



Approximate dynamic programming for management of high-value spare parts

Hugo Simao and Warren Powell

*Department of Operations Research and Financial Engineering,
Princeton University, Princeton, New Jersey, USA*

Approximate
dynamic
programming

147

Received January 2008
Revised June 2008
Accepted July 2008

Abstract

Purpose – An aircraft manufacturer faces the problem of allocating inventory to a set of distributed warehouses in response to random, nonstationary demands. There is particular interest in managing high value, low volume spare parts which must be available to respond to low-frequency demands in the form of random failures of major components. The aircraft fleet is young and in expansion. In addition, high-value parts can be repaired, implying that they reenter the system after they are removed from an aircraft and refurbished. This paper aims to present a model and a solution approach to the problem of determining the inventory levels at each warehouse.

Design/methodology/approach – The problem is solved using approximate dynamic programming (ADP), but this requires developing new methods for approximating value functions in the presence of low-frequency observations.

Findings – The model and solution approach have been implemented, tested and validated internally at the manufacturer through the analysis of the inventory policy recommendations in different network scenarios and for different pools of parts. The results seem promising and compelling.

Originality/value – The uniqueness of this research is in the use of ADP for the modeling and solution of a distributed inventory problem. Its main value resides on the incorporation of the issue of spatial substitution in demand satisfaction within the problem of determining inventory levels in a distributed warehouse network.

Keywords Inventory management, Warehousing, Spare parts, Aircraft components

Paper type Research paper

1. Introduction

An aircraft manufacturer will be introducing during 2008 a new fleet of executive jets using entirely new parts, and needs the ability to respond to random demands as the parts fail. There are over 10,000 parts, ranging from small parts that will be needed with a somewhat predictable frequency, to high-value, low-volume parts (engines, major wing components, avionics, and major hydraulic systems) which will fail at a rate that changes over time, producing two sources of nonstationarity in the demand for the parts. First, there will be rising demand as more aircraft enter the marketplace. Second, the rate at which components fail will likely follow the standard “bathtub” curve, with higher failures at first due to initial design problems, lower failure rates during the main part of the lifetime, finally followed by higher failure rates as they reach the end of their lifetimes.

Further complicating the problem is that because these are new parts, the rate at which parts will fail is uncertain. While manufacturers will provide initial estimates for expected



lifetimes, it is quite common for actual lifetimes to differ, possibly significantly, from these initial estimates. It will be necessary to detect shorter than expected lifetimes as quickly as possible so that inventories can be adjusted. Lower-than-optimal inventories produce either higher transportation costs (as parts have to be rushed from suppliers) or delays in responding to customer requests.

Since the market for executive jets is composed in large part by individual owners, the manufacturer needs to set up a dedicated logistics network for the worldwide storage and distribution of spare parts. The manufacturer has proposed a network composed of a small number of strategically located central warehouses (possibly one or two per continent) and regional warehouses located at service locations (possibly one at each service station). If, on the one hand, the centralized warehousing allows for risk pooling in the demand satisfaction, on the other hand, the distributed inventories aim at increasing the level of service (LOS) of demand.

Routine component failures are detected during scheduled maintenance and thus the part replacement may be done during a time window as wide as four days. Centralized warehousing should be able to handle these types of demand well. Critical component failures pose a different problem, but still the aircraft can fly to the nearest service center and there be serviced within a time window of up to two days. It is, however, the failure of components that leave the aircraft-on-the-ground (AOG) that pose the most significant challenge to centralized warehousing. In general, parts need to be shipped from a warehouse to the airfield or airport where the failure happened within time windows no longer than six hours from the detection of the demand. These are the cases where distributed warehousing may pay off.

The challenge faced by the manufacturer in establishing an inventory policy for a distributed spare parts network resides in balancing the requirements of high levels of service for AOG and critical demand with the budget constraints of keeping a high level of inventory at the warehouses.

This problem is of fundamental importance, with applications to a broad range of resource allocation problems. This is the same problem class that arises in the allocation of equipment, people and resources so that one can respond quickly to events such as hurricanes, terrorist attacks and other emergencies.

In the context of presenting a dynamic allocation heuristic for the centralized safety stock problem (one central warehouse and several regional ones), Cao and Silver (2005) offer a brief survey of the work done in the subject. It comprehends modeling and solution approaches that range from variations in traditional inventory control policy to mathematical programming formulations to heuristics. In this paper, stemming from the authors' expertise with the application of approximate dynamic programming (ADP) to resource allocation problems (Godfrey and Powell, 2002; George and Powell, 2006; Powell *et al.*, 2007; Powell, 2007), an ADP model and algorithmic strategy are proposed to solve the problem in question.

