SPARE PARTS AND MAINTENANCE OPTIMIZATION IN A MOBILE TELEPHONE COMPANY

Ángel Díaz

Instituto de Empresa
Dept. Operations
& Technology Management
C/ María Molina, 12, 5º
28006, Madrid - Spain
angel.diaz@ie.edu

Abstract

Telecommunication companies depend on high availability of equipment to maintain service quality. In cellular communications electronic cards maintenance is basically reduced to exchanging parts as they fail. These parts are geographically dispersed in unmanned locations. Spares and maintenance policies are thus interrelated and tend to follow multiechelon configurations, following the architecture of the physical network. We describe the optimization of spare parts and maintenance policies performed by the Venezuelan mobile phone operator Movilnet. Both system issues (reduction of SKU, classification of items, network practices, and data integrity) and modeling issues, centered on the adaptation of the METRIC model to the simultaneous optimization of spare parts and personnel resources, are discussed. The systematic application of these resulted in important reductions in the purchasing of spare parts, without quality deterioration.

Key words

Spare parts optimization, MRO, multi-echelon inventory systems, mobile phone.

Importance of Spare Parts Logistics. The Case of Movilnet

Spare parts are members of the MRO (Maintenance, Repair and Operation) materials family. As such they represent an important portion of all purchasing activities and have been the recent object of e-market initiatives that try to reduce purchasing complexities through the advantages of the Internet.

Spare parts are characterized by slow flows with stochastic demand (making deterministic demand modeling –i.e., M.R.P.- difficult and requiring the use of Poisson or Negative Binomial distributions), frequent irregular patterns (making forecast difficult) and large variety (e.g., over eighty thousand at the venezuelan brewery Polar). An important type of spare parts in industrial practices is repairable items, due to increased requirements for high availability that can be achieved through off-line maintenance (exchange of failed modules that are then repaired off-line). This is the case of mobile phone operators where radio cards are geographically dispersed in hundred of unmanned bases, self-contained small units that handle the call flows. These cards are monitored from a central control unit, so that when one fails a technician is directed to take one spare card from central stock and to displace to the base where the card is exchanged. If the technician takes the wrong type of card downtime increases significantly. Different maintenance policies are thus possible: centralized personnel and spares, decentralized personnel and spares, or a combination of these. Further, the evaluation of the precise number of cards required and their locations is important as spares are frequently bought as a part of a package without any attempt at modeling their exact requirements. Movilnet, the leading mobile phone operator of Venezuela managed important reductions in spare parts requirement without quality loss through a systematic approach that included system and modeling issues.

System issues

Before modeling any spare parts problem, system issues have to be considered to reduce noise that would otherwise compromise the precision of the model. The following four-step improvement process used at Movilnet attempts to reduce this noise.

1. Coding. The objective of this step is to reduce the number of Stock Keeping Units (SKU). Easy access to on-line and non-electronic catalogues is a basic consideration, as is the use of tactics for parts redundancy avoidance. Among these figure the identification of common parts, the avoidance (when technically feasibly) of coding by brand, and the association of specific parts to their parent equipment, to facilitate the removal of obsolete items from the catalogue.

2. Classification. The objective of this step is to facilitate focus. Classifications can include: Repairables or consumables. Support materials can be consumable or repairable -according to the feasibility of a repair and re-utilization after a failure. The last type is of particular importance for industries or services with heavy utilization of equipment, both for economical implications and for their impact on the continuity of operations. Repairable items, usually fixed assets, must be traced individually, and their warranty status (i.e. new and under warranty or used and out of warranty) must be clearly identifiable on an item-by-item base.

Stockable or direct usage. Not all items must be stocked. Those with infrequent demand and low criticality can be bought on a per-need basis and sent directly to the user, avoiding stock that can result in immobilized parts. This is particularly important in project materials, where non-utilized parts are frequently returned and kept in the warehouse.

Pareto analysis by turnover and on-hand value. The traditional Pareto analysis can be applied on a utilization or on-hand value base. Calling these categories ABC for turnover (the traditional 20-80 rule of A items constituting a minority with the highest utilization), and XYZ for on-hand value (X items are few, but represent the majority of the on-hand value), consistency analysis can be performed, as suggested by Figure 1a (i.e., a high usage part with low existance level poses stock-out dangers).

