Effective Heuristics For Capacitated Production Planning with Multi-Period Production and Demand With Forecast Band Refinement

Philip Kaminsky
Department of IEOR
University of California
Berkeley, CA-94720
Ph: (510)-642-4927
Fax: (510)-642-1403
email: kaminsky@ieor.berkeley.edu

Jayashankar M. Swaminathan
Kenan-Flagler Business School
University of North Carolina
Chapel Hill, NC-27599
Ph: (919)-843-8341
Fax: (919)-843-4897
email: msj@unc.edu

February 2002

Effective Heuristics For Capacitated Production Planning with Multi-Period Production and Demand With Forecast Band Refinement

Abstract

In this paper we extend forecast band evolution and capacitated production modelling studied by Kaminsky and Swaminathan (2001) to the multi-period demand case. In their model, forecasts of discrete demand for any period are modelled as bands, defined by lower and upper bounds on demand, such that future forecasts lie within the current band. We develop heuristics which utilize knowledge of demand forecast evolution to make production decisions in capacitated production planning environments. In our computational study, we explore the efficiency of our heuristics, and also explore the impact of seasonality in demand and availability of information updates.
Keywords: Forecast Evolution, Production Planning, Discrete Demand, Capacity, Heuristics, Information Updates.

1 Introduction

As increasing competition has forced firms to operate their supply chains more efficiently, the need for improved forecasts has become clear. Indeed, firms have increasingly devoted resources to improving the quality of forecasts within the supply chain (Lee et al. 1997). These resources can be used in many ways. Subjective forecasts can be improved, for example, with improved sales force composites and more detailed customer surveys, or by soliciting the opinions of more experts. Objective forecasts can be improved, for example, by collecting more data and developing more detailed causal models. As a result, it is becoming more important to develop models and techniques which assist managers in determining the best places to focus their forecast improvement efforts, and which enable firms to make effective use of these forecasts in production planning. Clearly, any approach should account for the following facts about forecasts:

- Forecasts are typically wrong.
- A good forecast is more than a single number.
- The longer the forecast horizon is, the more wrong the forecast typically is.

In addition, in a capacitated production environment, there may not be enough production capacity to wait for the forecast of a particular event to improve as that event gets closer.

Motivated by examples in the electronics industry, Kaminsky and Swaminathan (2001) present a model in which forecasts are represented by a series of bands, encompassing both optimistic and pessimistic forecasts, which are updated each period as new information arrives. In particular, this forecast evolution model is defined by an upper and a lower bound on possible demand, which is called a band. As time moves forward, subsequent forecasts have a smaller range between the upper and lower bounds, or a smaller band (representing a better forecast), and are contained within the band defined by earlier forecasts. This models a forecasting process which gets refined over time as new information arrives. The manufacturing firm utilizing this forecast has a fixed capacity in each period and needs to decide how much to produce in each period, taking into account expected production, holding, salvage and stock-out costs. Kaminsky and Swaminathan (2001) explore the case with a single terminal demand and develop effective heuristics for the problem. In this paper, we consider the case where demand can occur in every period, develop alternative heuristics which fit this multi-period demand model, and computationally test their effectiveness. We also consider the impact on the system of improved forecast updates, capacity availability, and seasonality in demand.

The paper is organized as follows: In Section 2 we present related literature. In Section 3 we present the model, properties and our heuristics. In Section 4, we present our computational results and insights, and we conclude in Section 5.

2 Related Literature

Forecast evolution models have been studied by several researchers in the past. Hausman (1969) studies the problem in which improved forecasts are available before each decision stage, and he models the evolution of forecasts as a quasi Markovian or Markovian system. He suggests modeling a series of ratios of successive forecasts as independent lognormal variates and presents a dynamic programming formulation. Hausman and Peterson (1972) consider a single capacity constrained multi-product system with terminal demand with the forecasts for total sales following a lognormal model. They formulate the problem as a dynamic program and provide heuristics for solving the problem. Heath and Jackson (1994) introduce a martingale model for forecast evolution (MMFE) and present simulation results. Gullu (1996) considers a special case of the MMFE model and shows that in a capacitated environment the system performs better when one period demand forecast is utilized than when it is not. More recently, Toktay and Wein (2001) consider an MMFE model and characterize effective policies under heavy traffic assumptions for a capacitated single server.

Bayesian models for forecast updates in a inventory setting were first studied by Scarf (1959). Azoury (1985) extends some of Scarf's results. Bitran et al. (1986) study a terminal demand model for style goods under capacity restrictions and Bayesian forecast updates, where each of the product families is produced one period before the realization of demand. They present a stochastic mixed integer programming formulation and provide a decomposition scheme and bounds on the optimal solution. Fisher and Raman (1996) consider a similar two period problem from the fashion goods industry where the manufacturer needs to determine production quantities in two periods, with lower production costs in the first period but improved forecast information available in the second. Agarwal et al. (1999) consider a similar problem but analyze capacity outsourcing issues. Nurani et al. (1994) consider uncapacitated inventory problems with forecast updates and analyze the effectiveness of myopic heuristics. Lee and Whang (1998) study the postponement problem and identify additional benefits of postponement that are realized due to resolution of uncertainty of demands up to the point of postponement. Gurnani and Tang (1998) consider the benefits of updated demand forecast in a two period problem under uncertain costs. There are a variety of other papers that study capacitated non-stationary inventory problems; Swaminathan and Tayur (2001) discuss these papers, and provide an extensive review of other inventory models used for tactical production planning.

