A procedure for the optimization of the dynamic response of a Vendor Managed Inventory system

S.M. Disney and D.R. Towill

Logistics Systems Dynamics Group, Cardiff Business School, Cardiff University, Aberconway Building, Colum Drive, Cardiff, CF10 3EU, UK.

DisneySM@cardiff.ac.uk, Tel +44(0)29 2087 6083, Fax +44(0)29 2087 4301

Abstract

This paper considers the performance of a production or distribution-scheduling algorithm termed Automatic Pipeline, Inventory and Order Based Production Control System (APIOBPCS) embedded within a Vendor Managed Inventory (VMI) supply chain where the demand profile is deemed to change significantly over time. A dynamic model of the system using causal loop diagrams and difference equations is presented. The APIOBPCS ordering algorithm is placed within a VMI relationship and a near saturated search technique evaluates optimum solutions based on production adaptation cost, system inventory cost and distributors’ inventory costs. The procedure can also cope with supply chains that operate in a localized region (where small, frequent deliveries are possible) or on a global scale, where large batch sizes are needed to gain economies of scale in transport costs. Properties of the optimal systems are highlighted via various Bullwhip, customer service level and inventory cost metrics. Managerial insights are gained and a generic decision support system is presented for “tuning” VMI supply chains. An important feature of the optimization procedure is the ability to generate a number of competing ordering algorithm designs. Final selection of the “best” system is then made via managerial judgement on the basis of the simulated response to typical real-life demands. We finish with a discussion of how the procedure may be used in an industrial context to design and strategically manage VMI supply chains.

Key Words

Vendor Managed Inventory, APIOBPCS, Optimization, Production and Inventory Control

Introduction

Many companies are now compelled to improve supply chain operation by sharing demand and inventory information with suppliers and customers. Different market sectors have coined alternative terms covering essentially the same idea of Vendor Managed Inventory (VMI). This is a production/distribution and inventory control system where stock positions and demand rates are known across more than one echelon of the supply chain. It is this enhanced information which is used for the purposes of setting production and distribution targets. VMI comes in many different forms described by terms such as Synchronized Consumer Response, Continuous Replenishment Programs, Efficient Consumer Response, Rapid Replenishment (Cachon & Fisher, 1997), Collaborative Planning, Forecasting and Replenishment (CPFR), (Holmström, Framling, Kaipia & Saranen, 2000), and Centralized Inventory Management (Lee, Padmanabhan & Whang, 1997a) often being used, depending on the sector application, ownership issues and scope of implementation. However, in essence, they are all variants on the VMI theme. There is also an increasing amount of literature on the way such shared information can be utilized to advantage. This is based on findings from computer-based simulation packages, e.g., Mason-Jones & Towill (1997), Lambrecht & Dejonckheere (1999a and 1999b), Van Ackere, Larsen & Morecroft (1993), Waller, Johnson & Davis (1999), Kaminsky & Simchi-Levi (1998), and from OR based statistical theory, e.g., Chen, Ryan & Simchi-Levi (2000) and Lee, So & Tang (2000).
Here we use discrete control theory and simulation to design a VMI ordering system. Specifically, the performance of the VMI system when coupled with the APIOBPCS (Automatic Pipeline, Inventory and Order Based Production Control System) production scheduling system, (John, Naim & Towill, 1994), is investigated and optimized. The major assumptions made are that:

- The system is linear, thus all lost sales are backlogged and negative production orders may become negative hence excess inventory is returned without cost.
- demand may be either deterministic (when calculating inventory responsiveness) or stochastic (when calculating production adaptation costs), in which case a normal distribution that is independent and identically distributed is assumed.
- The production adaptation cost is proportional to the variance of the orders generated.
- The cost of holding inventory is proportional to the time weighted inventory deviation following a unit step input
- The ordering policy adapted is APIOBPCS

The APIOBPCS policy is actually a very general rule; for instance the order-up-to policy and many variants of the order-up-to model are a special case of APIOBPCS, Dejonckheere, Disney, Lambrecht & Towill, (2001). For example APIOBPCS mimics human behavior whilst playing the “Beer Game”, Sterman (1989) and Naim & Towill (1995) and is a general descriptor of much of UK industrial practice (Coyle 1977). Specific industrial applications are described in Olsmats, Edghill & Towill (1988) and del Vecchio & Towill (1990). The order-up-to policies are well known to be optimal in terms of inventory costs, Chen, Drezer, Ryan & Simchi-Levi (2000). However the more general APIOBPCS structure is capable of minimizing the sum of the inventory and production adaptation costs (Dejonckheere et al, 2001) and is therefore of more general application.

The structure of the paper is as follows: Firstly the VMI supply chain concept is described. Next the supply chains dynamic time based performance is quantified. A solution space search appropriate to APIOBPCS operation then evaluates most of the possible combinations of decision parameter values to identify good solutions to the production-ordering problem within VMI supply chains. The evaluation is based on two types of cost functions. One is a surrogate for production adaptation costs; the second is a surrogate for the cost of holding inventory throughout the VMI system. Thus, the system is designed to minimise inventory-holding costs and adaptation related costs covering the need to ramp production up and down to meet perceived needs. The latter may be seen as due to bullwhip type behavior, and the former due to buffering it via inventory positioning. Additionally the optimization routine allows for the investigation of different ratios of production adaptation and inventory costs (where production adaptation may be weighted more, equal to, or less than inventory costs as shown later in Table 1).

The effect of different delivery frequencies between the two echelons in the VMI supply chain is also investigated. As the move to a VMI scenario alters the fundamental structure of the supply chain ordering mechanism, new VMI ordering decision needs to be “re-tuned” or optimized. In effect, for supply chains with volatile consumer demands, the manner in which inventory is moved from the manufacturer to the distributor in the VMI relationship creates extra variation in the demand signal (this is because of the predictive element at the distributor) to which the manufacturer has to respond. This paper specifically investigates the manner in which a manufacturer needs to “re-tune” his production and distribution control system within a VMI context. Additionally good designs are found for the case of different transportation lead-times between the manufacturer and the distributor and for different ratios of production adaptation and inventory costs. Since the latter are rarely known with great accuracy, studying a range of possible ratios yields considerable insight at the systems design stage. The methodology thus enables the final choice of parameters to be selected by comparing competing designs responding to real-world demand signals.

In summary, an evaluation and optimization procedure is highlighted “to make the best of”, (or to optimize, Sterman 1991) the APIOBPCS ordering decision within the VMI context, in order to minimise inventory holding and production adaptation costs and account for transportation batch sizes by altering:

- the way in which the re-order point at the distributor is calculated,
- system forecasting parameters and
- feedback parameters within the production ordering decision algorithm.

The set of results from the optimization procedure are analysed and some general managerial insights are gained, against the constraints that, the real system is linear and our model is representative of the system. These insights include the effect of the transportation lead-time, forecasting constants, inventory and WIP feedback on the bullwhip effect and inventory responsiveness. An important feature of the optimization procedure is the ability to generate a
number of competing designs via changing the weighting between inventory and production adaptation costs. Management judgement then assists in determining the “best” solution. This can be cross-checked for robustness by conducting a sensitivity analysis.