Criticality. Critical parts have a much greater impact on operations than non-critical parts. Here a multi-attribute definition of impact is required, as an item can be considered critical depending on the effect that a stockout has on the system (the so-called operational criticality), but also on its effect on safety, how difficult it is to buy or the length of the lead-time. Some authors have proposed aggregate measures, using techniques such as AHP (Cohen and Ernst, 1988, Flores et al., 1993). An item can be of type A and have low criticality (e.g., office supplies), or be of type C and have high criticality (e.g., a high reliability relay switch). It is interesting to relate these classifications. Although in Movilnet all radio cards are considered critical, Figure 1b (containing real data from Caracas' Subway, with the vertical axis showing the number of items) shows that typically less than 10% of all items have high turnover and criticality. This can dramatically reduce the required effort from, for example, seemingly unwieldy 40,000 items to a more manageable 4000.

See figure 1.

3. <u>Network optimization</u>. The objective of a network approach in warehousing is to provide costs reduction, due to portfolio effects (variance reduction due to centralization of stocks, with small, decentralized stocks to improve service quality), and increased purchasing power and control. On the external network (i.e., suppliers), exciting advances can be found in e-market sites that facilitate purchasing and reduce the cost of transactions in a significant way, as MRO can represent 80% of all purchasing transactions (Shapiro and Hesse, 1999). Market sites have been developed in the last two years include sites for industrial spares (tpn.geis.com), health sector spares and consumables (neoforma.com), automobile spares (covisint.com) and electronic parts (e2open.com). Supplier-client alliances effectively outsource the managing of spares to organizations that benefits from existing expertise and economies of scale (e.g., Caterpillar for DaimlerChrylser and TNT for Fiat, Parker, 1999). The advantages of a network approach for spares parts is illustrated in the small example in the next section that shows a 25% reduction in stock, achieved through a jointly optimized multi-echelon approach.

4. <u>Information systems</u>. The objective of this step is to achieve data integrity. Key practices are related to the successful adoption of ERP systems: real time processing, centralized databases and a reengineering and strategic, rather than system, approach. Additional issues are related to improper user practices, such as not requesting an item if inventories levels are inferior to needs, or of requesting the same material more than once (making real demand determination difficult).

A particularly important capability of a spare parts information system is forecasting, critical to the removal of noise in the system. Large quantity of items usually precludes the use of transversal techniques (i.e., multiple regression), so time series methods are commonly used. For this to be successful, trained analysts should focus on critical items using flexible forecasting system that allow the use of automatic diagnostic tools (e.g., tracking signals), different forecasting methods, grouping of periods (critical when demands are ocassional due to low failure rates, as in Movilnet) and data filtering. A trained analyst can recognized in the data the presence of impulses, frequently produced by the overlapping of a punctual demand (e.g., an overhaul, best treated as a direct usage) on top of the regular demand; and of ramps, produced by an augmentation of the installed base of equipment, of its usage, or of wear-out, as well as more classical trends and seasonal effects.

The irregular pattern of demand manifested in Figure 2 (four years demand of brake pads at Caracas' Metro) is due to the existence of a non-regulated warehouse. When parts arrive, maintenance withdraws them from the central warehouse and keeps them in their own. As the second warehouse is not connected to a centralized data base, a telling-tale pattern of peaks and valleys results, making forecasting very difficult. The obvious action is to eliminate, or connect information-wise, this warehouse.

See figure 2.

Modeling was attempted at Movilnet after careful consideration of these issues.

Modeling issues

Inventory models answer the important questions of when to replace the stock and in what quantity. Although in practice the three basic types of inventory models (independent demand, dependent demand and synchronized flows) are combined, independent demand models dominate MRO materials, due to the characteristics mentioned in the introduction. The application of independent demand models also differs from those used in production or retailing due to the slowness of demand (desemphasizing batching and leading to the use of Poisson distributions) and to the existence of cyclic processes for repairable items. These issues are examined bellow.