In the next section, the particular characteristics of the problem will be described. In Section 3, an ADP model formulation will be presented. The issue of low-frequency observations concerning the estimation of the value functions will be discussed in Section 4. In the final section, some considerations about testing and validation of the approach will be presented, as well as some concluding remarks.

2. Description of the problem

The spare parts distribution network proposed by the aircraft manufacturer for the new line of executive jets has five basic elements:

- (1) part suppliers;
- (2) part repair shops;
- (3) central warehouses located at distribution centers (DC);
- (4) regional warehouses located at the aircraft maintenance or service centers (SC);
and
- (5) locations where the demand for parts occur, which may be at service centers, or at airfields and airports where executive jets operate.

Figure 1 shows a hypothetical set of locations for the continental USA.

At any given time, parts ordered from suppliers will be shipped to the DC. In general, parts need to be certified for quality before being placed in inventory, and thus, assuming that certification will be economically viable only at the DC's, every part entering the system needs to go through a DC first. Parts have different production lead times, but in the aircraft industry some of these times may be quite long (like six or more months). This constraint highlights the importance of safety stocks.

When inventories at the service centers need to be replenished, parts will be shipped from the DC.

The failure of aircraft components may be detected either in the context of scheduled maintenance/inspections at the service centers, or in the course of pre-post flight checks at the airfields and airports.

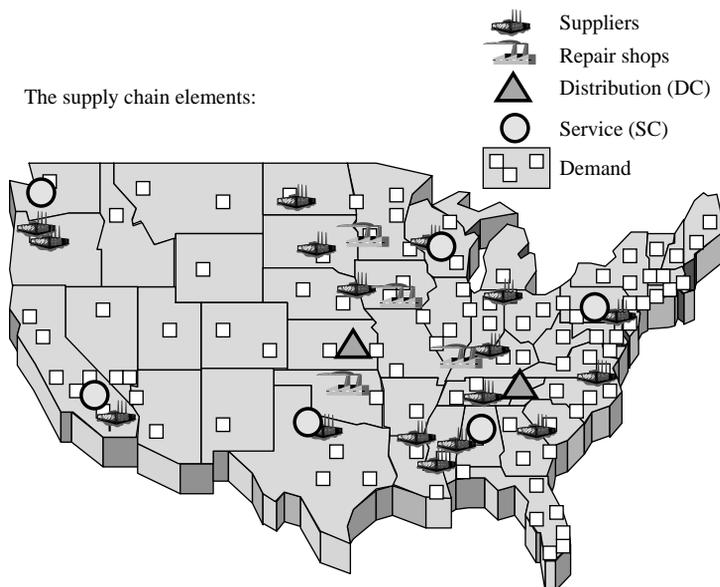


Figure 1.

When the demand for a part occurs at the site of a service center, the first recourse is to the inventory at that site. If that does not work, then in general the part will be shipped from a distribution center. Occasionally, it may be shipped from another service center. When the failed component is critical but the aircraft may still fly securely (for example, there is an alternate system in the aircraft), then the aircraft will fly to a service center as early as possible and the part will be replaced therein.

If, however, the failed component is such that the aircraft cannot fly until the part is replaced, then the failure is deemed of type AOG and the replacement part needs to be taken to the aircraft location as soon as possible (usually within a few hours), under penalty of severe contractual fines. In this case, the part may be transported with the mechanic who will change it, from the inventory at the service center from where he or she originates, or it may be shipped from another service center or a distribution center.

After replacement, if the part is an expendable, that is, a part that cannot be repaired or refurbished, then it will be discarded. If, however, it is a rotatable, that is, a repairable part, then it will be shipped to a repair shop, usually by way of a distribution or service center.

Figure 2 shows the logistics network resulting from the operations just described. The transportation services in the several links of this network will be provided or arranged by an external logistics company hired to manage the network and the warehouses. Depending on the specific link in the network, expedited modes of transportation may be needed at higher costs. Note, though, that the aircraft manufacturer will retain the responsibility and the ownership of the inventory control and purchase tasks.

3. An ADP model

3.1 Basic mathematical model

The fundamental problem consists of solving a very general class of multistage, stochastic dynamic programs. This model can be presented in the context of dynamic asset allocation using some basic notation.