Outline of the VMI system

To describe the VMI scenario; a manufacturer and a distributor collaborate to operate a particular VMI strategy. The consumer buys goods from the distributor’s stock. An important feature of the system is that the manufacturer manages the distributor’s stock. The distributor collects information on the downstream sales to the consumer that is used to provide a forecast of the future likely sales over the delivery lead-time. This forecast is used to set a reorder point, \( R \), which will be used to provide safety stock to ensure high availability of goods at the distributor. However when sales increase, the reorder point \( R \) should increase (assuming that the delivery lead-time is constant) so as to ensure high Customer Service Levels (CSL). So, \( R \) is based on a forecast generated by exponentially smoothing the customer sales over \( T_q \) time periods. This exponential forecast is then multiplied by a safety factor \( (G) \) that reflects the transportation lead-time between the two VMI echelons and the desired availability to determine \( R \). The distributor sales, inventory levels and reorder point are then passed to the manufacturer, who can then determine whether or not a delivery is required. When the distributor’s inventory is below \( R \), the manufacturer ships goods to inflate the distributor’s stock up to an Order-up-to point \( (O) \). However, by the time the delivery arrives the distributor’s stock level is not likely to equal the Order-up-to point, due to sales occurring since the time the re-order point triggered the dispatch.

The manufacturer then has the responsibility for determining how many products to make in order to balance the availability versus excess stock trade-off. This is done by summing the goods in-transit (GIT), the distributor’s inventory and the manufacturer’s finished goods inventory minus the re-order point \( R \), (This sum is termed the system inventory and is an important driver in VMI control). The system inventory is compared to a target system inventory, and a fraction \((1/T_i)\) of the error in system inventory is taken, summed together with a smoothed representation of demand (exponentially smoothed over \( T_a \) time units) and a fraction \((1/T_w)\) of the Work In Process (WIP) error to set the production targets. It is this particular way of using information that mimics the Automatic Pipeline, Inventory and Order Based Production Control Systems (APIOBPCS) principles (John et al., 1994). Remember, that although the inventory level refers to the inventory at the manufacturer and the distributor, the WIP level is specific to the manufacturer only.

It can be appreciated that the value of \( R \) should be dynamically updated to track the demand profile of the product. As the demand increases/decreases during the product lifecycle it is desirable to increase/decrease \( R \) so that the safety stock provides good customer service levels at the distributor without driving up inventory holding costs. \( R \) should be set as a multiple \((G)\) of average demand as forecasted by the exponential smoothing forecasting technique, such that there is an adequate cover on demand to protect the customer from the delivery lead-time. Any net changes in the stock re-order point, \( R \), therefore have to be added to customer sales. The traditional method, assuming a normally distributed demand, of calculating the re-order point \( R \) is to set \( R = \mu L + Z \sigma \sqrt{L} \), in time period \( i \), where \( Z \) is the standard normal variant or the number of standard deviations from the mean corresponding to the CSL required to provide the necessary cover, Wilkinson (1996). \( \sigma \) is the standard deviation of forecast error in time period \( i \), and \( L \) is the distribution lead-time plus one time unit to account for the order of events. \( \bar{D} \) is the average demand in time period \( i \). \( \bar{D} \) is determined via exponential smoothing (with constant \( T_q \)), as this is a method commonly used within the operations research community, (Chen, Ryan and Simchi-Levi, 2000). Therefore, \( R \) is related to a gain \((G)\) on average demand, via \( G = (\bar{D} L + Z \sigma \sqrt{L}) / \bar{D} \). Thus \( R \) can be reduced to a simple multiple \((G)\) of average demand \((\bar{D})\) such that there is an adequate cover on demand to protect the customer from the delivery lead-time. This means that, because any net changes made in the \( R \) have to be added to customer sales, the parameters that affect \( R \), (which have just been reduced to \( G \) and \( T_q \)) will have an impact on the dynamic performance of the system. Additionally, the constant APIOBPCS parameters \( T_a, T_i, T_w \) will also have an effect on the systems dynamic response. Thus, there is a need to identify “good” sets of algorithmic controller values \((T_a, T_i, T_q, T_w)\) for various values of \( G \) in order to reduce the sum of inventory costs and production adaptation costs. Additionally it is important to define “good” solutions for various ratios \((W)\) of production adaptation costs to inventory costs.

The philosophy of our approach to VMI is design summarised in the Input-Output diagram shown in Figure 1. The generic optimization process inputs are the VMI-APIOBPCS model, transfer function analysis capability, simulation software and a suitable objective function. For any given system application, the distributor re-order point gearing
(G), production delay (Tp) and costs weighting function (W) are specified as given. The optimization process then produces the recommended system design outputs. These are Tq (distributor sales averaging time); Ta (Factory Sales averaging time); Ti (Inventory adjustment time); and Tw (WIP adjustment time). The simulation capability may then be used to predict system response to any customer demand.

![Optimization process inputs](image)

**Figure 1. Input-output diagram of the VMI design process**

**Formal description of VMI as an integrated system**

The APIOBPCS structure can be expressed in words as, “Let the production targets be equal to the sum of an exponentially smoothed representation of demand (exponentially smoothed over Ta time units), plus a fraction (1/Ti) of the inventory error in echelon stock, plus a fraction (1/Tw) of the WIP error” (John et al, 1994). In a multi-echelon inventory decision-making environment, the inventory error corresponds to the desired inventory in the factory minus the actual inventory in the factory. However to extend the APIOBPCS model into VMI-APIOBPCS, the manufacturer’s finished goods now encapsulates the distributor’s inventory, the manufacturer’s finished goods as well as the stock in transport, minus the re-order point, R. Defining this as the system stock, the VMI-APIOBPCS can be expressed in words as, “Let the production targets be equal to the sum of an exponentially smoothed representation of demand (exponentially smoothed over Ta time units), plus a fraction (1/Ti) of the inventory error in system stock, plus a fraction (1/Tw) of the WIP error”. Note the difference in the stock calculation in the two scenarios, i.e. VMI compared with traditional two-echelon control. Figure 2 illustrates the flow of information in the VMI situation.

The previous verbal description can now be turned into a causal loop diagram as shown in Figure 3. This is particularly useful as a set of block diagrams and difference equation models can be developed directly from it. Difference equations can be used to develop a spreadsheet-based model of the system. The required equations are shown in Appendix A. Block diagrams are particularly powerful mathematical models of the system that can also be derived from the causal loop diagram and can be used to investigate system performance. More details on this can be found in Disney (2001), but are not presented here for sake of brevity. Note that as stated above that the average consumer demand (CONS) is estimated via a first order exponential smoothing function. The net change in this estimate, from one time period to the next, is then added to the consumer demand. This reflects the distributor’s effect on the demand signal in the VMI scenario and is a unique aspect of VMI identified by this contribution.
VMI Controls

Business Targets

Factory

Total System Stock (SINV)

Production lead-time

Production Orders

Despatches

Infinite Material Stocks

Distribution lead-time

Factory's Order Rate (ORATE)

Despatch Times

Sales (CONS)

Delivery

Re-order Point, R

Order-up-to-Point, OD

Incremental Change in Re-order Point, dSS

Virtual Sales, VCON

Virtual Sales, AVCON

Distributor (VMI)

Distributors Time to Average Sales, Tq

Distributors Safety Stock Gain, G

System Inventory, AINV

Factory’s Time to Average Sales, Ta

Time to Adjust for WIP Errors, Tw

Time to Adjust for Inventory Errors, Ti

Production Delay Estimate, TP

Manufacturer (APIOBPCS)

