On the bullwhip and inventory variance produced by an ordering policy

S.M. Disney*, D.R. Towill

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

Received 26 November 2001

Abstract

The Bullwhip Problem in supply chains is first outlined. A discrete control theory model of a generic model of a replenishment rule is presented. From this model, an analytical expression for bullwhip is derived that is directly equivalent to the common statistical measure often used in simulation, statistical and empirical studies to quantify the bullwhip effect. This analytical expression clearly shows that bullwhip can be reduced by taking a fraction of the error in the inventory position and pipeline position, rather than account for all of the errors every time an ordering decision is made as is common in many scheduling systems. Furthermore, increasing the average age of the forecast reduces bullwhip, as does reducing the production lead-time. Next an analytical expression for the variance of the inventory position is derived and used together with the bullwhip expression to determine suitable ordering system designs that minimise both bullwhip and inventory variance for a range of weightings between the two variances. The relationship between our Bullwhip metric and the control theory metric Noise Bandwidth is highlighted. This contribution then derives and exploits an analytical expression for the Integral of Time * Absolute Error (ITAE) criterion often used to quantify inventory responsiveness in response to a deterministic demand. The results form a Decision Support System that can be used to design balanced supply chain ordering decisions and clearly shows how the Bullwhip Effect can be reduced, at the expense of increased inventory recovery times. Thus, this paper presents the general solution to the bullwhip problem on a sound mathematical basis.

Keywords: Bullwhip effect; Supply chains; z-transforms; ITAE; APIOBPCS

1. Introduction

The “Bullwhip Effect” is a new term (but not a new phenomenon) coined by Lee et al. [1,2]. It refers to the scenario where the orders to the supplier tend to have larger fluctuations than sales to the buyer and the distortion propagates upstream in an amplified form. Lee et al. [1,2] state that there are five fundamental causes of bullwhip as shown in Fig. 1; non-zero lead-times, demand signalling processing, price variations, rationing and gaming, and order batching, to which other proven sources may be added. However, essentially the bullwhip effect is not a new supply chain phenomenon; Together the impact of Demand signal processing and non-zero lead-times has in the past been called the “Demand Amplification” or the “Forrester Effect” after Jay Forrester [3] who encountered the problem in many real-world supply chains and demonstrated it via DYNAMO simulation. The Forrester Effect is also encompassed by Sterman’s bounded rationality [4], terminology that is common in the field of psychology and used to describe players sub-optimal but seemingly rational decision-making behaviour. This contribution is concerned solely with the cause of the bullwhip effect due to demand amplification or the Forrester Effect,
The Bullwhip Effect

Price fluctuations or the Promotion Effect

Rationing and gaming or the Houlihan Effect

Order batching or the Burbidge Effect

Forrester Effect caused by lead-times and demand signal processing

Fig. 1. The four causes of the bullwhip effect.

although for completeness the other causes [5] are:

- **Order batching** is also known as the Burbidge Effect [6]. It refers to the practise of placing orders down the supply chain (or on the manufacturing process) in batches in order to gain economies of scale in set-up activities (such as setting up a machine or placing and receiving an order). It is often the result of an Economic Order Quantity calculation or similar technique. Burbidge [6] discusses the problems this causes in detail. Towill [7] outlines the contributions of Forrester and Burbidge for avoiding the Bullwhip Effect in an integrated approach termed “Forridge”.

- **Rationing and gaming**, or the Houlihan Effect was highlighted by Houlihan [8] who recognised that as shortages or missed deliveries occur in traditional supply chains, customers over-load their schedules or orders. This in turn places more demands on the production system that inevitably leads to more unreliable deliveries. Customers then increase their safety stock target that further distorts the demand signal via the Forrester Effect, giving rise to the Bullwhip Problem.

- **Price variations** or the Promotion Effect refers to the practise of providing products at reduced prices to stimulate demand. Lee et al. [1,2] discuss an industrial example of this issue.