Slow demands call for asymmetrical distributions, of which the dominating type is the Poisson distribution. In the small example bellow, a Poisson distribution is used for scenarios one and two. In spite of its limitations (e.g., variance to expected value fixed at one), Poisson distributions have the advantage of being easy to use and flexible in application when combined with another distribution in a compound Poisson (Feeney and Sherbroke, 1966).

The combination of large variety, slow usage and frequent high criticality makes the use of multiechelon systems for spare parts particularly attractive, as it can produce a portfolio effect, decreasing the amount of inventory required.

See figure 3.

Figure 3 shows a typical cycle for a repairable item (the following models can be also used for consumable items by substituting repair and transportation times for lead times). In this two processes happen simultaneously: the failed part is exchanged and pushed into the repair process (transit to depot, queuing at the repair station, repair) while simultaneously a part is pulled from the stock at the depot and sent to the base to fill the part used in the exchange. The base stock can be calculated by making the lead-time equal to the transportation time from the depot to the base plus the expected delay produced when there are no parts available at the depot. The calculation of the last quantity, $E(W_0)$, is a little more involved and is the heart of the METRIC (Multi-Echelon Technique for Recoverable Item Control) model, proposed by Sherbrooke (1968) and formulated in appendix 1.

The following small example illustrates the advantages of METRIC. A repairable item is required in two locations, each with demand □□of 0.1 units per week. The repair process takes 10 weeks. Three scenarios, represented in Figure 4 are used. Under the first no Depot is used and parts are delivered directly to the bases. In the second scenario a Depot (with one week transportation time to the bases) is used, but stock determination is performed independently at each position (i.e., the bases see the Depot as their supplier). Under the third scenario a joint optimization of the Depot and the three bases is performed, using METRIC. In all cases the objective Fill rate is at least 99%.

See figure 4.

Scenario 1: each base expected backorders (EB_i) is modeled independently using a Poisson distribution: $EB_i(s_i=0)=$ □$_i L_0=1.2$ (s_i is the number of spares at the base, while L_0 is time to repair site and back, two week, plus repair cycle time, 10 weeks). $EB_i(s_i=4)=0.007$, for a fill rate of 99.46%; Thus, for scenario 1 total stock is 8 parts.

Scenario 2: The bases are modeled as before, but with a lead-time of one week, resulting in $EB_i(s_i=1)=0.005$ for a fill rate of 99.53%. At the Depot expected backorders for a given level of spares, s_0, are $EB_0(s_0=0)=2$, and $EB_0(s_0=6)=0.006$, for a fill rate of 99,53%. Although total stock for scenario 2 is also 8 parts, total costs can be reduced through better purchasing and control. This configuration is extensively used in practice.

Scenario 3: Using METRIC, $EB_0(s_0=0)=2.2$ (transportation time from base to Depot is added to the repair cycle time), and $EB_0(s_0$□□$0)$ are calculated with a recursive expression (appendix 1). $EB_i(s_0, s_i=0)= \lambda_i[L_i+E[B_0(s_0)/\lambda]$ (where L_i is the time from Depot to base) and $EB_i(s_i$□□$0)$ are similarly calculated, resulting in a minimum total stock configuration that satisfies the fill rate requirements of two parts at the Depot and two parts at each base, for a 25% stock reduction over both scenarios one and two.

The use of Poisson distributions makes METRIC simple, but limits its usefulness, due to its fixed coefficient of variation and the inability to capture queuing times at the repair place. Graves (1985) proposed the use of a negative binomial distribution that fits a two-parameter distribution to the distribution of the bases outstanding orders. This approximation requires calculating the variance of the backorders at the depot with a simple recursive expression. This approximation is used in commercial software such as OPUS and OPRAL (Alfredsson, 1999). Graves (1985) and Axäter (1993) proposed exact methods that are unfortunately intensive in calculations, so the use of approximations is commonly observed in practice. The Modeling of limited repair facilities is a common-life restriction that cannot be supported by METRIC, as it underestimated the required spares (queuing time spent at the repair shop can be up to a whopping 90% of the total time spent in repair and can not be considered by the model). Although in the case of operating systems the queuing time can be observed from practice and added to the repair time, there is no obvious way to estimate this time in new systems, or when historical data is not available. The M/G/k multi-class model (Díaz and Fu, 1997), approximates the number of parts in the repair process to a negative binomial distribution, using expected values and variances of the waiting times that can be estimated directly from sample data. This model is not restricted to assuming exponential service times and ample repair facilities, and can handle items with different repair distributions being repaired in the same facility. The application of this model requires the estimation of the squared coefficient of variation of the waiting times in queue, C_W^2, from the class parameters, using either PK formulas or approximations. Applied to the Metro (subway system) of Caracas resulted in a proposed 50% reduction of required spares, over the original stock purchased following the recommendations of the supplier of the system, without quality deterioration (Diaz and Fu, 1997).