Production Delay, Tp

Desired WIP, DWIP

Target System Inventory, TINV

Completion Rate, COMRATE

WIP

Average Virtual Sales, AVCON

Factory’s Order Rate (ORATE)

Time to Adjust for Inventory Errors, Ti

Production Delay Estimate, TP

Desired WIP, DWIP

Target System Inventory, TINV

Completion Rate, COMRATE

WIP

Average Virtual Sales, AVCON

Factory’s Time to Average Sales, Ta

Time to Adjust for WIP Errors, Tw

Time to Adjust for Inventory Errors, Ti

Production Delay Estimate, TP

Manufacturer (APIOBPCS)

Distributor (VMI)

Distributors Time to Average Sales, Tq

Distributors Safety Stock Gain, G

System Inventory, AINV

Factory’s Time to Average Sales, Ta

Time to Adjust for WIP Errors, Tw

Time to Adjust for Inventory Errors, Ti

Production Delay Estimate, TP

Manufacturer (APIOBPCS)

Distributor (VMI)

Distributors Time to Average Sales, Tq

Distributors Safety Stock Gain, G

System Inventory, AINV

Factory’s Time to Average Sales, Ta

Time to Adjust for WIP Errors, Tw

Time to Adjust for Inventory Errors, Ti

Production Delay Estimate, TP

Manufacturer (APIOBPCS)

Distributor (VMI)

Distributors Time to Average Sales, Tq

Distributors Safety Stock Gain, G

System Inventory, AINV

Factory’s Time to Average Sales, Ta

Time to Adjust for WIP Errors, Tw

Time to Adjust for Inventory Errors, Ti

Production Delay Estimate, TP

Manufacturer (APIOBPCS)

Distributor (VMI)

Distributors Time to Average Sales, Tq

Distributors Safety Stock Gain, G

System Inventory, AINV

Factory’s Time to Average Sales, Ta

Time to Adjust for WIP Errors, Tw

Time to Adjust for Inventory Errors, Ti

Production Delay Estimate, TP

Manufacturer (APIOBPCS)
Dynamic response of the VMI system

An illustration of the dynamic response of the system to the unit step in consumption is useful at this stage. Such responses provide “rich pictures” of system behavior. In order to economize on the number of figures herein, those selected illustrate and anticipate the good dynamic performance discovered using an optimization technique developed in a later section. Thus the responses shown serve two purposes: firstly to highlight how the dynamic performance changes with various parameter settings as required to illustrate the text in this section and secondly to describe the optimal performance as given by our design technique.

Production order rate responses of the VMI-APIOBPCS system to a step input are shown in Figure 4 for various values of G, although we only need consider one of these responses now. A unit step input is a particularly powerful test signal that control engineers use to determine many properties of the system under study. For example, the step is simply the integral of the impulse function, thus understanding the step response automatically allows insight to be gained on the impulse response. This is very useful as all discrete time signals may be decomposed into a series of weighted and delayed impulses. Additionally the impulse response contains important frequency domain information that may be used to gain insights into responses to real-life random demand patterns (Dejonckheere et al, 2001). Within a supply chain context, the step response may be thought of as a genuine change in the mean demand rates, (for example, as a result of promotion or price reductions).

A detailed description of the individual controller contributions to ORATE within a single echelon APIOBPCS structure can be seen in Disney, Naim and Towill (1997). Hence only the effect of the distributor parameters needs to be presented here. Manifestly it can be seen that it takes time for the VMI system to adjust to the changes in downstream demand. Thus the system temporarily over produces in order to recover inventory deviations as directly observable in Figure 5 that shows the corresponding inventory responses for the same system.

Figure 4. ORATE response to a step input with optimum parameters when W=1 for various G

Figure 5. AINV response to a step input with optimum parameters when W=1 for various G.
Figure 4 shows that, as should be expected, reducing the transportation lead-time results in a better dynamic response. This is because as lower values of G (which can be used when there are short transportation lead-times) produce a response that has less overshoot to the step input and a shorter settling time. Both are desirable features in system control. This finding is also supported by Figure 5, which shows that there is significantly less inventory required for the optimum designs with lower cover (G) required on average demand when setting the distributor’s re-order point.

**VMI performance metrics**

In this section the optimal VMI-APIOBPCS control parameters will be determined. This methodology will ensure that the ordering system will minimize the sum of:

- the inventory holding costs at the distributor and the manufacturer, i.e. the total VMI holding costs
- the production adaptation costs at the manufacturer. The production adaptation costs are the costs associated with a variable production ordering rate, i.e. those associated with the Bullwhip Effect (Lee, Padmanabhan & Whang 1997a and 1997b).

Each will now be discussed in turn.

**The production adaptation cost metric**

The noise bandwidth ($\omega_n$) is traditionally a useful measure to characterize the frequency response of a system and hence the production adaptation costs. It is defined as the area under the squared frequency response of the system, Equation 1. Very importantly the noise bandwidth $\omega_n$ is a performance measure that is proportional to the variance of the ORATE when sales consist of pure white noise (constant power density at all frequencies) Garnell & East (1977) and Towill (1999). In conventional OR terms, pure white noise maybe interpreted as an independently and identically distributed (i.i.d.) normal distribution. Thus the noise bandwidth is directly equivalent to the Coefficient of Variation measure used by many authors to quantify the Bullwhip Effect when demand is i.i.d., Dejonckheere et al (2001). Stalk and Hout, (1990), state that the cost of variable production schedules is proportional to the cube of the variance of the schedule which under some conditions this approximates to a square law. Thus, the noise bandwidth may be reasonably considered a surrogate metric for production adaptation costs. These costs may include such things as such as hiring/firing, production on-costs, over-time, increased raw material stock holdings, obsolescence, lost capacity etc.

We make use of Shannon’s Sampling Theorem, (Shannon, Oliver & Pierce, 1948) which states that sampled data systems can only detect inputs of frequencies up to half the Nyquist Frequency of $\pi$ radians per sampling period from the Amplitude Ratio plot alone due to aliasing, hence the integral is only required for the frequencies 0 to $\pi$. In order to calculate the noise bandwidth the frequency response may be calculated at set values of $w$ selected up to the Nyquist frequency. $w_n$ is thus estimated via numerical integration with strips of 0.001 radians per time period. The system ORATE transfer function required can easily be derived from the causal loop diagram via the block diagram as in Disney (2001), but is omitted here for brevity. Hence ORATE can be calculated directly from the transfer function so simulation is unnecessary when estimating;

$$w_n = \int_{0}^{\pi} \left[ ORATE_f \right]^{2} \, dw, \quad \text{......... Eq. 1}$$

Figure 6 illustrates the $w_n$ for various values of Ta and Ti for the particular case when G=1 and W=1. It can be seen that the Noise Bandwidth of the system decreases as the forecast parameter (Ta) and as the inventory error controller (Ti) is increased. In other words a smoother dynamic response will be produced when the average age of the forecast is increased and smaller fractions of inventory errors are recovered over the sampling period. This is expected from the analysis by Towill (1982), since the feed-forward and feedback controls are both adjusted to produce the “best” system performance. Note that the “best” parameter settings selected therein are conservative with little bullwhip induced by the system.
The distributor’s inventory recovery metric

As described earlier, the distributor monitors and exponentially smooths sales. This is to determine a forecast of the likely future sales that in turn is used to set the re-order point R in a manner that ensures high CSL. The speculative stock held at the distributor should naturally be reflective of end consumer demand. Thus a metric is needed that reflects how well the safety stock target reflects the actual consumption. Also it needs to account for the cover provided on actual consumption to buffer the consumer from the delivery frequency (CONS*G). To obtain a realistic measure, a step change in customer demand is assumed. There will be an initial transient error in the reorder point calculation. A quantification of its magnitude will be representative of the responsiveness of the system. This can be quantified by taking the Integral of Time * Absolute Error (ITAE) (Graham and Lathrop, 1953) of this error response. Thus the ITAE is a measure of the amount of inventory that needs to be held in order to be able to cope with a step change in demand. A smaller value of the ITAE will imply that less stock is required to buffer the demand, and vice versa. Thus the ITAE may be considered to be a surrogate inventory cost metric.