Recently, many authors have been using Eq. (1) as a simple measure of the Bullwhip Effect [9], where \( \text{ORATE} \) refers to the orders placed on a supplier and \( \text{CONS} \) represents sales or consumption. Although this is a useful measure, it only quantifies output variance compared with the input variance and does not describe the structure that causes the variation increase. In Eq. (1), \( \sigma^2 \) denotes variance and \( \mu \) denotes the mean. For a stationary random signal, over long periods of time the means (\( \mu \)) cancel out as \( \mu_{\text{ORATE}} = \mu_{\text{CONS}} \) if the system is linear.

\[
\text{Bullwhip} = VR_{\text{ORATE}} = \frac{\sigma^2_{\text{ORATE}}}{\mu_{\text{ORATE}}^2} = \frac{\sigma^2_{\text{CONS}}}{\mu_{\text{CONS}}^2}.
\]  

(1)

This variance ratio (VR) measure can easily be applied to quantify fluctuations in actual inventory (AINV) as shown in Eq. (2). Note, that for the inventory variance amplification measure we are solely concerned with the ratio of the variance as mean inventory levels are arbitrary within our framework herein.

\[
\text{Inventory variance amplification} = VR_{\text{AINV}} = \frac{\sigma^2_{\text{AINV}}}{\mu_{\text{CONS}}^2}.
\]  

(2)

2. The inventory and production adaptation cost trade-off

Baganha and Cohen [10] postulate a puzzling idea: “inventory management policies can have a destabilising effect by increasing the volatility of demand as it passes up the chain”, whereas “one of the principal reasons used to justify investment in inventories is its role as a buffer to absorb demand variability”. In other words, inventories should have a stabilising effect on material flow patterns. So how is it that market variability is amplified rather than dampened? The answer to this conundrum is that inventories can have a stabilizing effect on demand variation provided the replenishment decision is designed carefully via common control theory techniques. However, they often do not have a stabilizing effect because they are poorly designed [11]. Indeed this has been shown by a number of authors using control theory including Simon [12], Vassian [13], Deziel and Eilon [14], Towill [15], Berry et al. [16], Disney et al. [17] and Towill and McCullen [18].

The real question to be answered in this situation is “To what extent can production rates be smoothed in order to minimise production adaptation costs without adversely increasing our inventory costs?” This is an important trade-off because if a perfectly level production rate (i.e. level scheduling) is used then large inventory deviations are found and hence large inventory costs are incurred. Conversely, if inventory deviations are minimised (by “passing on orders”), highly variable production schedules are generated and hence production adaptation costs are incurred. Of course, the exact position of the optimal trade-off depends on the properties of the particular supply chain it is operating within. This important production and inventory trade-off is illustrated conceptually in Fig. 2 adapted from Towill [19] by Towill et al. [20]. This figure, in one simple sketch, has covered the whole spectrum of possibilities from which a supply chain designer can select an appropriate ordering decision. For example, a Fixed Production Rate
corresponds to Lean Production [21,22] strategy where a level schedule is preferred; Pass on Orders corresponds to an Agile Production strategy where the aim is to minimise inventory deviations [18]; and Leagile Production [23] is the appropriate mix of Lean and Agile modes at different echelons in the supply chain [20]. There are also “shades of grey” between these extremes, which may be described as the “explicit or ideal filter” approach where low frequency variations are followed and higher frequency variations filtered out or attenuated. The approach taken by this paper may be considered to cover all of these strategies, i.e. Lean, Agile, Leagile and Explicit Filter, as all of these strategies have been implicitly investigated. The Order-Up-To policy and many variants of it are also encapsulated by this approach. See Dejonckheere et al. [11], for a detailed investigation of the Order-Up-To strategy.

2.1. Overview of the production and inventory control system

The particular production or distribution ordering procedure exploited in this study is a special case of the Automatic Pipeline Inventory and Order Based Production Control System (APIOPBCS) system coined by John et al. [24]. The APIOPBCS system can be expressed in words as “Let the production (or distribution) targets be equal to the sum of: averaged demand (exponentially smoothed over Ta time units), a fraction (1/Ti) of the inventory difference in actual stock compared to target stock and the same fraction (1/Tw) of the difference between target Work In Progress (WIP) and actual WIP”. Naim and Towill [25] then found that the APIOPBCS structure was directly equivalent to Sterman’s [4] Anchoring and Adjustment algorithm that he showed fitted human behaviour when playing the “Beer Game”. Mason-Jones et al. [26] have further exploited this fact using simulation. Disney and Towill [27] have investigated the stability of traditional (APIOPBCS) and Vendor Managed Inventory supply chains in discrete time via the z-transform and the Tustin transformation.

