Optimal Procurement with Demand Warning

Greg Calbert [a] and Rajesh Thiagarajan [a]

[a] *Planning and Logistics Discipline, Joint Operations Analysis Division, Defence Science and Technology Organisation*
Email: Greg.Calbert@dsto.defence.gov.au, Rajesh.Thiagarajan@dsto.defence.gov.au

Abstract: In the field of inventory management, substantial research has been conducted on the formulation of both prediction and procurement algorithms to meet lumpy unforeseen demands. Often these approaches, based on re-order point methods fall short, generating stock outs when procurement lead times are long or order quantities are constrained. Fortunately, for certain circumstances in the fields of military and commercial planning, there are warnings associated with demand spikes or lumps which can be used in the procurement plan.

In this paper we formulate a dynamic programming approach for the optimal procurement of stock, given warning of a demand spikes based on Advanced Demand Information. The dynamic programs formulated are based on direct rewards, both during the prior warning period and during the demand spike. During the warning period, moderate negative rewards for the procurement of stock are given to avoid the storage of unnecessary stock holdings and during the demand spike, large negative rewards are given to stock outs. The dynamic program model assumes orders take place only after receipt of goods, that is, production is inherently constrained, so multiple orders cannot take place within a single lead time. Consideration of capacity constraints is included in the dynamic program formulation.

Dynamic programs that break this production constraint assumption are formulated, allowing orders between the receipts of goods. This results in a large increase in program dimensionality which is estimated. We discuss extensions to the problem, including stochastic baseline demands, and combining re-order point methods with a dynamic programming 'meta-controller' to ensure inventory sustainability.

Keywords: *Logistics planning, dynamic programming, inventory policy*

1. INTRODUCTION

This paper examines inventory control policies in the context of military operations. Most literature on the control of inventory focuses on inventory procurement policies driven by commercial constraints, such as minimising capital stockholdings, avoiding stock outs and minimising overall costs (Porteus, 2002). Some modern inventory control literature has focussed on the incorporation of advanced demand information, that is, time series information of future demands (as opposed to forecasts of future demands) in developing procurement information (Ozer, et al, 2004).

In military logistics such advanced demand information may be available, given that military operations demands are planned, through rate of effort models of demand. Furthermore, a critical parameter influencing inventory management in military operations is the warning time. Some operations, such as disaster relief, have minimal warning time, whilst other operations may have moderate warning times and logistics requirements may be anticipated (Stevens and Ingram, 2013).

For this paper, we develop a dynamic program that controls inventory procurement decisions, assuming that a warning time is available for the procurement of stocks. It is assumed that advanced demand information is also available. Cost is not assumed to be a direct factor in the current modelling. Instead, the reward structure of the dynamic program reflects two factors. First, prior to a military operation no 'unnecessary' stocks should be held, as they are costly in terms of purchase, holding costs and perishability. Second, during the military operation, it is critical that minimal stock outs occur because of the importance of supply.

Following this introduction, we develop the states, transitions and reward structure of the dynamic program (Section 2). Results of the dynamic program are simulated in Section 3. Section 4 briefly describes extensions of this approach, including incorporation of within lead time, unconstrained production orders and re-order based approaches to inventory management. Section 5 provides a discussion and conclusions.

2. THE PROBLEM FORMULATION

In the formulation of this dynamic program, we assume that stocks can be delivered every lead time and not continuously, the later being the case for re-order point methods (Blumenfeld, 2001). This assumption may be removed, at the cost of a considerable increase in the state-space dimensions of the problem, which we discuss in Section 4. Suppose that these delivery or lead-time periods are labelled as $n = 1, \cdots, N$, where N is the terminal time of the problem, assumed to be the end of the military operation. It should be emphasised that the variable n is not the real time of the system. Rather, this variable specifies the intervals of time length L, where L is the lead time of delivery for a stock type. In this problem formulation, stock orders are not allowed in between deliveries, as is the case for normal reorder point methods, because production is constrained. Thus, if the warning time is WT time units and the scenario sustainment length is Sp time units then there are at most $N = \left\lceil \left(\frac{WT + Sp}{L} \right) - 1 \right\rceil$ deliveries in total.

In order to formulate the dynamic program, to determine the optimal procurement policy, the states of the problem are as follows.

- Let $\mathrm{X} \in \{\mathrm{B}, C\}$ specify the 'scenario state' of the system. Here, B specifies the state of the system before the start of the contingency. That is, a warning of a military scenario occurring has been issued, however, this scenario has not commenced. Let C specify the state of the system during the scenario.
- Let I_n specify the inventory level at period n and D_n the demand during period n to $n+1$.
- Let Q_n be the order placed at time $n \le N$, to be received at the next lead-time period, $n+1$.