Now ITAE is generally agreed to be the most intuitive criterion following a step, for assessing transient recovery, as there is inevitably a large error is present shortly after the step and the ITAE penalizes more heavily, errors that are present later, by a suitable weighting in the time domain, (Towill 1970). Furthermore ITAE also penalizes positive and negative errors equally, and is the simplest composite measure that is reliable, applicable and selective, (Graham & Lathrop 1953). So ITAE is used in our optimization procedure via the defined in Equation 2 as follows;

\[
 ITAE_{VCON} = \sum_{t=0}^{\infty} \frac{|E_{VCON}|t}{a}, \text{ Eq. 2.}
\]

where, \( t \) = time period, debased at step change,
\(|E_{VCON}|\) = modulus of the error in speculative safety stock targets at the distributor,
\( a \) = weight to scale the metric to a similar magnitude of the production adaptation costs, nominally set at 250 after some initial experimentation.

Throughout this paper the ITAE was calculated following a unit step input in CONS that increased from 0 to 1 widgets per time period at time = zero. The ITAE\(_{VCON}\) for various settings of G and Tq can be seen in Figure 7. It
shows that $G$ has the major impact of the reflectiveness of the distributors stock level target rather than the parameter defining the average age of the exponential forecast ($T_q$).

Figure 7. Reflectiveness of the distributors safety stock level as defined by the ITAE<sub>VCON</sub>

The system inventory recovery metric

It is desirable to satisfy demands (at the distributor and at the factory) from stock so the dissatisfied customer will not have to wait for product to be delivered. It is also important to ensure that the system inventory (i.e. the total inventory at the factory, goods in transit and inventory at the distributor, minus $R$) does not deviate from a target level of stock holding. If large deviations occur, large target stock positions will have to be assumed to achieve acceptable levels of availability. Specifically, we are looking for parameter settings that will ensure the system inventory level will recover quickly from errors in the target inventory levels that result from changes in demand levels. Again the ITAE following a unit step in demand is used but in this particular metric we will now be measuring the ITAE of the system inventory as defined by Equation 3 as follows:

$$ITAE_{INV} = \sum_{t=0}^{\infty} \frac{|E_{INV}| \cdot t}{b}, \text{ Eq. 3.}$$

where, $t =$ time period, debased at step change,

$|E_{INV}| =$ modulus of the error in system stock levels,

$b =$ weight to scale the metric to a similar magnitude of the production adaptation costs, nominally 500.

Equation 3 yields a metric that captures supply chain inventory costs and is illustrated by Figure 8. This graph shows that the system inventory recovers well with small values of $T_i$ (the rate at which inventory deviations are corrected), but if $T_i$ is set too low then the inventory recovery performance rapidly decreases. This is an important managerial insight that has been further explored by Dejonckheere et al (2001). The feed-forward path (the forecast) seems to have a less important role here. But generally inventory performance is improved with lower values of $T_a$.

The objective function to be maximized

As the three performance criteria outlined above are all defined as the smaller the better, the reciprocal of the Euclidean distance from zero in three-dimensional space is used as the overall performance vector. This vector is thus shown in Equation 4 below. The individual objectives have been weighted by experiment to initially balance contributors and thereby ensure that for the particular case when $W=1$, one metric does not swamp another.
Figure 8. System inventory recovery metric as defined by the ITAE

As different values of $G$ can be used within our VMI system, the objective is to maximize the score as given by Equation 4 below of the system for different values of $G$. This was done by conducting a near full solution space search using the algorithm shown in Appendix B. Equation 4 shows that the production adaptation cost metric incorporates a scaling factor, $W$. This is nominally set to unity to denote equal emphasis on production adaptation costs and inventory costs. However, $W$ has also been set to a range of other values to investigate the optimum solutions of inventory cost sensitive scenarios and production adaptation cost sensitive scenarios. Changing $W$ to generate a range of scenarios for further comparison is directly analogous to changing the weighting function used in “modern” control theory by Christensen & Brogan (1971), for exactly the same purpose.

\[
\text{score} = \frac{1}{\int_0^\infty \left( \frac{\partial \text{RATE}_j}{\partial \delta W} \right)^2 W \, dt} + \frac{\sum_{i=0}^\infty |E_{\text{AINV}}|^2}{500} + \frac{\sum_{i=0}^\infty |E_{\text{VCON}}|^2}{250}, \quad \ldots \text{Eq. 4.}
\]

The full optimization procedure is shown in detail in Appendix B. However a summary of the procedure is as follows:

- Set up VMI transfer function model for frequency response
- Set up integral of frequency response calculation
- Set up search procedure
  1. Define variables
  2. Get parameter ranges, $G$ and $W$ and the depth of search from user
  3. Reset initial conditions
  4. Calculate the system inventory cost via the ITAE of AINV in response to a step input via a difference equation model
  5. Calculate the distributor inventory cost via the ITAE of VCON in response to a step input via a difference equations model
  6. Calculate the production adaptation costs by calling the integral of the frequency response calculation
  7. Iterate around the loop 3-6, evaluating sets of parameter values, remembering the “best” set
Repeat the search procedure for difference values of G and W

Because the difference between 1/x and 1/(x+1) is small for large values of x to save analysis time, the density of the search was reduced for large values of the parameters (such as Ta, Ti, Tq, and Tw). Specific details of this can be determined from Appendix B, which lists the “Mathematica” (Wolfram Research, Champaign, IL) source code used for this optimization procedure.

**Optimization Results**

The results from the application of the optimization procedure are illustrated in Table 1. The results given are for the “best” designs for set values of the Weighting Function W, and the re-order point gearing G. They clearly show that the dynamic performance of the system is very much related to the value of G used to cover the delivery frequency between the distributor and the manufacturer. Note that less frequent deliveries from factory to distributor require larger values of G and hence the dynamic performance of the system decreases. However, for larger G’s the transport batch size will be increased and transport costs are lowered, so there are further trade-off’s to consider. If the delivery lead-time is known and fixed, G is then set to ensure the availability of goods from stock at the distributor. This can be easily done by using a traditional reorder point mechanism, as described earlier.

It is important to stress that it is the ratios of the APIOBPCS parameter (Ta, Ti and Tw) values to the production delay (Tp+1) that leads to generic solutions. Hence these ratios should be considered when using Table 1 to “tune” real-world VMI systems with different production lead-times. For example, if a particular production delay is 8 sampling periods and G is set equal to unity, then the appropriate parameters (based on a qualitative judgment of the dynamic response) would be Ta=(6/5)*9, Ti=(7/5)*9 and Tw=(42/5)*9. This scaling requirement is often overlooked by the OR community. Moreover, note that this scaling issue does not apply to the distributor parameters Tq and G, which operate independently from the production delay and only depend on the delivery lead-time between the manufacturer and the distributor.

