Inventory Management Using Receding Horizon Method

Carl Richardson

Abstract— **This report investigated the use of the Receding Horizon method to minimise the total expected operational cost for an inventory management problem. For a single prediction and control horizon pair, a variable sequence of stock updates was found which significantly reduced the total cost compared to the optimal fixed stock update found through Monte Carlo simulation.**

I. INTRODUCTION

The UN authorised the opening of a store on Mars, it's only purpose is to sell super apples. The challenge for the store manager is as follows: given the probability distribution of the weekly demand for super apples, an order must be placed at the end of each week if the stock level is less than or equal to three (determined by Monte Carlo investigation). The order number can vary from week to week. The store is open Monday-Friday and at the end of each week, the stock is counted and an order is placed. The new stock is delivered before the store reopens on Monday. The objective of this report is to use the Receding Horizon method (RHM) for the purpose of minimising the operational costs. First a prediction is made for what the optimal solution might look like; this is followed by a formulation of the RHM. Finally, the performance and optimal solution of a single prediction and control horizon pair are discussed before comparing the performance of a range of combinations.

II. OPTIMAL POLICY PREDICTION

For the first 50 weeks, only the two weekly operational costs (stock penalty and warehouse cost) must be considered. These differ by a factor of four, so the stock update should primarily ensure the stock penalty is minimised. The secondary consideration is to ensure the stock is not kept unnecessarily high so to increase the chance of not paying a warehouse fee. With this philosophy in mind, the distribution of the demand shows that $P(D \leq 4) = 0.96$ and the mean weekly demand is three. The results from the Monte Carlo investigation suggested the optimal threshold was less than or equal to three, so this will be used throughout these experiments. From this information, I would recommend to update the stock to four each time the level falls below the threshold. This leaves a 4% chance of having insufficient stock and a 16% chance of having the exact amount of stock. For the penultimate week, only the returns cost and stock penalty must be considered. Since these are expensive costs, I would suggest taking a bit more of a risk to try and avoid both. I would recommend updating the stock to three as this would give an 80% chance of having sufficient stock and 40% chance of having the exact amount.

III. RECEDING HORIZON FORMULATION

To use the RHM, a dynamic system must be defined according to equation 1. The next state, x_{t+1}, is a function of the current state and input, u_t. The system output, y_t, is another function of the current state and input. The dimension of each variable is only constrained by the problem and the functional relationships. This model provides the structure to simulate the system. In order to select an optimal input, a cost function must be defined in the form of equation 2. This is the sum of the stage costs, $l(x_t, u_t)$, from the current time point, t_0, over the prediction horizon, N_y. The cost function is optimised with respect to the set of inputs $\{u_{t_0}, ..., u_{t_0+N_y-1}\}$. The control horizon, N_u, determines how many unique control inputs of the cost function are optimised. If $N_u < N_y$, the set $\{u_{t_0+N_u+1}, ..., u_{t_0+N_y-1}\}$ will be repetitions of the control input $u_{t_0+N_u}$. After performing the optimisation step, the control input $u^*_{t_0}$ is applied to the system. Furthermore, this model is subject to constraints on the state, $x_t \in X$, and input, $u_t \in U$.

$$\begin{aligned} x_{t+1} &= f(x_t, u_t) \\ y_t &= h(x_t, u_t) \end{aligned} \tag{1}$$

$$J = \sum_{t=t_0}^{t_0+N_y-1} l(x_t, u_t) \tag{2}$$

For the inventory management problem, the stochastic system was defined by equation 3. The stock level and stock update, at the end of week t, were respectively represented by x_t and u_t. The demand for week t was represented by the discrete stationary random variable $d_t \backsim p(d)$. In terms of equation 1, d_t was considered as a second input modelling a measured disturbance. The cost function was defined by equation 4, it was the sum of the weekly expected operational costs over the prediction horizon. The operational costs were as follows. It cost 5 gold coins per weekend to store the super apples in a warehouse, w_t. A penalty of 20 gold coins was issued if the weekly demand was not met, p_t. On the last week of operation, 10 gold coins per apple was charged to return the remaining stock to Earth, r_t. Each of these costs was a function of the current state given by equation 3. The equations for the specific costs are defined in the appendix. Finally, the input was constrained to the set $u_t \in \{0, 1, 2, 3, 4, 5, 6\}$ and the state had the constraints $x_t \in Z$ and $x_t \geq 0$.

$$\begin{aligned} x_{t+1} &= x_t + u_t - d_{t+1} \\ y_t &= x_t \end{aligned} \tag{3}$$

$$J = \sum_{t=t_0}^{t_0+N_y-1} E_{p(d)}[w_t(x_t) + p_t(x_t) + r_t(x_t)] \tag{4}$$