The scenario states, inventory levels and order quantities are illustrated in the following figure 1.

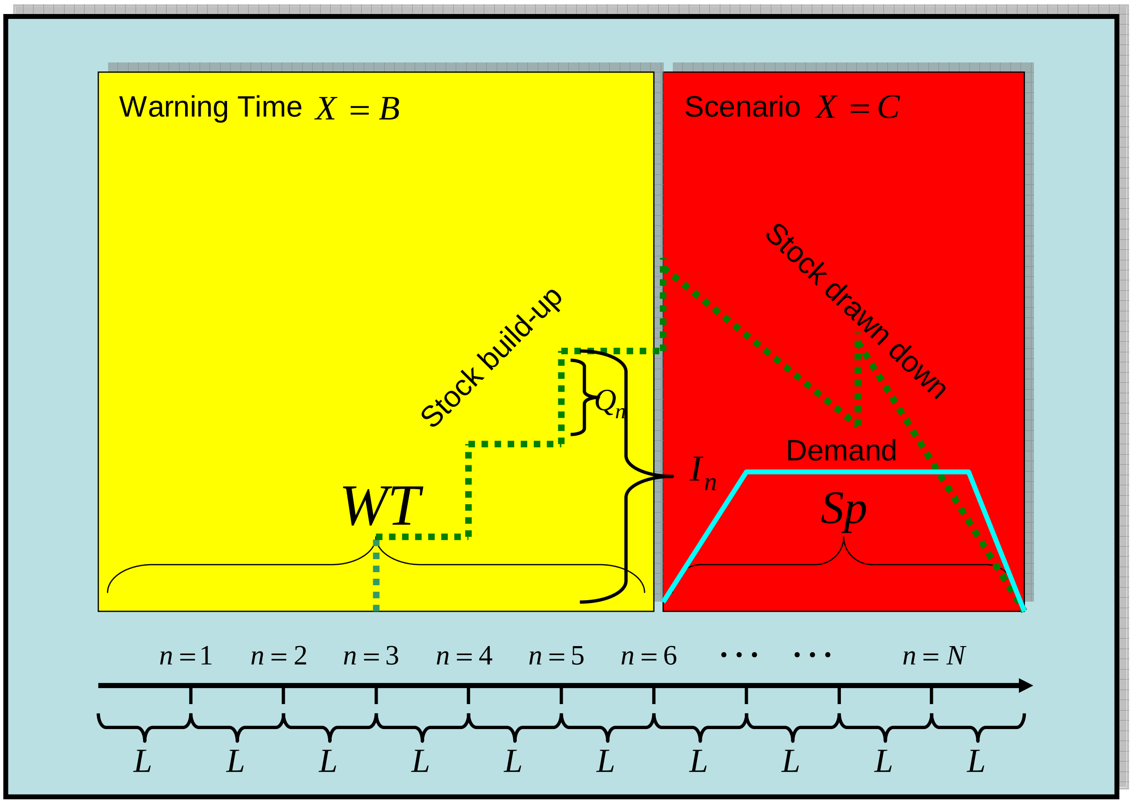

Figure 1: Problem formulation for the dynamic program and state variables.

2.1. The dynamic program

A reward function is specified for each time period that is dependent on the current inventory level and the scenario state, $\phi : \mathbb{R} \times X \to \mathbb{R}$ defined by

$$\phi(I_n, \{B, C\}) = \begin{cases} -k_{\mathrm{B}} I_n, I_n > 0, X = B, \\ 0, I_n \le 0, X = B, \\ k_c I_n, I_n < 0, X = C, \\ 0, I_n \ge 0, X = C. \end{cases} \tag{1}$$

It is assumed that $k_C >> k_{\mathrm{B}}$, $k_C, k_{\mathrm{B}} \in \mathbb{R}$ reflecting the requirement that the reward for having negative stock levels during the scenario is significantly greater than the stock holding costs prior to the scenario.