As mentioned above, Kaminsky and Swaminathan (2001) introduce a forecast evolution process based on bands, in which the range of possible demand is completely contained within a *forecast band*. This band, defined by upper and lower bounds, captures many of the important elements of real world forecast revisions, and in particular, the concept of improved information over time (Buckley et al. 1995). It is to be noted that earlier papers such as Bitran *et. al.* (1986) and Hausman and Peterson (1972) consider forecast evolutions which get updated and refined over time. The former captures it through the standard deviation from the actual demand and the latter captures it through the relative uncertainty in the forecast. The band model differs in that it restricts the future demand forecasts using a hard limit on the lower and upper bounds. In the discrete demand case, this allows one to simplify the capacitated production planning problem and develop effective solution approaches. In this paper, we use the forecast band model, but we generalize this to the multi-period case, where demand can occur in every period.

3 Basic Model

3.1 Forecast Evolution and Production Model

In this section we present the forecast evolution model. We assume that demand is discrete, and number time periods in reverse order, so that when we say time t, we mean that t periods remain. Period 1 is the final period. Some notation follows:

- a_l^j: the minimum possible demand for period j forecast in period l;
- w_l^j: width of the forecast for demand in period j given that we are in period l;
- $F_t^j(a_{t+1})$: probability distribution of a_t^j in the interval $[a_{t+1}^j, a_{t+1}^j + w_{t+1}^j - w_t^j]$;
- α_t^j: known reduction in width of forecast; $w_t^j = w_{t+1}^j - \alpha_t^j$
- x_t: on-hand inventory at the start of period t;
- q_t: production quantity in period t;
- C: production capacity in each period;

- π: per unit penalty cost for demand not met;
- c: per unit cost of production;
- h: per unit holding cost incurred on inventory held at the end of any period;
- $s(< c)$: per unit salvage cost for each unit remaining after demand is met;
- $I_k = (a_k^{k-1}, w_k^{k-1}, a_k^{k-2}, w_k^{k-2}, ..., a_k^1, w_k^1)$: information available at the end of period k;
- $V_t(x_t, I_{t+1})$: expected optimal cost incurred with t periods to go, where x_t is the inventory on hand, and I_{t+1} denotes the information available at the start of period t;

In the forecast evolution model, given that we are at the end of period t, we know that demand in period j can only take on discrete values on the interval $[a_t^j, a_t^j + w_t^j]$. In addition, each period the range of possible demand gets smaller in a way consistent with the process described above. That is, $w_t^j = w_{t+1}^j - \alpha_t^j$, where each $\alpha_t^j \geq 0, \forall t$ are such that $w_j^j \geq 0$. The final demand a_j^j is defined on the discrete interval $[a_{j+1}^j, a_{j+1}^j + w_{j+1}^j]$ with a probability distribution $F_j(a_{j+1}^j)$. This forecast evolution process captures the above two assumptions: (1) the forecast gets tighter as one gets closer to demand since w_t^j is increasing in $t - j$, and (2) the future forecast are contained in the current range since a_{t-1}^j lies in the range $[a_t^j, a_t^j + w_t^j - w_{t-1}^j]$. Note that $F_t(a_{t+1}^j)$ allows us to model changes in distribution over time, and α_t^j provides the ability to model the different points of time at which major or minor updates in forecast may be obtained. A firm that obtains all the forecast updates very close to the selling season will have $\alpha_t^j = 0$ for values of t further away from j , and a high value for α_t^j for values of t very close to j. Note that we assume that the progression of w_t^j is known, whereas the progression of a_t^j is probabilistic. This models a situation in which the firm has a good understanding of when additional information will be obtained (for example, a trade show), but doesn't know *a priori* what that information will be.

We assume that demand occurs every period, and that the manufacturer can produce between 0 and C units each period. We consider a T period finite horizon model, in which demand can occur in each of the periods. Periods are numbered backwards; T is the first period, and the last period is 1. The objective is to minimize the expected total cost during the finite horizon; at each period a production decision needs to made. More specifically, the order of events in each period t is as follows:

1. the production decision q_t is made
2. demand a_t^t is realized from the range $[a_{t+1}^t, a_{t+1}^t + w_{t+1}^t]$ with probability distribution $F_t^j(a_{t+1})$
3. the forecasts for each subsequent period i, a_t^i are realized from the range $[a_{t+1}^i, a_{t+1}^i + w_{t+1}^i]$ with probability distribution $F_t^i(a_{t+1})$
4. holding and penalty costs are charged

Let $V_t(x, I_{t+1})$ denote the optimal cost to go function with t periods remaining in the horizon when the on-hand inventory is x and the information available is I_t. We assume that all unmet demand is back ordered. Then,

$$V_t(x, I_{t+1}) = \min_{0 \le q_t \le C} \mathbf{E}\{q_t c + h(x + q_t - \xi_t)^+ + \pi(\xi_t - x - q_t)^+ + V_{t-1}(x + q_t - \xi_t, I_t)\}$$

and

$$V_1(x, I_2) = \min_{0 \le q_1 \le C} \mathbf{E}\{q_1 c + (h + s)(x + q_1 - \xi_1)^+ + \pi(\xi_1 - x - q_1)^+\}$$

The expectation $\mathbf{E}$ is with respect to all possible realization of demand in the current period, and all possible values of I_t. It is clear that this problem is quite challenging. It is a non-stationary inventory problem with capacity restrictions and forecast evolution. A threshold policy is optimal if the optimal cost to go function is convex, which in the discrete case reduces to showing increasing first differences. We will now show that the one period and t period objectives have increasing first differences.