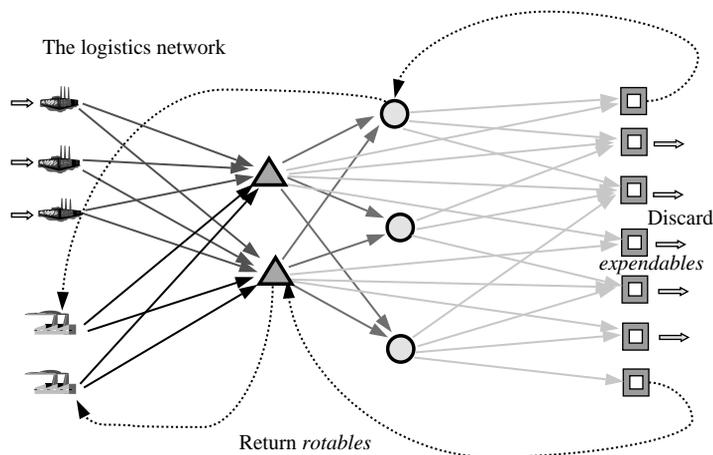


Figure 2.

Let:

a = The list of attributes of a part or component
(location, type, weight, supplier, production lead time,
age, and possibly repair shop).

R_{ta} = The number of components with attribute a at time t .

$\mathbf{R}_t = (R_{sta})_{a \in A}$ = The vector with the number of components
for each attribute in set A at time t .

This representation makes it fairly easy to add attributes as needed. The vector \mathbf{R}_t captures the state of all the inventories being managed. In addition, a modeling framework is needed where parameters such as failure rates are uncertain and are changing over time as new information arrives. All other forms of “knowledge” are represented using κ_t , and all information is captured using the state variable S_t , which means that the state of the system (the state of everything known about the system) is given by:

$$\mathbf{S}_t = (\mathbf{R}_t, \kappa_t)$$

Thus, \mathbf{R}_t tells about all the inventories of parts, while κ_t will capture other information relevant to decisions, including how many aircraft are in service, the age of each aircraft (which gives information about the likelihood of a failure), and the average lifetime of each part (which will change as new information becomes available).

Decisions are modeled using:

$d \in D$: A type of decision that can act on a component
(buy, move to a location, replace a part).

x_{tad} = The number of components with attribute a
that are acted on using a decision of type d at time t
(this is how purchase or move quantities are specified).

$\mathbf{x}_t = (x_{tad})_{a \in A, d \in D}$
= The vector of decision variables at time t .

A net contribution function to be maximized is associated with a set of decisions:

$C(S_t, \mathbf{x}_t)$ = Net contribution from choosing decision vector \mathbf{x}_t given the state vector \mathbf{S}_t .

The system evolves over time as new information arrives. The following generic notation is used:

\mathbf{W}_t = A vector of information that arrives during time interval t .

\mathbf{W}_t can include new demands, equipment failures, shipping delays, or delays in the time required to refurbish a component. Better information, or more reliable equipment, is modeled by changing the evolution of \mathbf{W}_t . The system evolves over time according to a transition function:

$$\mathbf{S}_{t+1} = S^M(\mathbf{S}_t, \mathbf{x}_t, \mathbf{W}_{t+1}).$$

Within this single function are included all the equations describing how inventories change, the possibility of a component “failing” (or being stolen) even while it is in inventory, new customer demands, and updates to estimates of parameters such as the expected lifetime of a part. Virtually, all of the engineering details (the physics of the problem) are captured in this single function.

Decisions are made using a decision function which is denoted by:

$X_t^\pi(\mathbf{S}_t)$ = A member of a set of decision functions ($\pi \in \Pi$) that returns a decision vector \mathbf{x}_t given the available information \mathbf{S}_t .

The problem then becomes to find the best decision function out of the set $(X_t^\pi(\mathbf{S}_t))_{\pi \in \Pi}$ that solves:

$$\max_{\pi \in \Pi} E \left\{ \sum_{t=0}^T C(\mathbf{S}_t, X_t^\pi(\mathbf{S}_t)) \right\}.$$

This problem class is exceptionally difficult. For discrete resource allocation problems, even deterministic versions of this problem can be computationally intractable. To handle the uncertainty that is inherent in this problem class, one needs to solve the stochastic version. The challenge is designing computationally tractable decision functions $X_t^\pi(\mathbf{S}_t)$ that solve this problem.