Among other assumptions that have been relaxed are: lateral resupply between bases (Sherbrooke, 1992, Axsäter 1990), non-repairable failures, cannibalization of parts, commonality of parts (Sherbrooke 1992), batching (Axsäter 1993) and multi-indenture (inclusion of subcomponent, or indenture).

Multi-indenture models assume each system to be composed of main components, or Line Repairable Units (LRU), and these of subcomponents or Shop-Replaceable Units (SRU). When an LRU fails, it is exchanged for a spare unit, taken from the base, or backordered if none are in stock. In both cases the LRU is repaired at the base or at the depot according to the type of failure. At the repair place, the faulty SRU is identified and substituted for a spare or backordered if none is available. The original formulation of the multi-indenture problem, called MOD-METRIC, is due to Muckstadt (1973).

Other modeling issues include optimization, performed either by Lagrangean multipliers (Fox and Landi, 1970), Sherbrooke's marginal analysis, or through complete enumeration in the case of small problems. Searching systematically with these optimization techniques, expected backorders and investment trade-off curves could be obtained, allowing managers to make strategic decisions. It is further possible to express expected backorders in terms of availability and cost by modeling the relationships of the parts to the system as a whole through the use of conventional reliability modeling techniques, as shown in the Movilnet case.

The next section shows how METRIC was adapted to the case of Movilnet, the leading mobile phone operator of Venezuela.

Simultaneous Modeling of Spares and Maintenances Policies. The Case of Movilnet

Fast growth, and continuous changes in technology and regulation, forces the providers of cellular services to maintain high service quality, while maintaining or reducing operating costs. The infrastructure of the leading operator in Caracas, Movilnet, is formed by two main facilities, known as MTX, where equipment, spares and personnel are centralized, and about 80 bases per MTX, small self-contained units where the radio cards are located (Rodríguez and Rodríguez, 1998). Card behavior is monitored from a central location, and in case of failure a technician obtains a card from the stock at the MTX and exchanges it at the base for the failed one. This is then sent in batches to Sweden (all cards are made by Ericsson), where they are repaired and returned to Caracas. At the moment of the study Movilnet (1998) was considering acquiring additional spares to be kept decentralized at some of the bases. Due to the very long repair cycle (two months) corrections for limited repair facility were considered unecessary.

See figure 5.

Three different scenarios were modeled:

<u>**First scenario**</u>**:** Parts and personnel are centralized in the MTX, which serves 80 bases. There is 1-p_r chance of making a wrong diagnostic before leaving for the base, in which case the repairpersons go back to the MTX to get the right part. This is a single echelon model, all equations are given in appendix 2.

L_{i0} is the transportation time from the MTX to the closest base in a cluster, or group of bases, L_{ii} the time between bases in a cluster and p_r the probability of a right diagnostic (in which case transportation time from MTX to a base is $L_{ii}+L_{i0}$. If the diagnostic is wrong, with probability 1-p_r, the time is treble, as repair persons have to go back to the MTX – second-time mistakes were not considered). $E[B_0(s_0=0)]$ is λL_r, the expected number of failures during the repair cycle. From empirical data, L_{ii} is 0.5 hr, L_{i0} is one hour and λ_i (base failure rate) is 0.0028 failures per day. The application of the model showed that p_r has little effect on backorders, dominated by a long repair cycle of an average two months, and was thus dropped from the other scenarios.