Lemma 3.1 *Let* $L(y, I) = cy + \mathbf{E}_\xi\{(h+s)(y-\xi)^+ + (\pi)(\xi - y)^+\}$*. Then* $L(y+2, I) - L(y+1, I) \geq L(y+1, I) - L(y, I)$*.*

Proof. Let $\hat{L}(y, \xi) = cy + \{(h+s)(y-\xi)^+ + (\pi)(\xi - y)^+\}$.

When $\xi \geq y+2$ then $\hat{L}(y+2,\xi) - \hat{L}(y+1,\xi) = \hat{L}(y+1,\xi) - \hat{L}(y,\xi) = c - \pi$. When $\xi \leq y$ then $\hat{L}(y+2,\xi) - \hat{L}(y+1,\xi) = \hat{L}(y+1,\xi) - \hat{L}(y,\xi) = c + (h+s)$. When $\xi = y+1$ then $\hat{L}(y+2,\xi) - \hat{L}(y+1,\xi) = c + h + s$ and $\hat{L}(y+1,\xi) - \hat{L}(y,\xi) = c - \pi$. Therefore,

$$(\hat{L}(y+2,\xi) - \hat{L}(y+1,\xi)) - (\hat{L}(y+1,\xi) - \hat{L}(y,\xi)) \geq 0 \ \ \forall \xi$$

Thus, $L(y+2, I) - L(y+1, I) \geq L(y+1, I) - L(y, I)$. ∎

Theorem 3.2 $V_1(x+2, I_2) - V_1(x+1, I_2) \geq V_1(x+1, I_2) - V_1(x, I_2) \; \forall \, x, I_2$*.*

Proof. $V_1(x, I_2) = \min_{0 \leq q_t \leq C} \mathbf{E}\{q_t c + (h+s)(x + q_t - \xi_t)^+ + \pi(\xi_t - x - q_t)^+\} = -cx + \min_{x \leq y \leq x+C} \mathbf{E}\{cy + (h+s)(y - \xi_t)^+ + \pi(\xi_t - y)^+\} = -cx + \min_{x \leq y \leq x+C} L(y, I_2)$

Let $g = \arg\min L(y, I_2)$. There are five cases for possible values of g.

Case (1): $g \leq x$. In this case, $V_1(x+2, I_2) - V_1(x+1, I_2) = -c + L(x+2, I_2) - L(x+1, I_2)$ and $V_1(x+1, I_2) - V_1(x, I_2) = -c + L(x+1, I_2) - L(x, I_2)$. From Lemma 3.1 we know that $V_1(x+2, I_2) - V_1(x+1, I_2) \geq V_1(x+1, I_2) - V_1(x, I_2)$.

Case (2): $g = x+1$. In this case, $V_1(x+2, I_2) - V_1(x+1, I_2) = -c + L(x+2, I_2) - L(x+1, I_2)$ and $V_1(x+1, I_2) - V_1(x, I_2) = -c$. Since $g = x+1$ is the minimal point and L is convex (increasing first differences), we know that $L(x+2, I_2) \geq L(x+1, I_2)$ therefore $V_1(x+2, I_2) - V_1(x+1, I_2) \geq V_1(x+1, I_2) - V_1(x, I_2)$.

Case (3): $g = x+C+1$. In this case, $V_1(x+2, I_2) - V_1(x+1, I_2) = -c$ and $V_1(x+1, I_2) - V_1(x, I_2) = -c + L(x+C+1, I_2) - L(x+C, I_2)$. Since $g = x+C+1$ is the minimal point and L is convex (increasing first differences), we know that $L(x+C, I_2) \geq L(x+C+1, I_2)$ therefore $V_1(x+2, I_2) - V_1(x+1, I_2) \geq V_1(x+1, I_2) - V_1(x, I_2)$.

Case (4): $g \geq x + C + 2$. In this case, $V_1(x+2, I_2) - V_1(x+1, I_2) = -c + L(x+C+2, I_2) - L(x+C+1, I_2)$ and $V_1(x+1, I_2) - V_1(x, I_2) = -c + L(x+C+1, I_2) - L(x+C, I_2)$. From Lemma 3.1 therefore, $V_1(x+2, I_2) - V_1(x+1, I_2) \geq V_1(x+1, I_2) - V_1(x, I_2)$.

Case (5): $x + 2 \leq g \leq x + C$. In this case, $V_1(x+2, I_2) - V_1(x+1, I_2) = -c$ and $V_1(x+1, I_2) - V_1(x, I_2) = -c$ therefore $V_1(x+2, I_2) - V_1(x+1, I_2) \geq V_1(x+1, I_2) - V_1(x, I_2)$.