This research addresses two challenges: the computational one just mentioned and the challenge of designing policies for ordering and storing components to produce a robust resource allocation strategy that meets specified service goals and budget constraints. Standard inventory policies of the (s, S) type (bring the inventory to amount S when it falls below s), which are known to be optimal in special cases, are not optimal, however, in the presence of changing demands and demand rates, and they do not handle the ability to serve a demand in different ways (different locations, rush orders, using refurbished parts). A much more general strategy, based on ADP, is described next.

3.2 Algorithmic strategy

The research reported in this paper uses a new strategy for formulating dynamic programs (based on a concept known as the post decision state vector) to solve problems that have the general form:

$$x_t^n = \arg \max \left(C(\mathbf{S}_t, \mathbf{x}_t) + \bar{V}_t^{n-1}(R_t^x) \right) \quad (1)$$

where R_t^x is known as the “post decision state variable” (literally, the resource state variable immediately after the decisions just made have been taken into consideration). This is governed by the transition function:

$$R_t^x = R^{M,x}(\mathbf{R}_t, \mathbf{x}_t).$$

$\bar{V}_t^{n-1}(R_t^x)$ is an approximation of the value of being put in state R_t^x starting from state \mathbf{R}_t . Using $\bar{V}_t^{n-1}(R_t^x) = 0$ is equivalent to having a standard myopic strategy where the impact of decisions on the future is ignored. The challenge now becomes to design a good approximation for $\bar{V}_t^{n-1}(R_t^x)$ that is computationally tractable and yet provides good solution to the optimization problem in equation (1).

Assuming that $\bar{V}_t^{j-1}(R_t^x)$ is separable on the elements of R_t^x , then, for the problem in question, each of the terms of this value function will represent the marginal value of having a certain amount of a given component available in inventory at a given location. For instance, one term could represent the marginal value of right landing gears stored at a DC, as a function of the total number of units available there at a given time. Assuming further that the marginal value of the last component to arrive in inventory is no greater than the value of the previous one, then concave, piecewise linear functions can be used to approximate each of the separable elements of $\bar{V}_t^{j-1}(R_t^x)$. Under these assumptions, problem (1) becomes a linear program, easily solvable by available commercial packages.

Figure 3 shows a network flow representation of the linear program corresponding to problem (1) for a single aircraft component. Note that though the set of constraints of problem (1) is separable by single components, the actual, final objective function, to be discussed in the next subsection, is not. As part of the ADP strategy to solve this problem, the planning horizon (of, say, one year) is also decomposed into smaller time intervals (of, say, two days each), whose subproblems are then solved in chronological sequence. Hence, the LP shown in Figure 3 represents the subproblem to be solved for a given component in any one of these time intervals. The flows assigned to any decisions (arcs) that result in inventory parts stored at a DC or SC (like, for instance, holding parts at a warehouse or relocating parts from a DC to an SC) have to flow also through the appropriate value function arcs (the parallel arcs connecting the warehouse nodes to the super sink).

Note further that the decision to purchase more units of the component in question (the arc at the top of the network, joining the supplier to a DC) also flows into a set of value function approximation arcs. This value function has a special meaning. It provides the marginal value of available units of this component in the whole system, rather than at any given warehouse.

Each of the decision arcs in the network of Figure 3, that is the arcs outbound from the nodes representing suppliers, warehouses and repair shops (left-hand side of the figure), may have a contribution, a cost or both associated to it. The contributions may represent actual rewards incurred through the decision (for example, when serving

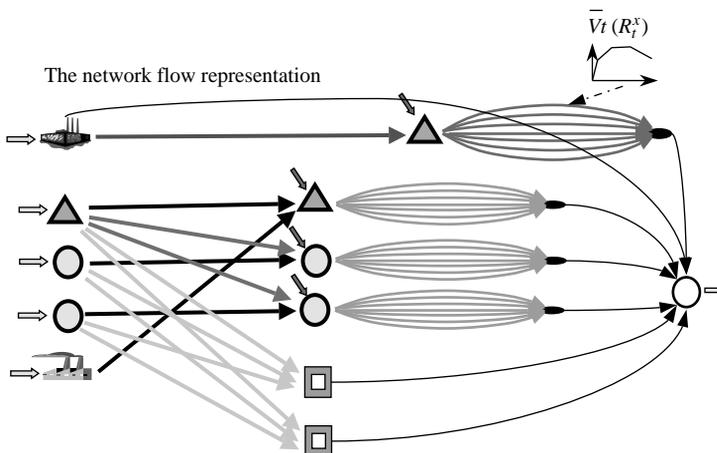


Figure 3.

demand) or artificial bonuses designed to encourage demand satisfaction. The costs may represent actual transportation, purchase or inventory costs, or penalties intended to discourage certain behaviors.