Second scenario: Parts are held at the MTX and at special bases that work as servers for a cluster of bases of variable size, but repairpersons remain centralized at the MTX. This is a multiechelon model, with $E[B_0]$ calculated as in scenario one, but with p_r=1, and different expected backorders formulas at the bases with and without spare parts (appendix 2).

For a given amount of spares at the depot, different sizes of clusters and of spares per cluster were modeled. The size of the cluster was found to have little impact on results, and the resulting expected backorders were on average slightly smaller than those for the same total quantity of spare parts under scenario one (a 7% improvement with 18 spares, for example).

Third scenario: Parts and repairpersons are now located at the MTX and at the server bases. This case is similar to scenario two, with a shorter transportation time to the bases.
This model improves scenario two slightly (10% improvement with 18 spares, with respect to scenario one), with the added complexity and cost of decentralizing repairpersons and spares.

Results: The estimation of spare levels is largely determined by the repair cycle at Ericsson, and thus the quality of the diagnostic (pr) has little effect on resulting spare levels. The second and third scenarios showed marginal improvements over the basic single-echelon model, preferred for simplicity and ease of control. Results of the first model were thus validated against simulation with 10 replications, finding less than three-percent average differences. Expected backorders and unavailability (penalty) costs, plotted against spare levels are shown in Figure 6. Penalty costs are simple to estimate, as there are 30 cards per base, with a utilization of 19.53% analog calls,15.42% digital calls and 65.05% idle time, and the cost of a call is an average $15.76 per hour. Unavailability with 10 spares, for example, is 0.16% (3.746 cards off out of a total of 2400), for a penalty cost of $339,465.24 per year.

See figure 6.

It is also possible to estimate the optimal number of spares, as each card has a cost of about $3000, and an estimated useful life of four years (due to technological obsolescence). This results in a minimum cost over four years, estimated at current costs, of almost 79 thousand dollars with 23 spares.

It was therefore recommended to hold around 23 spares in a centralized location (the MTX), for an expected availability of 99.999%. It is noteworthy that the actual number of spares for that card was 90, some 65 above real needs. Movilnet consequently reversed the proposed policy of incrementing and decentralizing cards, and has used the excess cards for the expansion of the network, realizing important economies. Movilnet further implemented information systems for these models, not supported by their ERP system. Movilnet also performed extensive modifications to their ERP package to support bar code-based tracking of the spares and the application of the Metric models. Movilnet is currently considering a second supplier of radio card and the economic implications of considering radio cards as consumables (discarding them after each failure).

Conclusions

Spare parts optimization is not as easy as 1-2-3. Considerable effort is required to remove noise from the system, but once these systematization steps have been achieved, the proper application of inventory models, and particularly of multi-echelon inventory models for repairable items, can produce appreciable savings in spares investment.

Appendix 1. METRIC formulation

$E(B_0(s_0))$ the expected number of backorders at the depot (units demanded but not satisfied when there are s_0 spares at the Depot) can be calculated by:

$$E(B_0(s_0) = \sum_{x=s_0+1}^{\infty} (x - s_0) \frac{(\lambda_0 L_0)^x}{x!} e^{(-\lambda_0 L_0)}, \text{ where}$$

L_0=average lead time at the central warehouse (or repair time at the depot, including repair time and transit time, in the case of repairable items)
λ_0=demand rate at the depot (sum of demands at bases, $\Sigma\lambda_i$),
s_0=spares target level at the depot

Backorders at each base i when there are 0 spares at the base and s_0 at the Depot are simply obtained with:

$$E[B_i(s_0, s_i=0)]= \lambda_i[L_i+E[B_0(s_0)/\lambda_0], \text{ where}$$

L_i=average lead time from central warehouse to base i

With simple recursive expressions the process is repeated for possible combinations of s_0 and s_i to obtain $E(B_i(s_0,s_i))$, expected backorders at base i when there are s_0 spares at the Depot and s_i at that base.