Before formulating the dynamic program we must also define the function, $f : \mathbb{Z} \to X$ that specifies the time or period under which a transition to the occurrence of the scenario takes place, which is simply defined by

$$f(n) = \begin{cases} \mathrm{B}, n < WT, \\ C, n \ge WT. \end{cases} \tag{2}$$

Given these definitions, it is now possible to define a value-function for the dynamic program, $\mathbf{V} : \mathbb{R} \times S \to \mathbb{R}$ as the undiscounted sum

$$\mathbf{V}(I_o, \mathrm{B}) = \sum_{n=1}^{N} \phi(I_n, f(n)). \tag{3}$$

The optimal value function can be found using Bellman's equation (Bellman, 1954), which is

$$\mathbf{V}(I_n, f(n)) = \max_{0 \le Q_n \le MaxOQ} \left(\phi(I_n, f(n)) + \mathbf{V}\left(I_{n+1}, f(n+1)\right) \right), \tag{4}$$

with the inventory dynamics $I_{n+1} = I_n + Q_n - D_n$ and maximum order quantity $MaxOQ$.

We also impose a terminal reward penalty,

$$\mathbf{V}(I_N, C) = -k_N \left| I_N \right|, \tag{5}$$

where $k_N \in \mathbb{R}$ is a small positive constant. This reward function ensures inventory levels approach zero at the end of the contingency. Capacity constraints can also be factored in. If Cap is the maximum capacity, then

$$\mathbf{V}(Cap, f(n)) = -k_{Cap}, \tag{6}$$

where $k_{Cap} \in \mathbb{R}$ is a large positive number.

The optimal policy $Q_n^* = \pi^*(I_n, f(n))$ was solved by backwards iteration of the dynamic program (Mangel and Clark, 1989). Simulations are then used to calculate the behaviour of the policy.

3. RESULTS AND SIMULATIONS

The behaviour of the dynamic program is assessed by generating a random demand, over the sustainment period. Figure 2 shows the inventory level for a demand of 100 units of stock, a warning time prior to the scenario of 40 time units, a scenario length of 20 time units. The maximum inventory capacity was assumed to be 100 units.

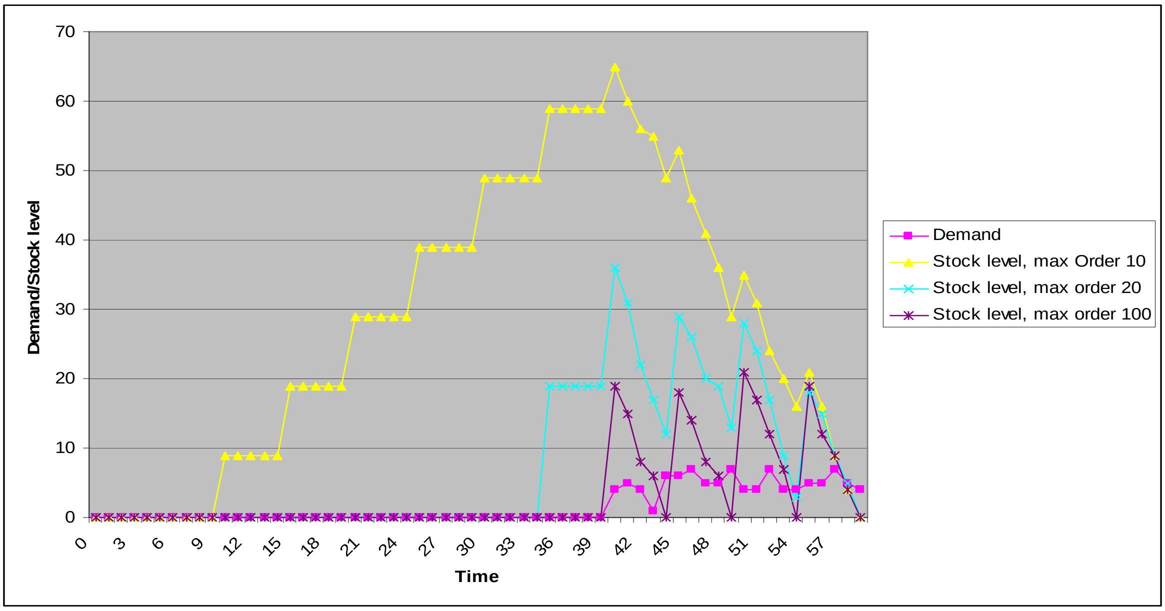

Figure 2: Demand and stock level, assuming maximum order quantities of 10, 20 or 100 units.

Figure 2, shows that the policy exhibits stock build-up behaviour, if the maximum order quantity is below that of the demand. This is seen when the maximum order quantity is 10 units, where procurement commences at time 10, 30 units of time before the commencement of the scenario. When the maximum order quantity is 100 units, a single order of 100 units will not be considered the optimal policy, as there is no extra reward in ordering over and above what is required to keep the stock level dynamics at positive levels.