3.3 Final mathematical model

The primary goal of inventory control systems is to determine the stock policy that minimizes transportation and storage costs, while serving the maximum demand possible. An interesting trade-off arises when there are constraints on the amount of capital invested in inventory. The question then becomes that of finding an inventory policy that minimizes costs while keeping the money invested below a maximum target and the demand satisfied above a minimum target, both of them aggregated over all components in the system and over the whole planning horizon. One of the appealing aspects of this proposition is the possibility of trading off a lower LOS (fill rate) of an expensive, long-lasting component, for higher levels of service of less expensive, more often-failing components.

In order to achieve the above-mentioned composite goal, two extra terms will be added to the objective function of the ADP model outlined in the previous subsection.

Let:

$x_{\tau a}^p$: no. of parts with attribute a purchased at time τ

p_a : purchase cost of part with attribute a

B : maximum capital to be invested over the planning horizon

$b(x) = \sum_{\tau} \sum_a p_a x_{\tau a}^p$: actual capital invested in all parts over the horizon.

The objective function of problem (1) can now be modified as follows:

$$\max_x (C(\mathbf{S}_t, \mathbf{x}_t) + \bar{V}_t(R_t^x) - \theta \times H_1(b(x) - B))$$

where $H_1(\dots)$ is a convex, nondecreasing function of its argument. Note that, the added term exerts a negative pressure on the inventory levels. What is missing is another term to push the inventory levels up towards a solution in which a minimum target of aggregate fill rate for each type of LOS is reached (recall that there are three levels of service of demand: AOG, critical and routine).

Now let:

$x_{\tau a^j}^r$: actual no. of parts with attribute a^j (replaced under LOS j) at time τ

$R_{\tau a^j}^r$: total no. of parts with attribute a^j (that need replacement under LOS j) at time τ

F_j : minimum desired fill – rate under LOS j over the planning horizon

$f_j(x) = \left(\frac{1}{\sum_{\tau} \sum_{a^j} R_{\tau a^j}^r} \right) \sum_{\tau} \sum_{a^j} x_{\tau a^j}^r$: actual fill – rate achieved under LOS j

over the horizon.

Adding another term to the objective function of equation (1), the final objective function becomes:

$$\max_x \begin{pmatrix} C(\mathbf{S}_t, \mathbf{x}_t) + \bar{V}_t(R_t^x) - \theta \times H_1(b(x) - B) \\ - \sum_j \gamma_j \times H_2(F_j - f_j(x)) \end{pmatrix} \quad (2)$$

where $H_2(\dots)$ is also a convex, nondecreasing function of its argument. The challenge, or the art, in solving the problem with objective function (2) is to find the right balance among the parameters θ and γ_j , and the net contribution terms.

It is important to note, though, that the desired target values for capital invested in inventory (B) and fill rates per LOS (F_j 's) have been aggregated not only over all the parts (summation over j), but also over all time intervals in the planning horizon (summation over τ). However, equation (2) is the objective function for each time subproblem t . This means that when solving the subproblem at time t with equation (2), decisions made in time intervals $\tau < t$ will be already known, but decisions to be made in time intervals $\tau > t$ will have to be approximated. In practice, a Gauss-Seidel iterative approach has proven to work very well.

4. Estimating value functions

Unless there is some exogenous knowledge that allows for the value functions $\bar{V}_t^{n-1}(R_t^x)$ to be determined beforehand, they will have to be computed iteratively, using dual values or numerical derivatives estimated through the repeated solution of the mathematical programs that compose the subproblems. In this section, an overview of the estimation procedure for the problem in question will be provided.

Using the assumptions stated in Subsection 3.2, namely that a value function is separable in the elements of R_t^x and that each term of the value function is a concave, piecewise linear function of the number of available units of that element, let \bar{v}_{ta}^{n-1} represent the value function associated to attribute vector \mathbf{a} at time t used in the solution of subproblems in iteration n of the iterative calibration procedure. Now, let \hat{v}_{ta}^n represent a dual value (or numerical derivative) of the attribute vector \mathbf{a} at time t obtained through the solution of a slightly modified subproblem in iteration n .