The APIOPBCS model uses three components to generate orders in the supply chain. The first type of information is a forecast. We have chosen to generate this via exponential smoothing, because it is readily understood by practitioners, is the most common technique found in the literature and industry [9], and is relatively simple mathematically. The second component of the order rate is a fraction (1/Ti) the discrepancy between TINV and AINV. The fraction is used because it is easily understood and known to be quite capable of “locking on” to target inventory (TINV) levels if the production lead-time is known. The third component of the order rate is a fraction (1/Tw) of the discrepancy between target and actual WIP (or error between the TINV on order but not yet received and the AINV on order but not yet received in the language of the Beer Game). The fraction is used because it is easily understood and known to be quite capable of “locking on” to target WIP levels if the production lead-time is known.

The APIOPBCS model is particularly powerful because it can represent, by setting particular controller values to specific values, the wide range of supply chain strategies outlined in the previous section such as Lean and Agile supply chains. The Order-Up-To model and its variants are also an important subset of the APIOPBCS model that has been studied in detail by Dejonckheere et al. [11]. The extension of APIOPBCS to Vendor Managed Inventory scenarios is described and investigated in Disney [28]. Dejonckheere et al. [29] further extend APIOPBCS to supply chains where
end consumer sales (EPOS) data is passed to and exploited by all members of the supply chain, in a strategy termed Information Enrichment. Mason-Jones and Towill [30] have also studied variations of this strategy via simulation.

2.2. The importance of the Deziel and Eilon variant

The particularly important special case of APIOBPCS, which we term DE-APIOBPCS, was studied by Deziel and Eilon [14], hence the DE addition to the acronym. However, the Deziel and Eilon model considers a different order of events to the APIOBPCS model as shown by Disney [28] using the Final Value Theorem. Deziel and Eilon’s contribution seems to have been largely missed by the academic community, having not been mentioned in three major reviews of control theory applications to the production and inventory control problem, Axsäter [31], Edghill and Towill [32] and Riddals et al. [33]. APIOBPCS strictly follows the order of events used in the Beer Game, Sterman [4], and hence the lead-time (Tp) is increased by one period and there is a one period delay in the WIP feedback loop as noticed by Vassian [13]. The DE-APIOBPCS model refers to the case when Tw = Ti making it a subset of APIOBPCS, and is described by the block diagram shown in Fig. 3.

The DE-APIOBPCS model has some very special properties, outlined in Disney [28], investigated further by Disney and Towill [27,34] and exploited in an industrial setting by Disney et al. [35]. In particular, the system is guaranteed to be stable, Disney and Towill [27], and is robust to a surprising number of non-linear effects, Disney and Towill [34]. Additionally, the production lead-time simplifies out of the ORATE transfer function, reducing the complexity of the mathematical analysis. Note also that the mean AINV level is equal to TINV and has no effect on our results herein. Hence, we may arbitrarily set TINV equal to the mean Consumption level for simplicity later.

Deziel and Eilon’s (unfortunately largely ignored) contribution, which Professor Will Bertrand (Eindhoven University) drew to the author’s attention via Jeroen Dejonckheere and Professor Marc Lambrecht (Leuven University, Belgium) through our Order-Up-To research, is a significant contribution that uses z-transforms. In particular, Deziel and Eilon mention in a “throw away” sentence in Appendix 2 of [14] the method of linking the field of control theory to the bullwhip problem. This contribution is essentially the exploitation of Deziel and Eilon’s statement without proof that the sum of square of the system’s impulse response is equal to the variance of the output divided by the variance of the input when the input is an independent and identically distributed random variable.

Further investigation revealed that Deziel and Eilon based this claim on the work of the Russian mathematician and engineer, Yakov Tyspkin [36], whose work is also extensively exploited in this paper. Tyspkin’s work, which was originally written in Russian [37], was also noticed by Jury [38], who provides an illuminating presentation of these mathematical tools, specifically using the z-transform. The unique contribution of this paper is to offer a new, conceptually very clear, mathematical proof behind Deziel and Eilon’s statement. We also explicitly recreate Tyspkin’s proof, which can be found scattered throughout his work [36], and relate it to the bullwhip problem, highlighting the link between bullwhip, the impulse response, noise bandwidth, the frequency response, the system’s response to random input signals and the auto-correlation function. Additionally, we use some of Tyspkin’s techniques to derive an analytic expression of the Integral of Time-Weighted Absolute Error (ITAE), which has been often applied to the inventory signal to measure inventory responsiveness.