Figure 3 shows the policy of the dynamic program, denoted as a contour map over the inventory level and the delivery number of order period, denoted by $\pi^*(I_n, f(n))$.

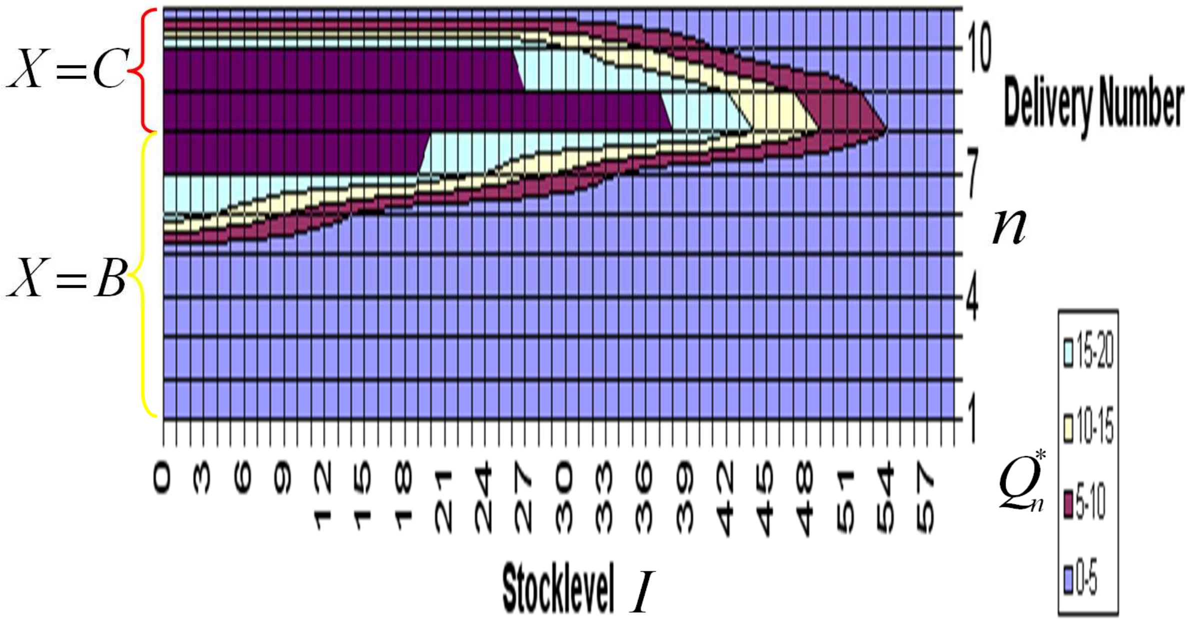

Figure 3: The policy of the dynamic program (optimal order quantity Q_n^*), as a contour map. The horizontal axis specifies the inventory level, I and the vertical axis, the delivery period, n. Colours in the map indicate the order quantity that is the policy. The parameters for this simulation were $L = 5, WT = 40, Sp = 20,$ and total demand is 100 units. The maximum order quantity, $MaxOQ$ was 20 units.

The structure of the inventory policy is as follows. Prior to the commencement of the scenario, $X = B$, (where the delivery period is between 1 and 7) only small orders are made when the inventory level is large. Substantial orders are made when the current inventory level is low and we are in the scenario sustainment period. If the inventory level is high and we approach the terminal delivery period, no further orders are made, to ensure minimal stock levels at the terminal period, to satisfy equation (5).

4. EXTENSIONS TO INCLUDE WITHIN LEAD TIME ORDERS AND RE-ORDER POINTS

There are various extensions of this simple model to include more complexity. One such extension is to assume that orders can take place, not just after the receipt of a previous order, but at any time during the decision process. The complexity of the dynamic program increases markedly, because both the inventory dynamics and the state of past orders have to be included in the state description. The value function is now specified as $\mathbf{V}(t,[t_1,Q_1,t_2,Q_2,\cdots,t_k,Q_k],I_t,f(t))$ were the pairs (t_j,Q_j) specify the j^{th} order's time of delivery and order quantity. If we assume that only one order can be placed per unit time, then $0 \le t_1 < t_2 < \cdots < t_k \le L.$ The dynamic program will then take the form

$$\mathbf{V}(t,[t_1,Q_1,t_2,Q_2,\cdots,t_k,Q_k],I_t,f(t)) = \max_Q \left(\phi(I_t,f(t)) + \begin{cases} \mathbf{V}(t+1,[t_2-1,Q_2,t_3-1,Q_3,\cdots,L,Q],I_t+Q_1-D_t,f(t+1)), t_1=0 \\ \mathbf{V}(t+1,[t_1-1,Q_1,t_2-1,Q_2,\cdots,t_k-1,Q_k,L,Q],I_t-D_t,f(t+1)), t_1>0 \end{cases} \right). \quad (7)$$