Thus $V_1(x+2, I_2) - V_1(x+1, I_2) \geq V_1(x+1, I_2) - V_1(x, I_2)\ \forall x, I_2$.∎

With Theorem 3.2, we have shown that with one period objective function $V_1(x, I_2)$ has increasing differences and therefore a threshold policy is optimal. Next, we show this is true any $V_t(x, I_{t+1})$.

Theorem 3.3 $V_t(x+2, I_{t+1}) - V_t(x+1, I_{t+1}) \geq V_t(x+1, I_{t+1}) - V_t(x, I_{t+1}) \forall\ x, I_{t+1}$.

Proof. We will show this by induction. Clearly from Theorem 3.2 we know that the relationship is true for $t = 1$. Let us assume that it is true for $t - 1$ and we will prove it for t. Let $G(y, I) = cy + \mathbf{E}_\xi\{(h)(y-\xi)^+ + (\pi)(\xi - y)^+\}$.

$$
\begin{aligned}
V_t(x, I_{t+1}) &= \min_{0 \leq q_t \leq C} \mathbf{E}\{q_t c + h(x + q_t - \xi_t)^+ + \pi(\xi_t - x - q_t)^+ + V_{t-1}(x + q_t - \xi_t, I_t)\} \\
&= -cx + \min_{x \leq y \leq x+C} \{G(y, I_{t+1}) + \mathbf{E}V_t(y - \xi_t, I_t)\}
\end{aligned}
$$

Note that $G(y, I)$ is identical to $L(y, I)$ excepting that the holding cost is changed from $h+s$ to h. It can be easily shown that $G(y+2, I) - G(y+1, I) \geq G(y+1, I) - G(y, I)$ similar to Lemma 3.1. $V_{t-1}(y - \xi_t, I_{t-1})$ had increasing differences (by the induction hypothesis) so $T(y) = G(y, I_{t+1}) + \mathbf{E}V_t(y - \xi_t, I_t)$ has increasing differences. Let $g = \arg\min T(y)$. There are five cases, for each possible value of g.

Case (1): $g \leq x$. In this case, $V_t(x+2, I_{t+1}) - V_t(x+1, I_{t+1}) = -c + T(x+2) - T(x+1)$ and $V_t(x+1, I_{t+1}) - V_t(x, I_{t+1}) = -c + T(x+1) - T(x)$. Therefore, $V_t(x+2, I_{t+1}) - V_t(x+1, I_{t+1}) \geq V_t(x+1, I_{t+1}) - V_t(x, I_{t+1})$.

Case (2): $g = x+1$. In this case, $V_t(x+2, I_{t+1}) - V_t(x+1, I_{t+1}) = -c + T(x+2) - T(x+1)$ and $V_t(x+1, I_{t+1}) - V_t(x, I_{t+1}) = -c$. Since $g = x + 1$ is the minimal point and T

is convex (increasing first differences), we know that $T(x+2) \geq T(x+1)$ therefore $V_t(x+2, I_{t+1}) - V_t(x+1, I_{t+1}) \geq V_t(x+1, I_{t+1}) - V_t(x, I_{t+1})$.

Case (3): $g = x + C + 1$. In this case, $V_t(x+2, I_{t+1}) - V_t(x+1, I_{t+1}) = -c$ and $V_t(x+1, I_{t+1}) - V_t(x, I_{t+1}) == -c + T(x+C+1) - T(x+C)$. Since $g = x+C+1$ is the minimal point and T is convex (increasing first differences), we know that $T(x+C) \geq L(x+C+1)$ therefore $V_t(x+2, I_{t+1}) - V_t(x+1, I_{t+1}) \geq V_t(x+1, I_{t+1}) - V_t(x, I_{t+1})$.

Case (4): $g \geq x + C + 2$. In this case, $V_t(x+2, I_{t+1}) - V_t(x+1, I_{t+1}) = -c + T(x+C+2) - T(x+C+1)$ and $V_t(x+1, I_{t+1}) - V_t(x, I_{t+1}) = -c + T(x+C+1) - T(x+C)$. Therefore, $V_t(x+2, I_{t+1}) - V_t(x+1, I_{t+1}) \geq V_t(x+1, I_{t+1}) - V_t(x, I_{t+1})$.

Case (5): $x + 2 \leq g \leq x + C$. In this case, $V_t(x+2, I_{t+1}) - V_t(x+1, I_{t+1}) = -c$ and $V_t(x+1, I_{t+1}) - V_t(x, I_{t+1}) = -c$ therefore $V_t(x+2, I_{t+1}) - V_t(x+1, I_{t+1}) \geq V_t(x+1, I_{t+1}) - V_t(x, I_{t+1})$.

$V_t(x+2, I_{t+1}) - V_t(x+1, I_{t+1}) \geq V_t(x+1, I_{t+1}) - V_t(x, I_{t+1}) \forall x, I_{t+1}$ follows by induction.

∎

Corollary 3.4 *For every period t there exists a threshold level $y_t^*(I_{t+1})$ such that it is optimal to produce in period t if $x_t \leq y_t^*(I_{t+1})$ and to not produce if $x_t > y_t^*(I_{t+1})$.*

As the state space for this problem is very large, optimally solving the dynamic programming recursion for this model is very complex and time intensive. We are thus motivated to consider a variety of heuristics for this problem in order to handle realistic size problems.