Now Table 1 can be used as a decision support system to design a VMI ordering system as follows. Suppose that in a given situation we have as is given the following parameters;

- the transportation lead-time is four time periods
- a suitable customer service level can be achieved by setting G to 8,
- the production adaptation costs are considered to be slightly more important than the two inventory based costs,

Then we proceed to investigate the cell in the column when G=8 and the row when W = (say) 5. This gives a set of numbers (7,27,6,63,0.14502), which are the parameter values for Ta, Ti, Tq, and Tw followed by the “score” achieved via the optimisation procedure the particular parameters. In practice the system designer would then undertake a cross-check to display important dynamic responses associated with this parameter set. In particular the systems designer should make a qualitative judgment on the suitability of the dynamic response in terms of inventory and production adaptation costs, whilst endeavoring to match the system response to economics of the value stream.

It is noted that the “slice” of the solution space used in Table 1 suggests that the result when W=0.01 and G=4 appear somewhat strange. Therefore further investigation and conformation for values close to this case are presented in Disney (2001) that verifies that this result is indeed correct.

**Insights drawn from the optimization results**

Summarizing the optimization results shown in Table 1 in the manner shown in Figures 9 to 12 enables some general managerial insights on VMI design to be gained. Specifically:

- Figure 9 shows that if high production adaptation costs or infrequent deliveries are present, factory demand should be forecasted via a larger smoothing constant. This is an intuitively obvious result as by slowing down forecasts smoother dynamic responses are expected.
- Figure 10 implies using a longer smoothing history to set the re-order point at the distributor in the presence of high production adaptation costs. Again this is as expected and for the same reason as before.
Table 1. Optimization results for different weights between production adaptation and inventory costs (W) and different delivery frequencies (G).

The results are presented in the format of [Ta,Ti,Tq,Tw,Score].
Figure 11 demonstrates that it is likely to be beneficial if high production adaptation costs are present to use WIP information in the VMI scenario (due to there being costs to enable WIP information to be collected). It is intuitive that WIP information would be beneficial in situations where production adaptation costs are higher. Additionally the graph also shows that it is important to use WIP information when inventory costs are heavily weighted if there are infrequent deliveries between the two echelons (i.e. when G is large).

Figure 12 shows that inventory errors should be recovered slowly if infrequent deliveries or high production adaptation costs are present. This again is an obvious result in accordance with our previous intention.

Experiments in which a number of variables are held constant whilst the effects of others are determined are helpful in understanding system behavior. But they are not a substitute for using the power of the optimization process across a wide range of variables. However, Figures 9 through 12 are helpful in deciding how many alternative designs may be considered before the final parameter selection. They may also highlight potential areas of conflict when there are constraints imposed on system performance such as restrictions inherent in the production and distribution functions.

**Bullwhip and CSL estimation in VMI systems**

It is useful at this stage to highlight the performance of the various optimum designs via some key performance indicators as illustrated in Tables 2 to 6. System performance is highlighted by illustrating the performance of the nominally weighted scenario (W=1), the case where production adaptation costs are five times as important (W=5), the case where inventory holding costs are five times as important (W=0.2) and the case where ratio is 100 to 1 (W=0.01 and W=100), for a range of values of G. Table 7 highlights the performance characteristics for optimal systems with differing values of W (W=0 to $\infty$) when G =1. This has been normalized to emphasize the available range of solutions.

The key performance indicators can be broadly split into two areas, the first in relation to the Bullwhip Effect, and the second to CSL. The following Bullwhip measures have been selected:

- **Peak ORATE overshoot to a unit step input.** This is a useful measure to identify the transient response in the time domain to a change in demand rates.
- **Maximum amplitude in the ORATE frequency response.** This describes the worst-case scenario of a systems performance at amplifying the Bullwhip Effect as the system performance at the resonant frequency is highlighted.
- **Noise Bandwidth.** As described earlier, the noise bandwidth is a useful measure of production adaptation costs.
- **Coefficient of Variation.** In response to real inputs the Coefficient of Variation measure (Chen, Ryan & Simchi-Levi, 2000) is often used by the OR community to quantify the Bullwhip Effect as follows.

\[
\text{Bullwhip} = \frac{\text{VAR(ORATE)}}{\text{VAR(CONS)}}, \quad \text{Eq 5.}
\]

The Coefficient of Variation is defined as the ratio of the variance of the output (ORATE) to the variance of the input (CONS). A random, normally distributed demand signal was used in this bullwhip measure. The signal had a mean of 16.003 and a variance of 2.544 for 100 time periods, thus it may not be i.i.d. (if it were, this simulation based coefficient of variation measure would be exactly equal to the noise bandwidth measure dividing by $\pi$, (Dejonckheere et al, 2001)). This metric was realized by developing a spreadsheet model of the difference equations shown in Appendix A.

To evaluate CSL the measures chosen were:

- **% Availability.** Using the same demand signal as for the Co-efficient of Variation measure above, the % availability was estimated via as the percentage of time system inventory levels was below zero when the target inventory position was set to equal the mean demand, as first outlined by Cheema (1994).
- **Maximum deficit following a step increase in demand.** The maximum deficit in actual systems inventory is a useful measure of system responsiveness to changing demands.
- **ITAE.** As described earlier, the ITAE of the systems inventory is a useful measure of inventory recovery.
<table>
<thead>
<tr>
<th>Type</th>
<th>Design</th>
<th>$\phi = G=1$</th>
<th>$\phi = G=2$</th>
<th>$\phi = G=4$</th>
<th>$\phi = G=8$</th>
<th>$\phi = G=16$</th>
<th>$\phi = G=32$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bullwhip measures</td>
<td>Peak ORATE overshoot to a unit step</td>
<td>2.5</td>
<td>2.51</td>
<td>2.4541</td>
<td>3.5417</td>
<td>4.4545</td>
<td>5.83</td>
</tr>
<tr>
<td></td>
<td>Max amplitude in frequency response</td>
<td>3.0425</td>
<td>2.99308</td>
<td>2.88697</td>
<td>5.2254</td>
<td>6.68916</td>
<td>8.75075</td>
</tr>
<tr>
<td></td>
<td>Noise Bandwidth</td>
<td>17.3271</td>
<td>16.5514</td>
<td>15.5854</td>
<td>27.8812</td>
<td>35.3112</td>
<td>54.7947</td>
</tr>
<tr>
<td></td>
<td>Co-efficient of variation</td>
<td>5.3884</td>
<td>5.1689</td>
<td>4.875</td>
<td>8.9898</td>
<td>11.6292</td>
<td>18.1556</td>
</tr>
<tr>
<td>CSL measures</td>
<td>% Availability</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>88.45</td>
<td>77.5</td>
<td>65.56</td>
</tr>
<tr>
<td></td>
<td>Maximum deficit following a unit step</td>
<td>-5.9688</td>
<td>-6.9375</td>
<td>-8.875</td>
<td>-12.75</td>
<td>-18.89</td>
<td>-36</td>
</tr>
<tr>
<td></td>
<td>ITAE</td>
<td>0.2680</td>
<td>0.4440</td>
<td>2.3185</td>
<td>0.7762</td>
<td>1.3371</td>
<td>3.8734</td>
</tr>
</tbody>
</table>