To simplify calculations, deterministic lead times and Poisson distributions are assumed. By a theorem due to Palm (1938), the distribution of the outstanding orders at the depot (i.e., those parts in repair) is Poisson when the failure process is Poisson and the repair facilities are ample, for any given distribution of repair times (which can then be considered to include the transportation time from the base to the depot). If some failures can be repaired locally at the bases, the model can be easily modified as follows:

$$E[B_0(s_0=0)]=[\Sigma(\lambda_i(1-p_i))]L_0 \text{ and } E[B_i(s_0,0)]= \lambda_i\{p_i r_i+(1-p_i)\ [L_i+E[Bo(s_0)/\lambda_0]\},$$

where p_i is the probability of a local repair at base i, r_i is the average local repair time at that base and λ_0 the total demand at the depot, or $\Sigma(\lambda_i\ (1-p_i))$.

Appendix 2. Adaptation of METRIC to Movilnet´ scenarios

<u>Scenario 1</u>: Centralized spares and personnel
Expected backorders at the bases:

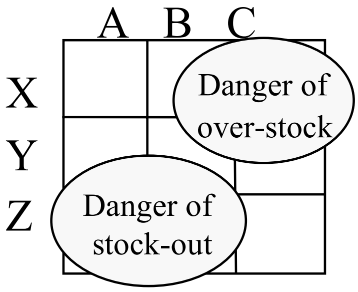
$$E[B_i(s_0)]= \lambda_i[3(L_{ii}+L_{i0})-2(L_{ii}+L_{i0})p_r+E[B_0(s_0)/\lambda]$$

<u>Scenario 2</u>: Selctive decentarlization of spares, centralized personnel

$E[B_c(s_0, s_i)]= E[B_0(s_0)](\lambda_c/\lambda)+\lambda_c L_{i0}$
$E[B_i(s_0, s_i)]= E[B_c(s_0, s_i)] + \lambda_i(L_{ii}+ L_{i0})$ (base without spare parts)
$E[B_i(s_0, s_i)]= E[B_c(s_0, s_i)] + \lambda_i L_{i0}$ (base with spare parts)

<u>Scenario 3</u>: Selctive decentralization of parts and personnel

$E[B_c(s_0, s_i)]= E[B_0(s_0)](\lambda_c/\lambda)+\lambda_c L_{i0}$
$E[B_i(s_0, s_i)]= E[B_c(s_0, s_i)] + \lambda_i L_{ii}$ (base without spare parts)
$E[B_i(s_0, s_i)]= E[B_c(s_0, s_i)]$ (base with spare parts)



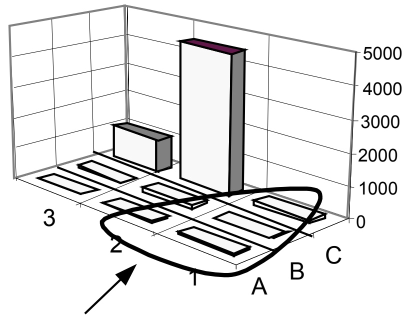


Figure 1. Stock (XYZ)-Turnover (ABC) Pareto (left) and Criticality (123)-Turnover (ABC) Pareto (right) diagrams facilitate inventory analysis when facing large variety of parts

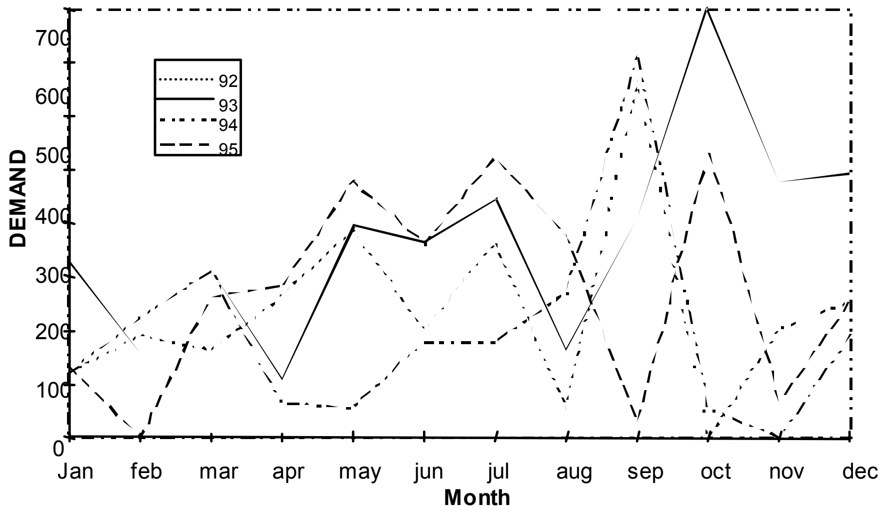


Figure 2. Demand for brake pads at Caracas' Metro. This confusing pattern results from an uncontrolled structure of multiple warehouses