3.2 Heuristics

We consider three heuristics for this model. In particular, the first two heuristics involve reduction of the state space by simplification of the model in various ways. The final heuristic, motivated by the heuristics in Kaminsky and Swaminathan (2001), as well as approaches first proposed by Morton and Pentico (1995), solves a series of related single period uncapacitated problems, and then uses this information to make production decisions.

- **H1: Solving the problem optimally for some small number of periods in the future**. We solve the P period finite horizon problem optimally given I_{t+1} (that is, ignoring forecast evolution), where $P << T$. Note that a myopic approach is a special case of this heuristic where only one period is considered, and that salvage cost is added only if the final period is being considered.

- **H2: Solving the problem for lower and upper bounds for a limited time horizon.** In this heuristic, we assume that the demands in future periods are deterministic, fixed at the lower bounds or upper bounds, and find the optimal production for this (deterministic) system for P' periods, where $P' < T$. Note that since this deterministic problem is easier to solve than the stochastic problem, the H2 horizon P' can be longer than the H1 horizon P. In Heuristic **H2l**, if we are at time t, we fix demand at time $l \leq t$ at a_t^l, while heuristic **H2u** fixes demand at $a_t^l + w_t^l$. As in Heuristic H1, salvage cost is added only if the final period is being considered.

- **H3: Solving the non-stationary problem without capacity constraints and extrapolating.** Kaminsky and Swaminathan (2001) solve the terminal demand version of this problem by first solving an uncapacitated single period problem, and then adapting that solution to the multi-period capacitated production model. Conceptually, they do this "spreading the manufacturing backwards" so that the desired inventory can be manufactured in the presence of capacity constraints. In this heuristic **H3**, we use a similar approach.

  Of course, one difficulty in doing this is that the uncapacitated problem is challenging to solve. Fortunately, Morton and Pentico (1995) develop a heuristic which they computationally demonstrate produces very good solutions to the uncapacitated multi-period nonstationary stochastic inventory problem. This heuristic works by solving this multi-period problem as if it is a single period problem with demand equal to the cumulative remaining demand, and adjusted holding costs.

In our heuristic **H3**, we combine this idea of "spreading back" production from Kaminsky and Swaminathan (2001), with a modified version of the Morton and Pentico heuristic.

First, we present our modified version of the Morton and Pentico heuristic. Recall that this heuristic is based on solving a modified single period problem. Now, let Φ_{t1} be the cumulative distribution of demand over the periods from t to 1 inclusive, assuming that the distribution in each subsequent period is conditional only on I_t (that is, without forecast evolution, as in heuristic **H1**). Then the optimal inventory level for a single period model with modified holding costs, the cumulative demand, and maximum production equal to the total available capacity, is calculated. Specifically, the inventory level is calculated in two ways:

$$\bar{y}_{t1}^{m} = \arg \min_{0 \leq y \leq TC} E_{d_{t1}} \{cy + (h - s)(y - d_{t1})^{+} + (\beta\pi)(d_{t1} - y)^{-}\}$$

where d_{t1} is distributed according to Φ_{t1}. Also, the one period myopic inventory level is calculated

$$\bar{y}_{t}^{m} = \arg \min_{0 \leq y \leq TC} E_{d_t} \{cy + (h)(y - d_t)^{+} + (\beta\pi)(d_t - y)^{-}\} \qquad t > 1$$

and

$$\bar{y}_{1}^{m} = \arg \min_{0 \leq y \leq TC} E_{d_1} \{cy + (h - s)(y - d_1)^{+} + (\pi)(d_1 - y)^{-}\}$$

where d_t is distributed according to Φ_t. Observe that shortages in this period also impact costs in later periods, as demand is backlogged. Therefore, we multiply the penalty cost by a factor β to account for its impact on future periods. To determine β, we first determine the load factor λ as follows:

$$\lambda = \min\{2, \max\{1, \frac{y_t^m}{.8tC}\}\}$$

which represents the approximate fraction of 80 % of remaining capacity which will be required, bounded by one on the low end and two on the high end. If this remaining capacity is tight, back order will be more damaging in the long run. We then determine the penalty multiplier β by multiplying the load factor by the number of periods remaining, bounded on the low end by two:

$$\beta = \min\{\lambda t, 2\}$$

Following Morton and Pentico's EOS heuristic, we next take the uncapacitated order-up-to level to be:

$$\bar{y}_t = \min\{\bar{y}^m_{t1}, \bar{y}^m_t\}$$

Of course, the production schedule suggested by this heuristic will typically be infeasible for the capacitated problem. We thus need to move some of the production earlier where necessary to account for these capacity constraints. We detail this approach, heuristic **H3**, below.