Table 2. Properties of optimal systems when $W=0.01$

<table>
<thead>
<tr>
<th>Type</th>
<th>Design</th>
<th>$\phi = G=1$</th>
<th>$\phi = G=2$</th>
<th>$\phi = G=4$</th>
<th>$\phi = G=8$</th>
<th>$\phi = G=16$</th>
<th>$\phi = G=32$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bullwhip measures</td>
<td>Peak ORATE overshoot to a unit step</td>
<td>1.831</td>
<td>1.966</td>
<td>1.78232</td>
<td>1.81621</td>
<td>2.788</td>
<td>2.699</td>
</tr>
<tr>
<td></td>
<td>Max amplitude in frequency response</td>
<td>2.25446</td>
<td>2.4633</td>
<td>2.01439</td>
<td>2.0118</td>
<td>3.57042</td>
<td>3.0209</td>
</tr>
<tr>
<td></td>
<td>Co-efficient of variation</td>
<td>1.0113</td>
<td>1.2366</td>
<td>1.1513</td>
<td>1.4991</td>
<td>2.1795</td>
<td>2.7591</td>
</tr>
<tr>
<td>CSL measures</td>
<td>% Availability</td>
<td>100</td>
<td>99.8</td>
<td>99.8</td>
<td>96.87</td>
<td>84.15</td>
<td>70.65</td>
</tr>
<tr>
<td></td>
<td>ITAE</td>
<td>0.699956</td>
<td>0.8518</td>
<td>2.96105</td>
<td>10.5947</td>
<td>4.00452</td>
<td>25.5316</td>
</tr>
</tbody>
</table>

Table 3. Properties of optimal systems when $W=0.2$

<table>
<thead>
<tr>
<th>Type</th>
<th>Design</th>
<th>$\phi = G=1$</th>
<th>$\phi = G=2$</th>
<th>$\phi = G=4$</th>
<th>$\phi = G=8$</th>
<th>$\phi = G=16$</th>
<th>$\phi = G=32$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bullwhip measures</td>
<td>Peak ORATE overshoot to a unit step</td>
<td>1.69</td>
<td>1.75</td>
<td>1.70</td>
<td>1.69</td>
<td>1.85</td>
<td>2.22</td>
</tr>
<tr>
<td></td>
<td>Max amplitude in frequency response</td>
<td>2.02</td>
<td>2.1</td>
<td>1.93</td>
<td>1.87</td>
<td>2.02</td>
<td>2.49</td>
</tr>
<tr>
<td></td>
<td>Noise Bandwidth</td>
<td>1.48115</td>
<td>1.61538</td>
<td>1.90241</td>
<td>2.07832</td>
<td>2.63246</td>
<td>3.00962</td>
</tr>
<tr>
<td></td>
<td>Co-efficient of variation</td>
<td>0.4974</td>
<td>0.5443</td>
<td>0.6321</td>
<td>0.6882</td>
<td>0.8750</td>
<td>1.0275</td>
</tr>
<tr>
<td>CSL measures</td>
<td>% Availability</td>
<td>100</td>
<td>100</td>
<td>99.2</td>
<td>97.85</td>
<td>88.65</td>
<td>80.23</td>
</tr>
<tr>
<td></td>
<td>ITAE</td>
<td>1.48753</td>
<td>1.5176</td>
<td>1.57631</td>
<td>1.76783</td>
<td>2.64735</td>
<td>4.12833</td>
</tr>
</tbody>
</table>

Table 4. Properties of nominally weighted optimal systems ($W=1$)

<table>
<thead>
<tr>
<th>Type</th>
<th>Design</th>
<th>$\phi = G=1$</th>
<th>$\phi = G=2$</th>
<th>$\phi = G=4$</th>
<th>$\phi = G=8$</th>
<th>$\phi = G=16$</th>
<th>$\phi = G=32$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bullwhip measures</td>
<td>Peak ORATE overshoot to a unit step</td>
<td>1.488</td>
<td>1.4837</td>
<td>1.53168</td>
<td>1.56213</td>
<td>1.66264</td>
<td>1.9471</td>
</tr>
<tr>
<td></td>
<td>Max amplitude in frequency response</td>
<td>1.65016</td>
<td>1.64274</td>
<td>1.70452</td>
<td>1.71692</td>
<td>1.84925</td>
<td>2.2051</td>
</tr>
<tr>
<td></td>
<td>Noise Bandwidth</td>
<td>0.9052</td>
<td>0.904365</td>
<td>1.011884</td>
<td>1.15326</td>
<td>1.27588</td>
<td>1.60367</td>
</tr>
<tr>
<td></td>
<td>Co-efficient of variation</td>
<td>0.3087</td>
<td>0.3067</td>
<td>0.3448</td>
<td>0.3885</td>
<td>0.4218</td>
<td>0.5331</td>
</tr>
<tr>
<td>CSL measures</td>
<td>% Availability</td>
<td>100</td>
<td>100</td>
<td>99.41</td>
<td>99.43</td>
<td>97.46</td>
<td>87.26</td>
</tr>
<tr>
<td></td>
<td>ITAE</td>
<td>3.1644</td>
<td>4.1392</td>
<td>5.62214</td>
<td>15.6780</td>
<td>32.6348</td>
<td>66.0079</td>
</tr>
</tbody>
</table>

Table 5. Properties of optimal systems when $W=5$
Inspection of the simulation results of the optimal systems show that as the more weight is given to the production adaptation costs, Bullwhip reduces (however it is measured). Additionally, reducing the transportation lead-time between the manufacturer and the distributor reduces Bullwhip. This verifies the Time Compression paradigm, Towill (1996). Also reducing the distribution lead-time increases all of the inventory responsiveness or CSL metrics. As expected the CSL measures all degrade as more emphasis is placed on the production adaptation costs. Thus, managers may use the decision support system in Table 1 with some confidence.

### Making the final selection

Table 7 compares the properties of the optimal systems generated by varying W when G=1. To highlight these “best” values these results have been normalized. For example, if we are using ORATE peak as our bullwhip measure, then the increase as W is varied is greater than 6:1. On the other hand selecting W=1 gives an ORATE peak only about 50% greater than the minimum (for W=0). Note that the standard deviation ratios calculated from noise bandwidth (\( \sigma_n \)) and from the coefficient of variation (\( \sigma_{CV} \)) are, as expected, in close agreement. So either may be used with confidence to estimate the range in ORATE amplitude observable in response to random demand. Again comparing W=1 and W=0 solutions, we see that the amplitude range is increased by nearly 4:1 with respect to the minimum value. So the normalized format is very helpful in determining simple rules of thumb relating cause-and-effect.