2.3. The mathematical model of DE-APIOBPCS

Manipulating the block diagram (Fig. 3) where \( z = 1/(1 + Ta) \), to find the production order rate, ORATE (Eq. (3)), and the actual inventory level, AINV (Eq. (4)) transfer function, yields

\[
\begin{align*}
\text{ORATE} &= \frac{z^{1-Tp}(-Ta - Tp - Ti + z + Ta + Tpz + Tz - (1 + Tz(-1 + z)))}{z^{Tp}((Ta(-1 + z) + z))}, \\
\text{AINV} &= \frac{(1 + Ti(-1 + z))}{(-1 + z)(Ta(-1 + z) + z)}. 
\end{align*}
\]

Taking the inverse z-transform of Eq. (3) gives the ORATE difference equation or lattice function

\[
|n| = \frac{(1 + Ta)^{(-1/Ta)}y[(1 + Ta + Tp) - (\frac{Tp}{Ta})y(Ti + Tp)]}{(1 + Ta)(1 + Ta - Ti)Ti}.
\]

Unfortunately the AINV transfer function is a transcendental equation and therefore its inverse cannot be found without specifying the lead-time, Tp. Even when Tp is defined the difference equation is a rather large and complicated and thus is not shown here.

3. Analytical solution to the bullwhip problem

From Eq. (1) we have

\[
\text{Bullwhip} = \frac{\sigma^2_{\text{OUT}}/\mu_{\text{OUT}}}{\sigma^2_{\text{IN}}/\mu_{\text{IN}}}
= \frac{\frac{1}{N_{\text{OUT}}} \sum_{n=0}^{N_{\text{OUT}}} (f_{\text{OUT}}[n] - \mu_{\text{OUT}})^2}{\frac{1}{N_{\text{IN}}} \sum_{n=0}^{N_{\text{IN}}} (f_{\text{IN}}[n] - \mu_{\text{IN}})^2}.
\]
As \( \mu_{\text{OUT}} = \mu_{\text{IN}} \) for a stationary random demand signal, this reduces to

\[
\text{Bullwhip} = \text{VR}_{\text{ORATE}} = \frac{\sum_{n=0}^{N} (f_{\text{OUT}}[n] - \mu_{\text{OUT}})^2}{\sum_{n=0}^{N} (f_{\text{IN}}[n] - \mu_{\text{IN}})^2}.
\]  

(7)

Furthermore, because we are assuming the system is linear, we can arbitrarily set the remaining \( \mu_{\text{OUT}} = \mu_{\text{IN}} = 0 \) to yield

\[
\text{Bullwhip} = \text{VR}_{\text{ORATE}} = \frac{\sum_{n=0}^{N} f_{\text{OUT}}^2[n]}{\sum_{n=0}^{N} f_{\text{IN}}^2[n]}.
\]  

(8)

Restating this in the inverse \( z \)-transform notation we have

\[
\text{Bullwhip} = \text{VR}_{\text{ORATE}} = \frac{\sum_{n=0}^{N} Z^{-1}\{F^* (z) I(z)\}^2}{\sum_{n=0}^{N} Z^{-1}\{I(z)\}^2},
\]  

(9)

which simplifies (from the definition of the \( z \)-transform) to

\[
\text{Bullwhip} = \text{VR}_{\text{ORATE}} = \frac{\sum_{n=0}^{N} Z^{-1}\{F^* (z)\}^2}{\sum_{n=0}^{N} \delta^2[n]},
\]  

(10)

where \( \delta \) is the Kronecker-Delta function, those sum is obviously 1, thus,

\[
\text{Bullwhip} = \text{VR}_{\text{ORATE}} = \sum_{n=0}^{N} Z^{-1}\{F^* (z)\}^2 = \sum_{n=0}^{N} f^2[n].
\]  

(11)

As \( N \) approaches infinity the bullwhip measures refers to the population rather then the sample, thus,

\[
\text{Bullwhip} = \text{VR}_{\text{ORATE}} = \sum_{n=0}^{\infty} f^2[n].
\]  

(12)

Thus the sum of the square of the discreta of the system impulse response is equal to the bullwhip measure. Applying Eq. (12) to the ORATE lattice function (Eq. (5)) yields Eq. (13), an analytic expression for the bullwhip effect:

\[
\text{Bullwhip} = \text{VR}_{\text{ORATE}} = \frac{2T_a^2 + 3T_i + 2T_p + 2(T_i + T_p)^2 + T_a(1 + 6T_i + 4T_p)}{(1 + 2T_a)(T_a + T_i)(-1 + 2T_i)}.
\]  

(13)
that shows that Bullwhip does reduce when Ta and Ti increase.