Suppose we are in period t. We implement the following procedure to determine the order up to level in this period. First, we calculate $\bar{y}_i, i = t, t-1, ..., 1$ as described above for the uncapacitated case, using the appropriate demand distributions and convolutions, assuming that demand doesn't continue to evolve, but is instead distributed in future periods based on the current bands. Next, we adjust these order-up-to levels to account for the capacity constraint using the following procedure:

1. **Calculate** $\bar{y}_i, i = t, t-1, ..., 1$ as described above.
2. **New levels** are determined recursively (from 1 to t) using the following formulation:

$$y_j = \min\{C, \bar{y}_j + \sum_{i=1}^{j-1}(\bar{y}_i - y_i)\}, \qquad j = 1, 2, ..., t-1$$

$$y_t = \min\{C, \bar{y}_j + \sum_{i=1}^{j-1}(\bar{y}_i - y_i) - x_t\}$$

Note that although this approach calculates a series of y_i values, these have to be recalculated for each t since the distribution and starting inventory changes as information is updated.

4 Computational Study

We present the results of computational studies, in which we tested our heuristics under different settings, and explored the benefits of improving forecasts earlier in the time horizon. We use the following parameters for our computational study:

Time Horizon: In order to test the effectiveness of our heuristics we compare heuristic solutions to the optimal solution. For these tests we use a four period time horizon, the largest time horizon for which our problem can practically be solved to optimality. For additional tests in which use only our heuristics, we utilize a 24 period time horizon.

Cost Parameters: We use production cost $c = 50$, penalty $\pi = 75, 150, 250$, salvage $s = 0, 10, 25, 45$ and holding cost $h = 0, 2, 4, 8, 12$ in the study, to determine the effect of relatively high and low penalty, salvage, and holding costs (similar to Morton and Pentico 1995).

Demand: We assume that demand in period t is uniformly distributed on the interval $[a_t^t, w_t^t]$. We set initial demand and forecast updates so that final demand has a width of 5. We set the initial width equal to a quantity no larger than 11, and less if there is not enough time to reduce the width to 5. Thus, for 4 period time horizons, we set the w values as follows:

$$w_5^4 = 5, w_5^3 = 6, w_5^2 = 7, w_5^1 = 8$$

Similarly, for 24 period horizons, we set initial w values as follows:

$$w_{25}^{24} = 5, w_{25}^{23} = 6, w_{25}^{22} = 7, w_{20}^{21} = 8, w_{25}^{20} = 9, w_{25}^{19} = 10$$

and for the remaining values

$$w_{25}^{18} = 11, i = 1, 2, ..., 18.$$

In order to reduce the bands to the appropriate value in time, we set α's as follows. For the 4 period horizon:

$$\alpha_l^t = 1, t = 1, 2, 3, l = t + 1, ..., 4$$

with remaining α's set to 0.

For the 24 period horizon:

$$\alpha_l^t = 1, t = 1, 2, ..., 18, l = t + 1, t + 2, ..., t + 6$$

$$\alpha_l^t = 1, t = 19, 20, ..., 23, l = t + 1, ...t + (24 - t)$$

with remaining α's set to 0.

All that remains is to set the values for the a^t values. Recall that the a^t values represent a lower bound on demand, and thus can be set to simulate stationary demand, as well as a variety of seasonal demand variation. For the four period time horizon, we consider the following three cases.

- the **BASE** case: $a_5^t = 4, t = 1, 2, 3, 4$.
- the **MODERATELY SEASONAL** case: $a_5^t = 5, t = 1, 3$, and $a_5^t = 3, t = 2, 4$.
- the **SEASONAL** case: $a_5^t = 8, t = 1, 3$, and $a_5^t = 0, t = 2, 4$.

Similarly, for the 24 period time horizon, we consider the following three cases:

- the **BASE** case: $a_{25}^t = 4, t = 1, 2, ..., 24$
- the **MODERATELY SEASONAL** case: $a_{25}^t = 5, t = 1, 2, 3, 4, 9, 10, 11, 12, 17, 18, 19, 20$, and $a_{25}^t = 3, t = 5, 6, 7, 8, 13, 14, 15, 16, 21, 22, 23, 24$.
- the **SEASONAL** case:$a_{25}^t = 8, t = 1, 2, 3, 4, 9, 10, 11, 12, 17, 18, 19, 20$, and $a_{25}^t = 0, t = 5, 6, 7, 8, 13, 14, 15, 16, 21, 22, 23, 24$.

Capacity: We examine three capacity cases, **HIGH**:$C = 13$,**MEDIUM**:$C = 9$,**LOW**:$C = 7$. To understand these capacity values, compare them to the $a_t + w_t$ values for the different scenarios described above.

4.1 Performance of the Heuristics

We tested the effectiveness of the heuristics for two different time horizons. For the four period horizon, we compared heuristics and the optimal solution for all combinations of the parameters listed above. Each parameter combination was tested through five simulation runs. In the simulation runs, at each period, the ordering decision was made based on the algorithm, and then random demand and forecast updates were realized. In addition, we compared the four heuristics (without the optimal solution) for all of the combinations of the parameters listed above for a twenty four period time horizon, using the same simulation-based approach. In all cases, the horizon for heuristic **H1** was three periods. For the four period problems, heuristics **H2l**, **H2u** and **H3** were solved for all periods. For the 24 period problems, heuristics **H2l** and **H2u** were calculated using a six period horizon.