<table>
<thead>
<tr>
<th>Type</th>
<th>Design</th>
<th>@ W=0</th>
<th>@W=0.01</th>
<th>@W=0.05</th>
<th>@W=1</th>
<th>@W=5</th>
<th>@W=20</th>
<th>@W=10</th>
<th>@W=∞</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bullwhip</td>
<td>Peak ORATE overshoot to a unit step</td>
<td>6.059</td>
<td>2.164</td>
<td>1.749</td>
<td>1.585</td>
<td>1.463</td>
<td>1.288</td>
<td>1.151</td>
<td>1.053</td>
</tr>
<tr>
<td></td>
<td>Max amplitude in frequency response</td>
<td>10.979</td>
<td>2.569</td>
<td>2.255</td>
<td>1.904</td>
<td>1.706</td>
<td>1.394</td>
<td>1.200</td>
<td>1.073</td>
</tr>
<tr>
<td></td>
<td>Co-efficient of variation</td>
<td>2605.650</td>
<td>158.950</td>
<td>54.802</td>
<td>29.832</td>
<td>14.673</td>
<td>9.106</td>
<td>5.245</td>
<td>2.307</td>
</tr>
<tr>
<td>CSL</td>
<td>% Availability</td>
<td>99.6086</td>
<td>99.02</td>
<td>98.24</td>
<td>97.06</td>
<td>96.48</td>
<td>92.37</td>
<td>90.27</td>
<td>87.34</td>
</tr>
<tr>
<td></td>
<td>ITAE</td>
<td>39.9941</td>
<td>49.84</td>
<td>60.236</td>
<td>84.173</td>
<td>156.160</td>
<td>145.21</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6. Properties of optimal systems when W=100

Figure 13 shows the ORATE responses of the optimal system when G=1 and when W=1, W=0.01 and W=100. The system was driven by the previously mentioned normally distributed random demand signal (with a mean of 16.003 and a variance of 2.544). Note that the range of the responses varies as expected from Table 7: the range is proportional to \( \sigma_n \). The effect of W on the required production variation is clearly shown, with larger W producing a smoother order rate. This is a good example of the use of the optimization procedure to produce a set of competing designs. Then the “best” of the competing designs is selected to match the business strategy of the value stream. In
current management language this strategy can cover the complete spectrum from “agile” to “lean” supply (Christopher & Towill, 2000). For example, the practical implication in the selection of the optimum VMI parameters corresponding to \( W = 0.01 \) results in an “agile” system with high capacity requirements. In contrast selecting the optimum parameters corresponding to \( W = 100 \) generates a “lean” type response with quasi-level scheduling. Setting \( W = 1 \) is a good compromise design requiring relatively smooth changes in production levels coupled with the need for minimum reasonable system inventory.

![Figure 13. Various optimal (when G=1) ORATE responses to a random demand with different inventory/production weights (W). Note W=100 yields a “level schedule” type design; W=0.01 yields an “agile” type design; W=1 gives a good compromise design](image)

If required the optimization process may be used in an iterative mode to make the final selection of system parameters. So responses such as Figure 13 can be repeated for a range of “weights” inventory and production adaptation costs. In this context Tables 2 through 7 provide ideas on how the parameters affect bullwhip and other system performance criteria. The iterative mode is summarized in Figure 14 and charts the system design starting at the business context level and ending with VMI implementation. Note that there are two feedback loops in the design methodology. The VMI design loop assumes that production and distribution processes are fixed and gives the “best” design under these particular conditions. However, if the “best” design is still not good enough, then either the production facility, or the distribution facility (or both) must be re-engineered to remove these constraints on VMI performance.

**Conclusions**

This paper has designed a VMI system for various different ratios of production adaptation costs and inventory holding costs. A decision support system has been proposed to determine the optimum design parameters in the VMI-APIOBPCS system by relating it to the gain on demand when setting safety stocks at the distributor via a look-up table (Table 1) and summarized in Tables 2 to 7. The benefits of achieving frequent deliveries between the customer and the supplier can be clearly seen. This is because dynamic performance, judged by both via the evaluation metric and via other Bullwhip and CSL measures, improve significantly as \( G \) decreases, i.e. as more frequent deliveries are made. Some managerial insights have been gained consequential to the use of information within the VMI scenario. The paper offers guidelines for setting VMI system “controller” values to minimise particular profiles of production adaptation costs and inventory holding costs.

Thus, this contribution has presented a Decision Support System that allows the appropriate “tuning” of a VMI system’s parameters across the whole spectrum inventory and production adaptation cost-sensitive supply chains. This is so whether they operate in a localized region (where more frequent deliveries are economical, and hence there are low values of \( G \)), or on a global scale (where distribution costs are significantly higher), and where high
values of G are needed to gain economies of scale in transport costs. The optimization procedure can be used to search out a range of competitive systems that can then be compared for effectiveness in response to simulated “real-world” inputs. The final choice of design parameters would then be based on the results of the dummy “real-world” tests. Finally, we have highlighted how the VMI design procedure may be used within an industrial context.

Figure 14. Using the VMI optimization routine within the business context

Nomenclature

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AINV</td>
<td>Actual systems inventory</td>
</tr>
<tr>
<td>APIOBPCS</td>
<td>Automatic Pipeline, Inventory and Order Based Production Control System</td>
</tr>
<tr>
<td>AVCON</td>
<td>Average Virtual Consumption.</td>
</tr>
<tr>
<td>COMRATE</td>
<td>Completion Rate</td>
</tr>
<tr>
<td>CONS</td>
<td>Consumption or Market Demand</td>
</tr>
<tr>
<td>CSL</td>
<td>Customer Service Levels</td>
</tr>
<tr>
<td>DES</td>
<td>Dispatches between the Manufacturer and the Distributor</td>
</tr>
<tr>
<td>DINV</td>
<td>Distributors Inventory Holding</td>
</tr>
<tr>
<td>DSS</td>
<td>Decision Support System</td>
</tr>
<tr>
<td>dSS</td>
<td>Incremental change in the Reorder point R</td>
</tr>
<tr>
<td>DWIP</td>
<td>Desired Work In Progress</td>
</tr>
<tr>
<td>E</td>
<td>Error</td>
</tr>
<tr>
<td>EINV</td>
<td>Error in Inventory Holding</td>
</tr>
<tr>
<td>EWIP</td>
<td>Error in Work In Progress</td>
</tr>
</tbody>
</table>
FINV Factory Inventory  
G Gain (Distributors Safety Stock/Average Consumption)  
GIT Goods In Transit  
ITAE Integral of Time*Absolute Error.  
ITAE Integral of Time * Absolute Error  
O Order-up-to-point  
OR Operations Research  
ORATE Order Rate  
R Re-order point  
T Transport Quantity  
Ta Consumption Averaging Time Constant  
Tbar or \( \bar{T} \) Estimate of the production lead-time  
Ti Inverse of Inventory Based Production Control Law Gain  
TINV Target System Inventory Holding  
Tp The production lead-time in units of sampling intervals  
Tq Exponential smoothing constant used at the distributor to set R  
Tw Inverse of WIP Based Production Control Law Gain  
VAR Variance  
VCON Virtual Consumption  
VMI Vendor Managed Inventory  
VMI-APIOBPCS Vendor Managed Inventory, Automatic Pipeline, Inventory and Order Based Production Control System  
W Ratio of production adaptation to inventory costs  
w Frequency  
WIP Work In Progress  
w Noise Bandwidth