IV. RESULTS

The first experiment simulated the stock dynamics, with prediction and control horizons both set to five. As with both experiments, the optimisation step was performed using Matlab's genetic algorithm as it provided the ability to optimise a piecewise function with integer constraints. Figure 1 shows the histogram of total costs, over 100 trials. The sample mean and standard deviation were 260.95 and 20.75 respectively. Figure 2 illustrates the stock level and stock updates, for a single trial; the corresponding total cost was 250 coins. The second experiment simulated the stock dynamics for a range of control horizons uniformly separated over each prediction horizon. The same realisations of the demand sequences for the stock and predicted stock dynamics were repeatedly used for each combination to ensure a fair comparison. Figure 3 illustrates the resulting distribution of total costs.

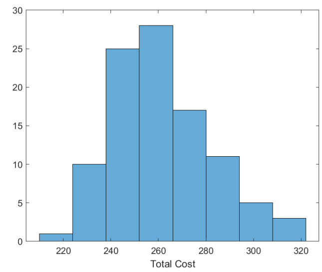

Fig. 1. Total costs, after 100 trials, for $(N_y, N_u) = (5,5)$

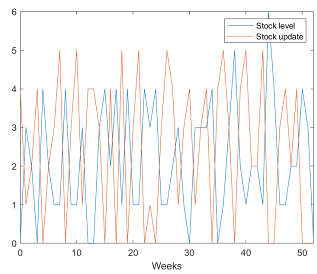

Fig. 2. Weekly stock level and stock update for $(N_y, N_u) = (5,5)$

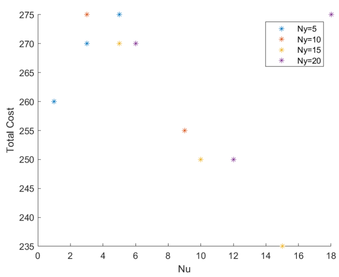

Fig. 3. Total cost for a range of horizon combinations

V. ANALYSIS

From the results of the first experiment, the 95% confidence interval (CI) of the true expected total cost, for the horizon pair $(5,5)$, was 260.95 ± 4.15. This suggests using a variable stock update was beneficial since the optimal fixed re-order policy, determined in assignment 1, had a true expected total cost in the 95% CI 280.73 ± 1.97. Figure 2 illustrates what the optimal choice of stock updates looked like. In many cases the sum of the end of week stock and the stock update equalled four; however, this was certainly not true for all cases. There were plenty of examples where the stock level dropped to one and an update of five was made. This suggests the expected horizon cost was minimised by being certain a stock penalty couldn't occur. Furthermore, on the 51^{st} week, the stock level dropped to three and no update was made; this supported the prediction that it was worth taking a risk to avoid the penalty and return cost.

The results of the second experiment show that when N_y was equal to 10 and 15 a negative linear relationship was illustrated between the control horizon and total cost; but when N_y was equal to 5 a positive linear relationship was illustrated between the same variables. When N_y was 20, no relationship was shown. This suggests, that for this particular system, there were detrimental effects for not looking far enough into the future and for considering events to distant into the future. A prediction horizon of 10 or 15 provided a happy medium where performance improved by increasing the control horizon. The minimum total cost was 235 coins which occurred when the horizon pair was $(N_y, N_u) = (15,15)$. Furthermore, since this was a stochastic system, the predicted stock dynamics rarely matched the observed dynamics; thus the stock updates found through the optimisation step didn't necessarily correspond to the optimal stock updates of the observed dynamic system.

VI. CONCLUSION

Results confirmed that a variable stock update significantly reduced the mean total cost compared to the fixed order update found by Monte Carlo simulation. Furthermore, the performance of different horizon pairs did not always reduce the cost as the horizon increased. It was suggested that the stochastic nature of the system was the reason behind this. In order to further investigate the varying performance of different horizon pairs, the same experiment could be repeated for many demand and predicted demand sequences so that statistics about the performance could be recorded. This would confirm whether the conclusions drawn from the second experiment were specific to that particular simulation or were a common theme of the system.

APPENDIX

$$w_t = \begin{cases} 5 & x_t > 0 \\ 0 & otherwise \end{cases} \tag{5}$$

$$p_t = \begin{cases} 20 & x_t < 0 \\ 0 & otherwise \end{cases} \tag{6}$$

$$r_t = \begin{cases} 10x_t & t = 52 \\ 0 & otherwise \end{cases} \tag{7}$$

In the implementation of the system, the costs were calculated after each update to the stock level. If the demand exceeded the previous weeks stock and the update, the updated stock would fall below zero at this point in the program, hence equation 6 works. Following the calculation of the costs, the stock level was reset to zero if it was less than zero beforehand. This ensured the non-negative constraints were met.