multi-echelon inventory models for repairable items have been developed since three decades ago. However, even the most recent models have restrictions to their successful application.

The dominant model both in literature and in practical applications is the METRIC. This has been extensively used even nowadays although it understates the backorders at the base and consequently overstates system operability.

We have defined a methodology based on a simulation model that can be applied to all multi-echelon problems besides the number of echelons and its characteristics. The model allows the simultaneous management of repairable or consumable items and considers, at the same time, a small population of items subject to failures and a limited repair facility. Balancing the complexity and the assumptions the proposed model allows the study of real cases. It permits the modelling of every kind of systems with the desired detail's level and lets a wide range of probability distribution functions to be associated to the decision variables. Thinking of an open system, it also integrates a replenishment policy where demands can be emitted automatically interacting with suitable software.

The model contribution is proved trough its application to the stock management problem of the 1900 Locomotives components. Not only does the model permit to establish a relationship between the operational locomotives' average number and the stock inventory position for the components but also shows the system overcrowded points as for the case of the Barreiro Repair Shop due to the unavailable Traction Engine required for the locomotives repair.

Further research can be focus on the study dependence of the failure probability distribution function from the number of working items. Also the introduction of cost is one of our priorities. In addition, we would like to study the model adaptation to acyclic multi-echelon structure in particular to a supply chain system.

REFERENCES

ALEXOPOULOS, C. & SEILA, A. 1998. Output data analysis, in *Handbook of simulation – Principles, methodology, advances, applications and practice*. New York: John Wiley & Sons.

BANKS, J., CARSON, J., NELSON, B. & NICOL, D. 2001. *Discrete-event system simulation*. New Jersey: Prentice Hall Inc.

BOX, G. & JENKINS, G. 1970. *Time series analysis, forecasting and control*. San Francisco: Holden-Day Inc.

CUNHA, M. E. 2003. *Utilização de modelos de gestão de stocks para areas de multi-escalão de produtos reparáveis*. Ph.D. thesis, IST - Lisbon University.

CUNHA, M. E., PÓVOA, A. P., LOPES, J. A. 2004. A multi-echelon system's simulation model for reparaible and consumable items management. *Proceedings of the ICS'2004 – Industrial Simulation Conference*, 353-357, Malaga, Spain, J. Marin & V. Koncar (eds).

DÍAZ, A. & FU, M. C. 1997. Models for multi-echelon repairable item inventory systems with limited repair capacity. *European Journal of Operational Research*, 97: 480-492.

EXTEND^{TM}, Version 4.1.1. 1998. *Imagine That Inc.* A personal computers simulation programme with Windows environment.

GROSS, D., MILLER, D. & SOLAND, R. 1983. A closed queueing network model for multi-echelon repairable item provisionig. *IIE Transactions*, 15: 344-352.

GUPTA, A. & ALBRIGHT, S. 1992. Steady-state approximations for a multi-echelon multi-indentured repairable item inventory system. *European Journal of Operational Research*, 62: 340-353.

GUIDE Jr, V. & SRIVASTAVA, R. 1997. Repairable inventory theory: models and applications. *European Journal of Operational Research*, 102: 1-20.

LAW, A. & KELTON, W. 2000. *Simulation Modeling & Analysis* (3rd ed). Boston: McGraw-Hill International Editions.

MUCKSTADT, J. A. & THOMAS, L. J. 1980. Are multi-echelon inventory methods worth implementing in systems with low-demand-rate items?. *Management Science*, 26: 483-494.

RICHARD, F. 1976. A stochastic model of a repairable item inventory system with attrition and random lead times. *Operations Research*, 24: 118-130.

RUSTENBURG, W., van HOUTUM, G. & ZIJM, W. 2001. Spare parts management at complex technology-based organizations: an agenda for research. *International Journal of Production Economics*, 71: 177-193.

SHERBROOKE, C. 1968. METRIC: A multi-echelon technique for recoverable item control. *Operations Research*, 16: 122-141.

SIMON, R. 1971. Stationary properties of a two-echelon inventory model for low demand itens. *Operations Research*, 19: 761-773.

SLAY, F. 1984. VARI-METRIC: an approach to modelling multi-echelon resupply when the demand process is Poisson with Gamma prior. *Logistics Management Institute*, Washington, D. C. Report AF301-3.

SPSS, Version 11.0.0. 2001. *LEAD Technologies Inc.* A personal computers statistics programme with Windows environment.

EXTEND^{TM}, Version 4.1.1. 1998. *Imagine That Inc.* A personal computers simulation programme with Windows environment.

WANG, Y., COHEN, M. & ZHENG, Y. 2000. A two-echelon repairable inventory system with stocking center dependent depot replenishment lead times. *Management Science*, 46: 1441-1453.

WELCH, P. 1983. The statistical analysis of simulation results, in *The Computer Performance Modeling Handbook*. New York: Academic Press.