Inspection of Fig. 4 and Eq. (13) shows that Bullwhip is symmetrical about Ti = Ta + 1, where minimum variance occurs when Ta and Ti is large.

4. Predicting inventory variance

Applying Eq. (8) to the AINV transfer function (Eq. (4)) yields Eq. (14), an expression for inventory variance.

\[
VR_{\text{AINV}} = 1 + Tp + \frac{2Ta^2(-1 + Ti)^2 + Ti(1 + Tp)^2 + Ta(1 + Tp)(1 + (-1 + Ti)Tp)}{(1 + 2Ta)(Ta + Ti)(-1 + 2Ti)}. \tag{14}
\]

Inspection of Eq. (11) shows that:

- the inventory variance is always greater than 1, i.e. inventory levels will always vary more then the variation in the demand signal. In other words, it is not possible to achieve zero inventory policy without making the customer wait for the product with this ordering decision.
- The first unit of inventory variance is due to the order of events delay, i.e. because demand for this planning period is not known until the end of the planning period, inventory has got to vary at least as much as the demand levels.
- Inventory levels will also vary by at least \( \sigma_{\text{CONS}}^2(1 + Tp) \), no matter how well the parameters Ta and Ti are set to minimise the squared impulse response deviation.
- Inventory variance increases as the production lead-time increases.

The inventory variance function will now be enumerated as it was for the bullwhip effect. The effect of Ta and Ti on the inventory variance is shown in Fig. 5 for the case when the production lead-time, \( Tp \), is 3 time periods long. Inspection of Fig. 5 and Eq. (14) shows that the inventory variance is symmetrical about Ti = Ta + 1, where minimum variance occurs when Ta is large and Ti is small and vice versa.

It should be remembered that the \( VR_{\text{AINV}} \) and \( VR_{\text{ORATE}} \) can always be evaluated by directly modelling the squared time domain impulse response in a spreadsheet to very quickly understand how a system will respond to random excitations without using transform techniques. This is an important result for analysts as the alternative routes (either by simulating the response to a random signal or, as highlighted later the, estimation of the squared frequency response) are computationally much more expensive than a simple impulse response. This is especially the case for reasonably low values of Ta and Ti (i.e. Ta and Ti < 30), as the sum typically converges within 300 iterations.

5. Balancing inventory variations and the bullwhip effect

The obvious question to ask now is “what combinations of the design parameters (Ta and Ti) minimise the sum of inventory variance and bullwhip?” Of course the financial implications of variance in the production order rate will be different for the variation in the inventory levels. Therefore, different weightings \( (K) \) for the Bullwhip and inventory variance sum used to evaluate a performance measure (termed \( \text{SCORE} \) in Eq. (15), will be considered via the Ta–Ti parameter plane, i.e.

\[
\text{SCORE} = (K * VR_{\text{ORATE}}) + VR_{\text{AINV}}. \tag{15}
\]

It will be assumed that \( VR_{\text{ORATE}} \) is always more costly then \( VR_{\text{AINV}} \), hence only a weight, \( K > 1 \), applied to \( VR_{\text{ORATE}} \) needs to be considered.

From Fig. 6 it can be seen that the good designs, in terms of the bullwhip effect and inventory variance, for DE-APIOPBCS occur when 4 < Ta < 10 and 4 < Ti < 10 for reasonable values of \( K \) when \( Tp = 3 \).