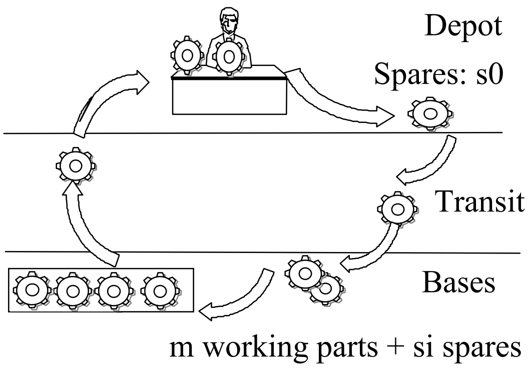


Figure 3: the repair cycle in in a multiechelon setting

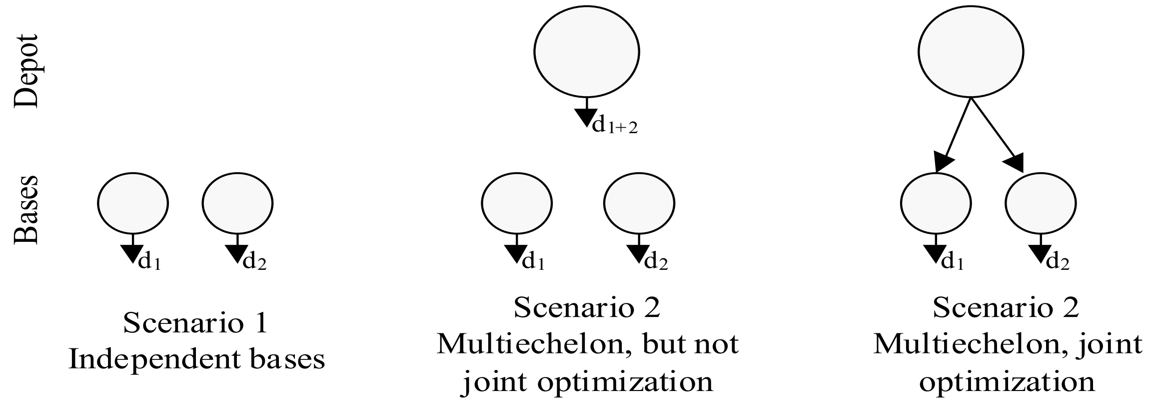


Figure 4. Different scenarios for decentralized stock

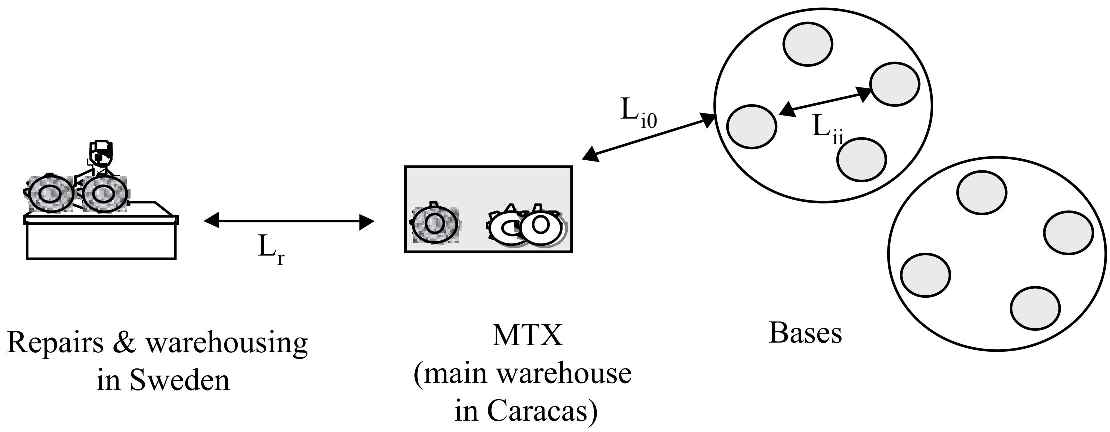


Figure 5. Movilnet multiechelon setting

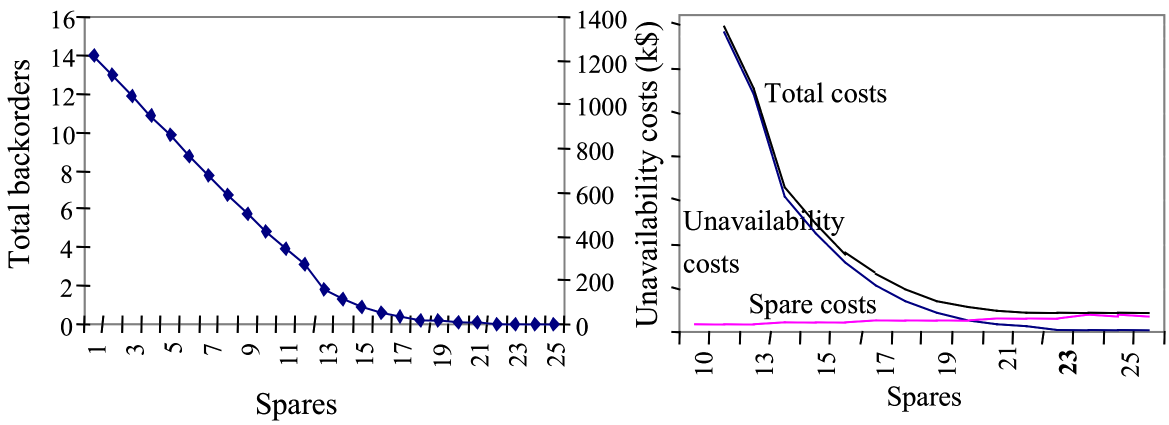


Figure 6. Backorders and unavailability costs – Movilnet

References

Alfredsson, P. "OPRAL – A Model for Optimum Resource Allocation. Proceedings of the International Conference on Industrial Logistics, 1999.

Axsäter, S. Continuous Review Policies for Multi-Level Inventory Systems with Stochastic Demand. In S. Graves, A. Rinnooy Kan and P. Zipkin (Ed.) Logistics of Production and Inventory, North Holland, Amsterdam, 1993.

Axsäter, S. Modelling emergency lateral transshipments in inventory systems. Management Science v36 n11 p1329 Nov. 1990.

Díaz, A. and Fu, M. Multiechelon Inventory Systems for Repairable Items with limited repair facilities. European Journal of Operations Research, 97 (1997) 480-492.

Feeney G., Sherbroke, C. The (s-1, s) Inventory Policy under Compound Poisson Demand. Management Science v12 p391 1966.

Fox, B. and Landi, D. Searching for the multiplier in one constraint optimization problems. Operations Research v18 n2 p253. 1970.

Graves, S. A. Multi-echelon Inventory Model for a Repairable Item with one-for-one Replenishment. Management Science V.31-10. 1985.

Muckstadt, J. A model for multi-item, multi-echelon, multi-indenture inventory system. Management Science v20 p472. 1973.

Palm, C. Analysis of the Erlang traffic formula for busy signal assignment. Ericsson Techniques v6 p39. 1938.

Parker, J. DaimlerChrysler picks Caterpillar. Traffic World, v259, n13 p16. Sep 1999.

Rodríguez, M. and Rodríguez, C. Diseño de un sistema de optimización y control de inventario asociado a los sistemas de telefonía móvil celular. UCAB thesis, Caracas, 1998.

Sherbrooke, C. Metric: A Multi-echelon Technique for Recoverable Item Control. Operations Research, 16.2. 1968.

Sherbrooke, C. Multi-echelon Inventory Systems with Lateral Supply. Naval Reseach Logistics, V.39. 1992.