### Appendix A. The difference equations required for VMI-APIOBPCS

<table>
<thead>
<tr>
<th>Description</th>
<th>Difference Equations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumption Forecast</td>
<td>( AVCON_{t} = AVCON_{t-1} + \frac{1}{1+Ta} (VCON_{t} - AVCON_{t-1}) ),</td>
</tr>
<tr>
<td>Target (desired) WIP</td>
<td>( DWIP_{t} = AVCON_{t} * Tp ),</td>
</tr>
<tr>
<td>Actual WIP</td>
<td>( WIP_{t} = WIP_{t-1} + ORATE_{t} - COMRATE_{t} ),</td>
</tr>
<tr>
<td>Error in WIP</td>
<td>( EWIP_{t} = DWIP_{t} - WIP_{t} ),</td>
</tr>
<tr>
<td>Inventory error</td>
<td>( EINV_{t} = DINV_{t} - AINV_{t} ),</td>
</tr>
<tr>
<td>Order Rate</td>
<td>( ORATE_{t} = AVCON_{t-1} + \frac{EINV_{t-1}}{Ti} + \frac{EWIP_{t-1}}{Tw} ),</td>
</tr>
<tr>
<td>Completion Rate</td>
<td>( COMRATE_{t} = ORATE_{t-1} \cdot Tq ),</td>
</tr>
<tr>
<td>Actual Inventory Level</td>
<td>( AINV_{t} = AINV_{t-1} + COMRATE_{t} - CONS_{t} ),</td>
</tr>
<tr>
<td>Virtual Consumption</td>
<td>( VCON_{t} = CONS_{t} + \left[ VCON_{t-1} + \frac{1}{1+Tq} ((G<em>CONS_{t}) - VCON_{t-1}) \right] - \left[ VCON_{t-2} + \frac{1}{1+Tq} ((G</em>CONS_{t-1}) - VCON_{t-2}) \right] )</td>
</tr>
<tr>
<td>Typical Test Input</td>
<td>( CONS_{t} = \begin{cases} 0 &amp; \text{if } t \leq 0 \ 1 &amp; \text{if } t &gt; 0 \end{cases} ) for a step input.</td>
</tr>
<tr>
<td>Typical Target Inventory</td>
<td>( DINV_{t} = 0 ),</td>
</tr>
</tbody>
</table>
Alternatively when modeling disaggregate inventory levels and transportation dispatches in the VMI-APIOBPCS model the following difference equations can be used:

<table>
<thead>
<tr>
<th>Description</th>
<th>Difference Equations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecasted Re-order point at the distributor</td>
<td>( R_t = R_{t-1} + \frac{1}{1+T_q}((G*\text{CONS}<em>t) - R</em>{t-1}) )</td>
</tr>
<tr>
<td>Order-up-to point at the distributor</td>
<td>( O_t = R_t + TQ_t )</td>
</tr>
<tr>
<td>Distributor's inventory level</td>
<td>( DINV_t = DINV_{t-1} - \text{CONS}<em>t + \text{DES}</em>{t-T} )</td>
</tr>
<tr>
<td>Goods In Transit between factory and distributor</td>
<td>( \text{GIT}<em>t = \sum</em>{t=-T}^{T} \text{DES}_t ), where ( T ) is the transportation lead-time.</td>
</tr>
</tbody>
</table>
| Dispatches                                       | \( \text{DES}_t = \begin{cases} TQ_{t-t} \text{ if } DINV_{t-t} + \text{GIT}_{t-t} < R_{t-t} \\
| Transport quantity                               | \( \text{TQ}_{t} = \text{CONS}_t \text{ or ETQ}_t \)                              |
| System inventory levels                          | \( \text{SINV}_t = \text{FINV}_t + \text{GIT}_t + \text{DINV}_t \)                |
| Factory inventory levels                         | \( \text{FINV}_t = \text{FINV}_{t-1} + \text{COMRATE}_t + \text{DES}_t \)          |
| Virtual consumption                              | \( \text{VCON}_t = \text{CONS}_t + \text{dSS}_t \)                               |
| Net changes in the distributor's safety stock level | \( \text{dSS}_t = R_t - R_{t-t} \),                                              |
| Desired consumption for the factory              | \( \text{AVCON}_t = \text{AVCON}_{t-1} + \frac{1}{1+Ta} (\text{VCON}_t - \text{AVCON}_{t-1}) \), |
| Desired WIP                                      | \( \text{DWIP}_t = \text{AVCON}_t * \text{Tp} \)                                 |
| Actual WIP                                       | \( \text{WIP}_t = \text{WIP}_{t-1} + \text{ORATE}_t - \text{COMRATE}_t \)          |
| Error in WIP                                     | \( \text{EWIP}_t = \text{DWIP}_t - \text{WIP}_t \)                                |
| Order rate                                       | \( \text{ORATE}_t = \text{AVCON}_{t-1} + \frac{\text{FINV}_{t-1}}{\text{Ti}} + \frac{\text{EWIP}_{t-1}}{\text{Tt}} \) |
| Completion rate                                  | \( \text{COMRATE}_t = \text{ORATE}_{t(Tp)} \)                                     |
| Error in system inventory levels                 | \( \text{EINV}_t = \text{TINV}_t - \text{SINV}_t \)                               |
| Typical Test Input                               | \( \text{CONS}_t = \begin{cases} 0 \text{ if } t < 0 \\
| Typical Target inventory                         | \( \text{TINV}_t = 0 \)                                                            |

### Appendix B. Description of the optimization routine

The Mathematica source code required for the optimization routine is as follows;

```mathematica
apiobpcs[Ta_, Ti_, Tq_, Tw_, Tbarp_, G_] = 
E^\{Tw\} (-G - Tq + E^\{Tw\}(-1 + E^\{Tq\}))((-1 + E^\{Tw\})Tbarp Ti - (Ta + Ti)Tw + E^\{Tw\}(1 + Ta + Ti)Tw) 
(-Ta + E^\{Tw\}(1 + Ta))(-Tq + E^\{Tq\}(1 + Tq))(Tw + Ti(-1 - E^\{Tw\}(-1 + Tw) + E^\{Tw\}Tw));

score[Ta_, Ti_, Tq_, Tw_, Tbarp_, G_] = ((Sum[(Abs[apiobpcs[Ta, Ti, Tq, Tw, Tbarp, G]])**2 
, {w, 0.0001, Pi, 0.0001}])**0.0001);

seafind=Compile[{{maxcon}, {G}, {W}}, Module[{Ta=0, Ti=1, Tw=1, Tq=0, Tbarp=4, prevmax=100000.0, Taopt=0, Ti=opt=1, Twopt=0, Tqopt=0, cons=0.0, vecond=0.0, itaed=0.0, nd=0
, scored=0.0, cons=0.0, avcon=0.0, dwip=0.0, wip=0.0, ewip=0.0, einv=0.0, ainv=0.0, or
```
ate=0.0, comrate=0.0, cr1=0.0, cr2=0.0, cr3=0.0, cr4=0.0, dss=0.0, pvecon=0.0, vecon=0.0, itae=0.0, n=0, loopcount=0},

While[Ta<(maxcon*3),
    While[Ti<(maxcon*2),
        While[Tw<maxcon*5, 
            While[Tq<(maxcon/2),
                While[n<60, cons=If[n>2,1.,0.]; vecon=vecon+((1/(1+Tq))*((G*cons)-vecon)); dss=vecon-pvecon; vcon=dss+cons; avcon=avcon+((1/(1+Ta))*vcon-vecon)); comrate=cr4; cr4=cr3; cr3=cr2; cr2=cr1; cr1=orate; ainv=ainv-cons+comrate; einv=((0-ainv)/Ti); dwip=avcon*4; wip=wip+orate-comrate; ewip=((dwip-wip)/Tw); orate=avcon+einv+ewip; itae=itaes+(Abs[ainv]*(n-2)); If[orate>50, n=60]; n++];
        ]];
    ]];
While[Tq<7, Tw++, Tw=Tw+7];

References