6. The relationship between bullwhip and noise bandwidth

Noise Bandwidth has often been used in the literature, for example Disney et al. [17] and Dejonckheere et al. [11], as a means of quantifying bullwhip. To highlight how the Noise Bandwidth relates to the variance amplification ratios, the Lyapunov–Parseval relationship as described by
Tsypkin [36] will be exploited. The Lyapunov–Parseval relationship holds for a stable system (i.e. a system that has a negative abscissa of convergence). Here, we have an advantage as DE-APIOBPCS is always stable and has a simplified structure. Tsypkin’s work was concerned with the discrete laplace transform (DLT). It is directly equivalent to the z-transform, with a simple change of notation \((z = e^{\phi})\). However, Tsypkin preferred to use the DLT due to the ease of drawing analogies between differential equations and difference equations (or lattice functions in Tsypkin’s
terminology), and because theorems often possess a simpler form using this notation.

Departing from the definition of the basic DLT, the direct DLT of the difference equation, \( f[n] \), denoted \( F^*(q) \) is defined by Eq. (16) where

\[
D\{f[n]\} = F^*(q) = \sum_{n=0}^{\infty} e^{-qn} f[n]. \tag{16}
\]

Tsypkin shows that the sum of the ordinates or discrete of a lattice function is equal to the value of the transfer function at \( q = 0 \), i.e.

\[
F^*(0) = \lim_{q \to 0} F^*(q) = \sum_{n=0}^{\infty} f[n] \tag{17}
\]

and that the transform of the product of two lattice functions may be calculated from the transform of the two lattice functions in the manner shown by Eq. (18) if suitable convergence conditions exist:

\[
D\{f[n]f[n]\} = \frac{1}{2\pi j} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F^1(s)sF^2(q-s) \, ds. \tag{18}
\]

Thus by applying Eqs. (17)–(18), the sum of the square of the ordinates of a lattice function \( f[n] \) is given by

\[
\sum_{n=0}^{\infty} f^2[n] = \frac{1}{2\pi j} \int_{-\infty}^{\infty} F^*(s)F^*(-s) \, ds. \tag{19}
\]

Putting \( s = j\tilde{\nu} \) and \( c = 0 \) into Eq. (19) yields

\[
\sum_{n=0}^{\infty} f^2[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} F^*(j\tilde{\nu})F^*(-j\tilde{\nu}) \, d\tilde{\nu}. \tag{20}
\]

or, since

\[
F^*(j\tilde{\nu})F^*(-j\tilde{\nu}) = |F^*(j\tilde{\nu})|^2, \tag{21}
\]

we have, because the latter expression is even,

\[
\sum_{n=0}^{\infty} f^2[n] = \frac{1}{\pi} \int_{0}^{\pi} |F^*(j\tilde{\nu})|^2 \, d\tilde{\nu}. \tag{22}
\]

Or, in words the sum of the square of the ordinates of a lattice function is proportional to the area under its spectrum proportional to what modern day control engineers call the Noise Bandwidth, \( W_N \).

\[
\sum_{n=0}^{\infty} f^2[n] = \frac{1}{\pi} \int_{0}^{\pi} |F(j\tilde{\nu})|^2 \, d\tilde{\nu} = \frac{W_N}{\pi}. \tag{23}
\]

Therefore, the \( W_N/\pi \) is also equal to the bullwhip measure. This result has been exploited in previous literature, as in Disney et al. [17] and exploited and verified via simulation by Dejonckheere et al. [11]. The latter reference also includes a proof involving the use of Fourier Transforms. Remembering that the DLT is equivalent to the \( z \)-transform by \( z = e^{\tilde{\nu}} \), meaning this result will be true for the \( z \)-transform as they have the same definition, see Eq. (24).

\[
D\{f[n]\} = F^*(q) = \sum_{n=0}^{\infty} e^{-qn} f[n] = \sum_{n=0}^{\infty} z^{-n} f[n] = F(z). \tag{24}
\]

There are some more insights to be gained from Tsypkin’s [36] work that are not shown here but are shown in Appendix A.