Figure 4 shows the network corresponding to the slightly modified subproblem for a single component. Value functions for a single component need to be estimated for

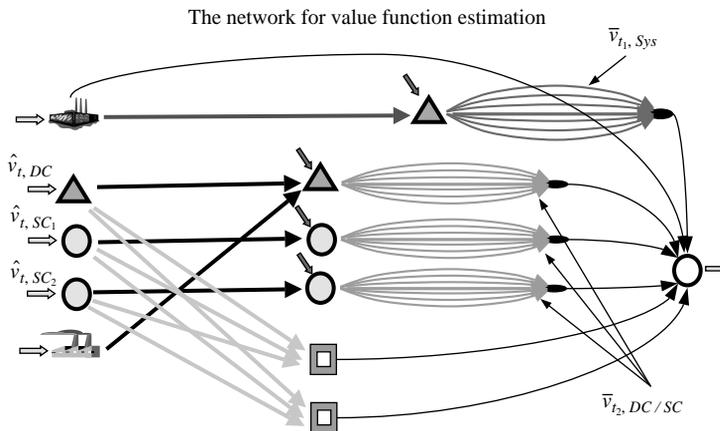


Figure 4.

each of the warehouses, including the central ones (the DC's). However, replenishment of the inventories at the regional warehouses (the SC's) happens only from the DC's. This means that the assumption of separability between the value function for a given component at a DC and the one at an SC is actually too strong. In order to comply with this assumption, though, when estimating the dual values \hat{v}_{ia}^n , the linear program needs to be modified, namely, the inventory replenishment arcs connecting DC's to SC's need to be removed (compare the network in Figure 4 with that in Figure 3).

Once the dual values are known, the update of the value functions themselves is done through basic exponential smoothing of the type:

$$\bar{v}_{ia}^n = (1 - \alpha_{n-1})\bar{v}_{ia}^{n-1} + \alpha_{n-1}\hat{v}_{ia}^n,$$

where α_{n-1} is the smoothing stepsize. The actual details of the choice and update of the stepsizes and the mechanics of updating piecewise linear functions, while maintaining concavity, are beyond the scope of this text, but the interested reader can refer to Godfrey and Powell (2002), George and Powell (2006) and Powell (2007) for them.

The value function representing the marginal value of a part for the whole system at a given time ($\bar{v}_{t,\text{Sys}}^n$) is updated using the pertinent value functions from each of the warehouses (\bar{v}_{ia}^n). The concave, piecewise linear function is generated by properly assembling in decreasing order all the slopes of the corresponding value functions at all the warehouses. In summary, at iteration n the \hat{v}_{ia}^n 's are used to update the \bar{v}_{ia}^n 's, and these are used to assemble $\bar{v}_{t,\text{Sys}}^n$.

4.1 The challenge of rare events

There is, however, an issue in the update of the value functions that is specific to this problem. Some aircraft parts fail very infrequently. This fact poses a statistical challenge to the iterative update of value functions as described in this paper. The estimation of the dual values \hat{v}_{ia}^n is done through the solution of the subproblems, which are in turn generated based on demand realizations (part failures) randomly sampled. If, for a given part, the number of observed failures in a sample path over the planning horizon is too low (like, for instance, one failure every couple of years!), then the \hat{v}_{ia}^n obtained will have a very low content of interesting information, namely the marginal value of a part at a warehouse will be determined solely by value of holding that part in inventory at that location! This would decrease dramatically the rate of convergence of the estimation of value functions for the part in question. An approximate solution to this issue is proposed in the next subsection.

4.2 Conditional marginal values

In order to increase the informational content of the dual values computed at each iteration, a conditional marginal value will be computed. The conditioning will be done with two simple events: the event of observing no failures of a part during the time interval of the subproblem in question; and the complement event to the first one, that is, the event of observing at least one failure over that time interval.