We find that for the 4 period problems (542 in all), **H1** and **H3** perform very well, with average objective values of 0.008% and 0.12% over the optimal strategy respectively, while **H2u** was on average 1.14% worse than the optimal strategy. **H2l** performs rather poorly, averaging 95.46% worse than the optimal strategy (see Table 1). It is not surprising that for a four period problem solving the three period horizon optimally, (**H1**) is likely to be very close to the overall optimal solution. On the other hand, by producing to meet the lower bound in the demand, **H2l** can be quite ineffective since the penalty cost is much higher than the holding cost. Also not surprisingly, the more sophisticated **H3** heuristic which extrapolates production based on an unconstrained heuristic performs very well. We also find that **H2l**, **H2u** and **H3** run much faster than **H1** (see Table 1, where times are seconds on a Pentium III-800 computer).

	H1	**H2l**	**H2u**	**H3**
Average Error	0.008%	95.46%	1.14%	0.12%
Average Time	0.22	0.004	0.004	0.009

Table 1: Average error and running times for H1, H2l, H2u and H3 for the four period problems.

For the problem with 24 periods, computing the optimal strategy was unrealistic in terms of time and memory constraints. As a result, we compare the average cost incurred by the heuristics across the 1620 problem instances that we tested. We find that **H2u** and **H3** perform significantly better than **H1** and **H2l**. All the heuristic run within a reasonable amount of time even for these longer horizon problems (see Table 2). This suggests that heuristic **H2u** and **H3** are efficient and effective even when the number of periods is high. When we compare these two heuristics, we see their performance is quite similar, with **H2u** having a slightly better average performance in the 24 period case, although **H3** actually performs slightly better than **H2u** in about 64% of the experiments.

	H1	**H2l**	**H2u**	**H3**
Average Error	43469	62126	35797	36000
Average Time	2.98	2.23	2.16	2.24

Table 2: Average error and running times for H1, H2l, H2u and H3 for the four period problems.

4.2 Forecast Updates, Capacity, and Seasonality

In this series of experiments, we evaluated the value of the timing of forecast information updates. We modelled the information updates in the forecast process by changing the width of the forecast band from period to period. We consider EARLY INFORMATION, INTERMEDIATE INFORMATION and LATE INFORMATION models. In the EARLY (INTERMEDIATE, LATE) INFORMATION case the band width reduces more per period in the first few (middle, last few) periods respectively, although in all cases, the bands decrease from their initial width to a width of 5. For example, for EARLY INFORMATION, all α's except for the final six before the demand period are 0. The final six $\alpha's$ are represented by the vector (321000). Similarly, the INTERMEDIATE and LATE INFORMATION cases are represented by (111111) and (000123) respectively. The INTERMEDIATE CASE is detailed above. Presumably the firm has to invest in better forecasting systems in order to obtain the refined updates earlier in the horizon. We do not consider such costs in our analysis but only focus on the benefits. Further, we also consider three levels of seasonality as indicated earlier which to an extent reflects the correlation in the demand across successive periods. We summarize our insights below:

- Unsurprisingly, total costs increase with increase in penalty and holding costs and costs decrease with increase in capacity.
- In general, increase in seasonality increases total costs when capacity is high. For example, based on the optimal solutions to the instances with 4 periods, in Table

3 we see a slight increase in cost when capacity is high. This is not always the case when capacity is tighter. For example, in the optimal solutions to the 4 period instances, when the capacity is medium (C=9) the cost first decreases and then increases and when capacity is tight (C=7) the costs decrease with seasonality. We observed similar results with other data sets. Apparently, higher seasonality provides a form of negative correlation, and the benefits of negative correlation are more dramatic when capacity is lower.

	Capacity = 7	Capacity = 9	Capacity = 13
Base Seasonal	1762	1464	1434
Intermediate Seasonal	1695	1453	1434
High Seasonal	1574	1466	1436

Table 3: Effect of Capacity and Seasonality when $h = 4, p = 75, c = 50, s = 10$ for the four period problem.

- We observe that obtaining information updates earlier is beneficial in most cases and we also find that the marginal benefit of obtaining earlier information is lower under higher capacity. For example, in Table 4 we observe this for an instance with 24 periods, when when $h = 4, p = 75, c = 50, s = 10$, when heuristic $H3$ is used. We see similar results for other instances and heuristics. Intuitively, obtaining earlier updates enables the firm to plan production more effectively. However, this is crucial only when the firm has limited capacity. When the firm has a large capacity base it has the ability to adequately respond to abrupt changes in the demand process, and hence the value of earlier information is minimal.

	Capacity = 7	Capacity = 9	Capacity = 13
Early Update	42596	11986	11712
Intermediate Update	44186	12737	11821
Late Update	46210	13387	11969

Table 4: Effect of Capacity and Information Updates when $h = 4, p = 75, c = 50, s = 10$ for the 24 period problem under intermediate seasonality.

- Interestingly, there are cases when obtaining early updates is not necessarily ben-

	Capacity = 7	Capacity = 9	Capacity = 13
Early Update	35383	12321	12597
Intermediate Update	36988	12375	11725
Late Update	38902	12654	11898

Table 5: Effect of Capacity and Information Updates when $h = 4, p = 75, c = 50, s = 10$ for the 24 period problem under high seasonality.

eficial. For example, consider the case in which capacity is high ($C = 13$) demand is highly seasonal, as in Table 5. This can be explained by observing that early information is less precise when seasonality is high, and when capacity is high, there is a chance of overreaction to early information updates. When capacity is lower, overreaction is constrained by available capacity.