### 7. An alternative inventory response measure

An alternative inventory response metric that has been used in recent literature (see [17,40]) is the Integral of Time “Absolute Error” (ITAE) and is defined as follows, Eq. (25);

\[
\text{ITAE}_{\text{AINV}} = \sum_{n=0}^{\infty} n|E|, \tag{25}
\]

where \( n \) is the time period and \( E \) is the error in the inventory levels measure as the deviation of AINV level from the TINV level. Historically, the ITAE has been calculated via simulation. However, Tsypkin’s approach offers a new analytical measure. First, note that the ORATE time domain response only contains exponential terms as we are assuming a pure time delay is used for the production delay and thus so must the AINV response, we conclude that the inventory response can only be negative. Departing from the DE-APROPBCSAINV transfer function (Eq. (4)) multiplied by the Heaviside Step function \((h)\), the transfer function of the absolute error in inventory levels is therefore given by Eq. (19) for the case when \( Tp = T\beta = 3 \).

\[
\left| \frac{\text{AINV}}{\text{CONS}} \right| = -\frac{3 + Ta + Ti + (2 + Ta + Ti)z + (1 + Ta + Ti)z^2 + (1 + Ta)Tiz^3}{(1 + Ti(-1 + z))z^2(Ta(-1 + z) + z)}. \tag{26}
\]

Taking the inverse \( z \)-transform of the Eq. (26), multiplying by the time index, \( n \), and taking the \( z \)-transform of the result yields Eq. (27). This represents the \( z \)-transform of the
absolute error multiplied by time.

\[ TAE^{*}(z) = Z \left\{ \frac{nZ^{-1} \left\{ AINV \right\}}{CONS} \right\} \]

\[
(2Ta(-1 + Ti)(3 + Ta + Ti) + 9z - (-2Ta(5 + Ta) + (3 + Ta)(2 + 5Ta)Ti + (3 + 5Ta)Ti^2)z
+2(2 + Ta^2 + Ti(5 + Ti) + Ta(3 + 4Ti))z^2 - (-1 + Ta^2) - 2(1 + Ta)(3 + Ta)Ti + (-2 + Ta^2)Ti^2)z^3
\]

\[
= \frac{2(1 + Ta)Ti(1 + Ta + Ti)z^4 + (1 + Ta)^2Ti^2z^5}{((1 + Ti(-1 + z))^2z^2(Ta(-1 + z) + z)^2)}. \quad (27)
\]

Exploiting Eq. (17), we replace \( z = e^q \) in Eq. (27) and find the limit as \( q \) approaches zero to yield Eq. (28). *This algebraic expression gives the Integral of Time*\(^*\) *Absolute Error analytically.* Interestingly, if \( Ta \) and \( Ti \) are constrained to integer values, the ITAE is also always an integer. This is a useful cross-check for accuracy to keep in mind when calculating ITAE via simulation. A further note of importance is relevant here. We could have also found the limit of Eq. (27) as \( z \) approaches 1, by exploiting the work of Grubbström [41,42], and colleagues at the Department of Production Economics in Linköping’s Institute of Technology concerned with the Net Present Value transform (for example, see Wikner [43] for an application of the NPV transform to IOPBCS). Other interesting research has emanated from Linköping concerning the use of transform methods as probabilistic moment generating functions to study stochastic issues in supply chains, for example, see Grubbström [44] or Tang [45].

\[ ITAE_{AINV} = Ta^2(3 + Ti) + Ta(12 + 7Ti + Ti^2) \]

\[ + 2(7 + 6Ti + 2Ti^2). \quad (28) \]

It can be clearly seen from Eq. (28) that quicker inventory responses result from smaller values of \( Ta \) and \( Ti \). Fig. 7 verifies this. Again it can be appreciated that using Eq. (28) rather than simulation to calculate the ITAE is considerably more accurate (side stepping accumulative rounding errors) and cheaper in terms of computer time.

However, this is in direct conflict with the Bullwhip measure and it would be interesting to investigate the \( Ta-Ti \) parameter plane by overlaying Fig. 7 with Fig. 4 as shown in Fig. 8. It unequivocally demonstrates the trade-off between inventory responsiveness and bullwhip.

It would be helpful to keep the \( Tp \) unspecified to determine the ITAE as a function of the production lead-time. However, this requires the solution to a transcendental equation and is difficult to achieve. Essentially, the problem is due to having to increase the time index of the time*\(^*\) absolute error lattice function, \( tae(n) \) in order to reflect \( Tp \) before taking the final \( z \)-transform as the order of \( TAE^{*}(z) \) is dependent in \( Tp \).