Suppose that for a given subproblem at time t , no failures of the part in question were observed in the sample path for that iteration. Then the network representation for this subproblem would look like the one in Figure 5. The dual values computed through this subproblem would correspond to the marginal values computed when no

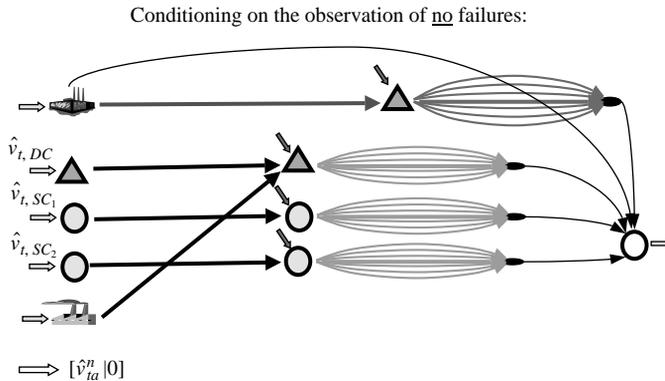


Figure 5.

replacements of part **a** are observed at time t . Call these conditional dual values $[\hat{v}_{ta}^n | 0]$. In order to compute the marginal values for the event of at least one failure of part **a** at time t , a single randomly generated part failure will be temporarily added to the network in Figure 5. The conditional dual values computed through the resulting linear program will be referred to as $[\hat{v}_{ta}^n | 1+]$.

Now, suppose instead that for the subproblem at time t , some failures of part **a** were actually observed in the random sample path for that iteration. For instance, take the case of Figure 6, where two failures are present. The dual values computed through the network in that figure would correspond to $[\hat{v}_{ta}^n | 1+]$. In this case, in order to compute $[\hat{v}_{ta}^n | 0]$, one would simply temporarily remove from the network the two failure nodes, and the respective inbound and outbound arcs, and solve the resulting linear program.

Assuming that the probabilities associated to these two complementary events can be explicitly determined, the equation used to compute the final dual value \hat{v}_{ta}^n is:

$$\hat{v}_{ta}^n = p_{ta,0} \times [\hat{v}_{ta}^n | 0] + p_{ta,1+} \times [\hat{v}_{ta}^n | 1+]$$

where:

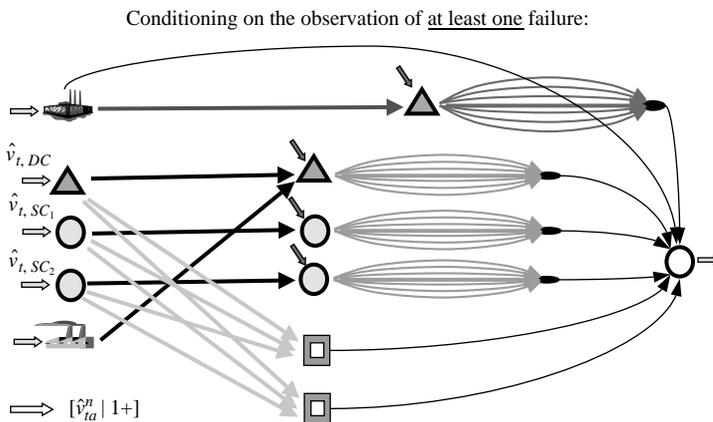


Figure 6.

- $p_{ta,0}$ = probability of NO replacement of part a during time interval t .
- $[\hat{v}_{ta}^n|0]$ = marginal value computed in a network where there is no replacement of part a during time interval t .
- $p_{ta,1+}$ = probability of AT LEAST ONE replacement of part a during time interval t .
- $[\hat{v}_{ta}^n|1+]$ = marginal value computed in a network where there is at least one replacement of part a during time interval t .

5. Concluding remarks

The main contribution of this paper is to present a novel, ADP-based approach to solving the problem of determining the best inventory policy for a distributed warehousing network, in the context of an aircraft spare parts distribution network. Complicating issues are the presence of high value, low-frequency demand parts and of parts that can be refurbished after replacement and then returned to inventory.

Though the model and solution strategy proposed in this paper have not yet been extensively tested against other models and approaches, they have been tested and validated internally at an aircraft manufacturer through the expert analysis of the policy recommendations in different settings and for different pools of aircraft parts. So far, the results seem promising and, given the absence of alternative methodologies, compelling.

As a brief illustration of the outcomes produced by the system, consider the yearly scenario of a pool of 478 rotatable parts being used by over 650 aircraft. Using a distribution network composed of one DC and 18 SC's spread over the continental USA, and setting the target value for the capital invested over a year to \$50 million and the fill-rate targets for AOG and critical failures to 98 percent, Figures 7-9 show some numerical results.

Figure 7 shows the progression of the actual yearly capital invested in parts as the iterations for the calibration of the value functions evolve. As the value functions learn, the value of $b(x)$ approaches the desired target value.