5 Conclusions

We present a model that captures forecast evolution information while making production decisions in a capacitated environment. We consider a forecast evolution model which is defined by a band (with upper and lower bounds) that captures the uncertainty in the forecast. As time moves forward, we assume the next forecast has a smaller width (representing a better forecast), and that the new band can lie anywhere inside the old band. This models a forecasting process which gets refined over time as new information arrives. We consider a manufacturing firm utilizing this forecast which has a fixed capacity in each period and needs to decide in which periods it should produce in order to meet demand in each period, taking into account expected production, holding, salvage and stock-out costs. We prove the existence of inventory threshold levels in each period below which the firm should produce and above which it should not. We develop several heuristics, based on either reducing the size of the dynamic programming state space, or on solving a series of uncapacitated problems, and then modifying the solutions for the capacitated case. Our computational study shows that several of these heuristics are very effective.

We empirically observe that early information updates in the forecasting process generally leads to decreased costs, although this is most dramatic when capacity is tight. Indeed, when capacity is high and demand is highly seasonal, over-reactions to early information updates may be detrimental to system performance. We also observe that in this model, increases in seasonality lead to increases in total costs when capacity is high, but may actually lead to decreased costs when capacity is tight. Finally, we find that the marginal benefit of earlier information is lower when capacity is high.

There are certain limitations in the model we present in this paper. First, we assume that forecasts get better or at least are as good in future periods. Although this assumption is reasonable under most conditions there may be situations where this condition may be violated. Second, we present a model where only one product is being produced in a capacitated setting. A multi-product system would be more appropriate in a general setting; it will be more complex to handle, however.

6 References

Agarwal N., S. Smith and A.A. Tsay, Multi-Vendor Sourcing in a Retail Supply Chain, *Working Paper*, Santa Clara University, CA, 1999.

Azoury, K.S., Bayes Solution to Dynamic Inventory Models Under Unknown Demand Distribution, *Management Science*, 31(9), 1150-1160, 1985.

Bitran G.R., E.A. Haas and H. Matsuo, Production Planning of Style Goods with High Setup Costs and Forecast Revisions, *Operations Research*, 34(2), 226-236, 1986.

Buckley S., R. Jayaraman, G. Lin, N. Nayak, R. Srinivasan, J.M. Swaminathan and M. Taubenblatt, Responding to Demand Uncertainty Through Supplier Flexibility and Safety Stocks, *INFORMS*, Los Angeles, April 1995.

Fisher M.L. and A. Raman, Reducing the Cost of Demand Uncertainty through Accurate Response to Early Sales, *Operations Research*, 44(1), 87-99, 1996.

Gullu R., On the Value of Information in Dynamic Production/Inventory Problems under Forecast Evolution,*Naval Research Logistics*, 43, 289-303, 1996.

Gurnani H. and C.S. Tang, Optimal Ordering Decisions with Uncertain Cost and Demand Updating, *Working Paper*, Hong Kong University of Science and Technology, 1998.

Hausman W.H., Sequential Decision Problems: A Model to Exploit Existing Forecasters, *Management Science*, 16(2), B93-B110, 1969.

Hausman W.H. and R. Peterson, Multiproduct Production Scheduling for Style Goods with Limited Capacity, Forecast Revisions and Terminal Delivery, *Management Science*, 18(7), 370-383, 1972.

Heath D.C. and P.L. Jackson, Modeling the Evolution of Demand Forecasts with Application to Safety Stock Analysis in Production/Distribution Systems, *IIE Transactions*, 26(3), 17-30, 1994.

Kaminsky P. and J.M. Swaminathan, Utilizing Forecast Band Refinement for Capacitated Production Planning, *Manufacturing and Service Operations Management*, 3(1), 68-81, Winter 2001.

Lee H.L. and J. Whang, Value of Postponement, *Product Variety Management: Research Advances* (ed. Tech Ho and Chris Tang), Kluwer Publishers, Norwell, MA, 65-84, 1998.

Lee H.L., V. Padmanabhan and J. Whang, Information Distortion in a Supply Chain: The Bull Whip Effect, *Management Science*, 43(4), 546-548, 1997.

Morton T.E. and D.W. Pentico, The Finite Horizon Nonstationary Stochastic Inventory Problem: Near-myopic Bounds, Heuristics, Testing, *Management Science*, 41(2), 334-343, 1995.

Nurani R.K., R. Akella and J.G. Shanthikumar, Effects of Using Future Forecast Information on Myopic Policies, *1994 MSOM Proceedings*, Pittsburgh, PA, 1994.

Scarf H., Bayes Solutions of the Statistical Inventory Problem, *Annals of Mathematical Statistics*, 30(2), 490-508, 1959.

Simchi-Levi D., P. Kaminsky and E. Simchi-Levi, *Designing and Managing the*

Supply Chain, Irwin-McGraw Hill, NY, 1999.

Swaminathan J.M. and S.R. Tayur, Tactical Planning Models, *OR/MS Handbook on Supply Chain Management*, edited by Steve Graves and de Tok, Elvesier Publishers, 2001 (forthcoming).

Toktay L.B. and L.M. Wein, Analysis of a Forecasting Production Inventory system with Stationary Demand, *Management Science*, 47(9), 1268-1281, 2001.