8. Conclusions

This paper has presented an analytic solution to the bullwhip problem for a specific ordering policy. It has shown
conclusively that the way to minimise the bullwhip problem with our policy is to increase the average age of forecasts and reduce the rate at which inventory and WIP correction are accounted for in the production/distribution-ordering algorithm. This can be achieved without unduly affecting the inventory variance. The value of Time Compression Paradigm has been highlighted as an effective means of reducing the bullwhip problem in our ordering policy. An analytic expression for inventory variance has also been derived and investigated. Analysts may easily exploit this procedure for determining bullwhip and inventory variance without resorting to control theory since it can now be carried out via simulation in a spreadsheet environment. The relationship between bullwhip and the Noise Bandwidth metric has been investigated. The ITAE performance index has been recast against inventory responsiveness. It means that bullwhip avoidance can be satisfactorily balanced against inventory responsiveness.

Appendix A

If we now turn our attention to the bullwhip measure, and its relation to the variance of output signal, the auto-correlation and spectral density function, some more insights can be gained from Tsypkin’s work [36]. From the definition of the bullwhip measure in Eq. (1) we have

$$VR = \frac{\sigma_{\text{OUT}}^2}{\mu_{\text{OUT}}} = \frac{\frac{1}{N} \sum_{n=0}^{N} (f_{\text{OUT}}(n) - \mu_{\text{OUT}})^2}{\frac{1}{N} \sum_{n=0}^{N} (f_{\text{IN}}(n) - \mu_{\text{IN}})^2}.$$  \hspace{1cm} (A.1)

As the mean values cancel out in the bullwhip measure and we are assuming the system is linear, we have the luxury of being able to set them arbitrarily. Note also that the mean inventory levels are also arbitrary. Letting the means equal zero and then giving the input signal (demand or CON5) a unity variance we can make (this operation is taken to simplify this presentation, but the results still work if we consider both the numerator and the denominator) the following simplifications.

$$VR = \frac{\sigma_{\text{OUT}}^2}{\sigma_{\text{IN}}^2} = \frac{\bar{f}_{\text{OUT}}(n)}{f_{\text{IN}}(n)} = f_{\text{OUT}}(n).$$  \hspace{1cm} (A.2)

We note that the auto-correlation function is defined by

$$R_{ff}(m) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{n=-N}^{N} f(n) f(n+m),$$  \hspace{1cm} (A.3)

where, setting $m = 0$ yields Eq. (A.4). Therefore, the value of the auto-correlation function at $m = 0$ is equal to the variance of the output. Additionally, if we assume the variance of the input is 1, it also equals variance amplification ratio (VR):

$$R_{ff}(0) = \frac{1}{N} \sum_{n=0}^{N} f_{\text{OUT}}^2(n) = f_{\text{OUT}}^2(n).$$  \hspace{1cm} (A.4)

The auto-correlation function is closely related to the spectral density via

$$f_{\text{OUT}}^2(n) = R_{ff}(0) = \frac{1}{\pi} \int_{0}^{\infty} S_{ff}(\tilde{w}) \, d\tilde{w}.$$  \hspace{1cm} (A.5)

Since the spectral density, $S_{ff}(\tilde{w})$, is an even function of $\tilde{w}$, and it can be written in the form

$$S_{\text{OUT}}(\tilde{w}) = \frac{N^2(\tilde{w})}{D^*(\tilde{w})} = |F^*(j\tilde{w})|^2,$$  \hspace{1cm} (A.6)

i.e. the variance of the output of a linear system to a random stationary input of unit variance with a mean of zero is equal to the Bullwhip measure

$$\frac{1}{\pi} \int_{0}^{\infty} |F^*(j\tilde{w})|^2 \, d\tilde{w} = \frac{W_N}{\pi} = VR.$$  \hspace{1cm} (A.7)

Thus the relationships between bullwhip and various control theory concepts, if the variance of the input is unity, may be summarised by

$$VR = \frac{\sigma_{\text{OUT}}^2/\mu_{\text{OUT}}}{\sigma_{\text{IN}}^2/\mu_{\text{IN}}} = \frac{1/(N\mu_{\text{OUT}}) \sum_{n=0}^{N} (f_{\text{OUT}}(n) - \mu_{\text{OUT}})^2}{1/(N\mu_{\text{IN}}) \sum_{n=0}^{N} (f_{\text{IN}}(n) - \mu_{\text{IN}})^2} = f_{\text{OUT}}^2(n).$$  \hspace{1cm} (A.8)

References