Figures 8 and 9 show the histograms of the actual fill-rates achieved for AOG and critical demand over 19 different demand sample paths. Clearly, on average, the solutions are very close to the desired targets.

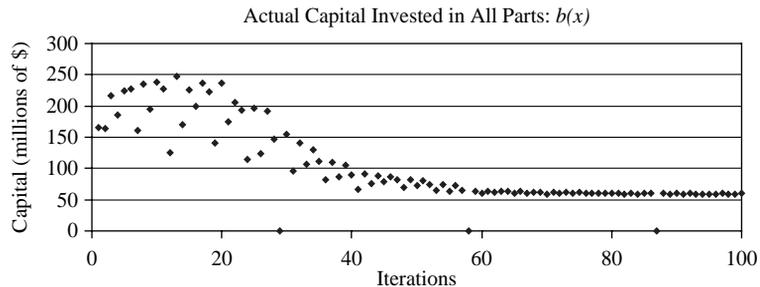


Figure 7.

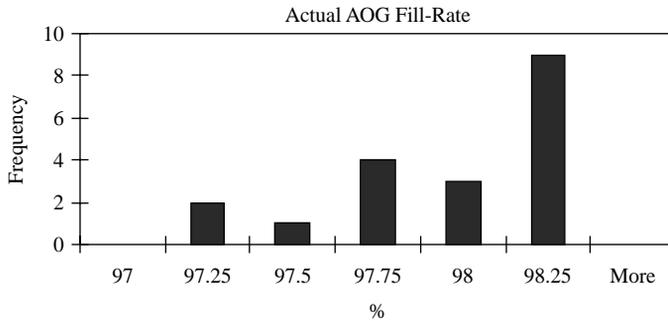


Figure 8.

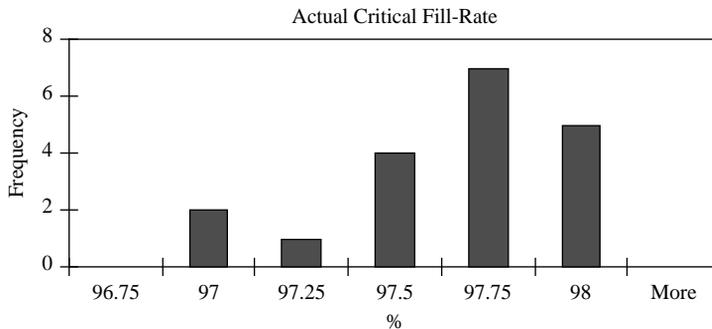


Figure 9.

Some challenges, though, remain. Given that embedded in the proposed ADP methodology there is a simulation of the failure processes of all parts being used in a fleet, computational times are very sensitive to the magnitude of the total number of parts and aircraft in use in a given year. For instance, each iteration of the value function calibration run whose results are shown in Figures 7-9 took, on average, 65 minutes to run on a 2.4 GHz AMD Opteron Processor 250, running Linux and using CPLEX 10.2.

The accuracy of the logistics network data is also relevant. The allocation of inventory to regional warehouses depends on the level and ease of accessibility of airfields and airports from those service centers, thus increasing the sensitivity of distributed warehousing solutions to the quality of those data.

Overall, however, ADP seems to offer a viable, and promising, way to solving dynamic, geographically distributed inventory problems, where there are wide ranges in the values of the stored products, in the frequency of their demands, and where some used products can be refurbished and returned to inventory.

References

- Cao, D. and Silver, E.A. (2005), "A dynamic allocation heuristic for centralized safety stock", *Naval Research Logistics*, Vol. 52, pp. 513-26.
- George, A. and Powell, W.B. (2006), "Adaptive stepsizes for recursive estimation with applications in dynamic programming", *Machine Learning*, Vol. 65, pp. 167-98.

Godfrey, G. and Powell, W.B. (2002), "An adaptive dynamic programming algorithm for single period fleet management problems I: single period travel times", *Transportation Science*, Vol. 36, pp. 21-39.

Powell, W.B. (2007), *Approximate Dynamic Programming*, Wiley, New York, NY.

Powell, W.B., Bouzaienne-Ayari, B. and Simão, H.P. (2007), "Dynamic models for freight transportation", in Barnhart, C. and Laporte, G. (Eds), *Handbooks in Operations Research and Management Science: Transportation*, Vol. 14, pp. 285-365.

Corresponding author

Hugo Simao can be contacted at: hpsimao@princeton.edu