The dynamic program specified in equation (4) requires approximately $2\left(\left\lceil\left(\frac{WT+Sp}{L}\right)-1\right\rceil\right)Cap$ variables for the value function. The dynamic program specified by equation (7) is of complexity $2(W+Sp)Cap(MaxOQ)^L L^L$ variables for the value function, where $MaxOQ$ is the maximum order quantity. It should be clear that the requirement to store the timing of all forthcoming orders in equation (7) means that the dynamic program does not scale well. For example, with $W + Sp = 100, Cap = 100, \max OQ = 20, L = 10,$ equation (4) requires approximately 10^3 variables whilst equation (7) requires approximately 10^{14} variables for the value function, which is clearly not computationally feasible.

Another extension of this dynamic program is to include re-order point methods in the formulation. Here, demand is stochastic, and hence the optimisation method is based on a Markov or semi Markov decision process (MDP). The MDP will be considered as a 'Meta inventory controller' placed over traditional prediction based inventory controller, which uses re-order point methods. As with multiple lead time orders, there will be a requirement for an expansion of the state description, to include not only inventory levels and the scenario state, but also the re-order controller's prediction of the demand over that inventory period (the prediction of demand must be calculated as demand is now stochastic) and transitions between demand predictions. The prediction in demand over period n is defined as $\bar{D}_n$ and Q^{ROP} is the re-order quantity (possibly the economic order quantity) if the inventory level drops below the re-order point, the Markov decision process looks like

$$\mathbf{V}(I_n, f(n), \bar{D}_n) = \max_Q \left\{ \sum_{D_n=0} \Pr(D_n \mid \bar{D}_n) \left[\phi(I_n, f(n)) + \sum \Pr(\bar{D}_{n+1} \mid D_n) \mathbf{V}(I_{n+1}, f(n+1), \bar{D}_{n+1}) \right] \right\} \quad (8)$$

where

$$I_{n+1} = I_n + Q - D_n + \begin{cases} 0, & \bar{D}_n L + \sigma < I_n, \\ Q^{ROP}, & \bar{D}_n L + \sigma \geq I_n, \end{cases} \quad (9)$$

which is the standard re-order point equation with safety stock parameter, σ. Here, the conditional probabilities $\Pr(D_n \mid \bar{D}_n)$ and $P(\bar{D}_{n+1} \mid D_n)$ need to be defined. Solving the extensions in equations (7) and (8) is the subject of current and forthcoming research.

5. SUMMARY AND FUTURE WORK

Dynamic programming is the basis for most inventory calculations (Porteus, 2002). We have formulated a dynamic programming approach tailored to military supply chain circumstances and have shown its effectiveness in determining an optimal procurement policy using simulations. A brief account of some extensions to the base model was presented. Particularly, an approach to combine the base model with re-order point method, typically used in real military systems, was discussed. Generally, the dimensionality of such MDPs will be too large to be solved analytically; therefore, the next step in this research will be the applicable of dynamic program approximation techniques, such as Neuro-dynamic programming (van Roy, et al, 1997).

REFERENCES

Bellman, R. (1954) The Theory of Dynamic Programming, Rand Corporation.

Blumenfeld, D. (2001) *Operations Research Calculations Handbook*, CRC Press.

Mangel, M. and Clark C.W. (1989). *Dynamic Modelling in Behavioural Ecology*. Princeton University Press.

Ozer, O and Wei, W. (2004). Inventory control with Limited Capacity and Advanced Demand Information. *Operations Research*, 52 (6), 988-1000.

Porteus, E. L. (2002) *Foundations of Stochastic Inventory Theory*, Stanford University Press.

Stevens, A. and Ingram, H. J. (2013) Reframing the Defence Discourse: Australia's "Black Swan" and its implications for preparedness and mobilisation. *Australian Defence Force Journal* (190) 30-43.

Van Roy, B. Bertsekas, D., Lee, Y, and Tsitsiklis, J.N. (1997) A Neuro-Dynamic Approach to Retailer Inventory Management. http://web.mit.edu/jnt/www/Papers/R-96-bvr-retail.pdf