A technique to develop simplified and linearised models of complex dynamic supply chain systems

Virginia L. M. Spiegler\textsuperscript{a}, Mohamed M. Naim\textsuperscript{b}, Denis R. Towill\textsuperscript{b,*}, Joakim Wikner\textsuperscript{c}

\textsuperscript{a}Formerly Cardiff Business School now Brunel Business School, Brunel University, London, UB8 3PH, UK
\textsuperscript{b}Logistics Systems Dynamics Group, Cardiff Business School, Cardiff University, Aberconway Building, Colum Drive, Cardiff, CF10 3EU, UK
\textsuperscript{c}Department of Management and Engineering, Linköping University, Linköping SE-58183, Sweden

Abstract

There is a need to identify and categorise different types of nonlinearities that commonly appear in supply chain dynamics models, as well as establishing suitable methods for linearising and analysing each type of nonlinearity. In this paper simplification methods to reduce model complexity and to assist in gaining system dynamics insights are suggested. Hence, an outcome is the development of more accurate simplified linear representations of complex nonlinear supply chain models.

We use the highly cited Forrester production-distribution model as a benchmark supply chain system to study nonlinear control structures and apply appropriate analytical control theory methods. We then compare performances of the linearised model with numerical solutions of the original nonlinear model and with other previous research on the same model.

Findings suggest that more accurate linear approximations can be found. These simplified and linearised models enhance the understanding of the system dynamics and transient responses, especially for inventory and shipment responses.

A systematic method is provided for the rigorous analysis and design of nonlinear supply chain dynamics models, especially when overly simplistic linear relationship assumptions are not possible or appropriate. This is a precursor to robust control system optimisation.

\footnotesize{*This article is dedicated to the work, friendship and memory of Professor Denis R. Towill.}

December 3, 2015
1. Introduction

In supply chains, the variability in the ordering patterns often increases as one moves up the chain, towards the factory and the suppliers (Dejonckheere et al., 2003). This variance is called the bullwhip effect, “one of the most widely investigated phenomena in supply chain management” (Chatfield and Pritchard, 2013). Even to this day the seminal works of Forrester (1958, 1961), that formed the foundation for System Dynamics, are almost always referred to synonymously with the bullwhip effect (e.g. Zhang and Burke 2011; Chatfield and Pritchard 2013). In citing Forrester’s works authors refer to the original production-distribution model, which may now be classified as a representation of a supply chain, as a bullwhip generator archetype and the use of System Dynamics simulation as a technique for exploring opportunities to mitigate unwanted dynamic behaviour.

While System Dynamics simulation is often used in the analysis and redesign of supply chain models that exhibit nonlinearities, quantitative analytical approaches are more often restricted to linear representations of supply chains. Hence, much of the research on supply chain dynamics either takes a ‘trial and error’, experimental, simulation approach to redesign (Forrester, 1961; Sterman, 1989; Shukla et al., 2009; Poles, 2013; Spiegler and Naim, 2014) or develops exact solutions of models that are already linearised approximations to the real-world situation (Towill, 1982; John et al., 1994; Disney and Towill, 2005; Gaalman and Disney, 2009; Zhou et al., 2010).

While the original Forrester supply chain model is often quoted as the embodiment of the bullwhip effect it has had little exposure with respect to its use as a benchmark for applying supply chain analysis and redesign methods, with the notable exceptions of Wikner et al. (1992) and Jeong et al. (2000). The former explore a simplification approach to understanding the causes of the bullwhip effect (Wikner et al., 1992), while the latter apply a linearisation approach but with an analysis totally reliant on simulation. Analytical tools to link system dynamics model structures to different system modes of behaviour have recently become available and explored for linear models (Saleh et al., 2010). However, there is still a need to expand the existing body of knowledge regarding robust control of nonlinear supply chains. “Nonlinearity can introduce unexpected behaviour in a system” (Forrester, 1961), causing instability and uncertainty and therefore needing to be rigorously analysed.
Our paper aims to present a technique to develop both simplified and linearised models of complex, nonlinear supply chain systems. We seek to gain greater insights into the underlying mechanisms that create supply chain dynamics and to provide guidelines for undertaking system dynamics simulation in a time effective and productive way. The simplified and linearised form may also be a precursor to robust optimisation of nonlinear decision rules in supply chains, a gap in the existing body of knowledge due to the complexities of dealing with seemingly intractable mathematics. We utilise the original Forrester model as a benchmark, as per Wikner et al. (1992) and Jeong et al. (2000), given it is a complex representation of a production-inventory control system with nonlinearities and it is highly cited for describing the behavioural dynamics of supply chains. The currency of the model is also evidenced by 333 citations related to supply chains in 2013-2014 (according to Google Scholar searched on 27 August 2014) and by Singhal and Singhal (2012) noting that Forrester’s papers still represent well the real-world phenomenon of fluctuations and oscillations since it contains a combination of simulated data and case study data to examine the flow of materials and information in a supply chain.

In summary, we aim to determine the methodological benefits of nonlinear control theory in supporting simulation based research on supply chain dynamics studies. This research is particularly relevant to operations research scholars exploring nonlinear dynamic systems. Moreover, future applications of the proposed technique may also benefit practitioners in improving supply chain performance. As Ivanov and Sokolov (2013) pointed out “useful tools for quantitative analysis of control and systems theory for a wide supply chain management research community remain undiscovered”. Our work addresses this shortcoming.

2. Nonlinear system dynamics

A nonlinear system is one whose performance does not obey the principal of superposition. This means that the output of a nonlinear system is not directly proportional to the input and the variables to be solved cannot be expressed as a linear combination of the independent parts (Atherton, 1975). In this section, we briefly review methods for analysing nonlinear system dynamics and highlight where certain methods have already been used in supply chain dynamics research.
When confronted with a nonlinear system the primary approach utilised by system designers is to identify an equivalent linear representation. A justification for this is that there are a variety of analytical techniques available in linear control theory that are not so readily applicable in the analysis of nonlinear systems. While linear control theory is well established, the literature lacks a unique nonlinear theory that strives for generality and applicability (Rugh, 2002; Hotz and Vogel, 2014).

The lack of generality, coupled with often indefinite research methods, has led to a confusion of terminologies making it a challenge to determine a listing of all existing techniques and their applicability in the analysis of nonlinear feedback systems. Table 1 lists those methods that have been sufficiently acknowledged in the literature. The table highlights the type of nonlinearity that each method addresses, the assumptions or limitations that need due consideration in their application and citations of where they have been applied in a supply chain dynamics context. The choice of each method may also depend on the degree of complexity involved in the setting up of a mathematical model, the type of data available for analysis and the analytical skills of the researcher or supply chain designer.

Most research on nonlinear supply chain systems has been undertaken via simulation methods. Table 1 gives only a small number of examples from a plethora of papers that utilise simulation to analyse nonlinearities in supply chains. This research has led to the understanding of particular phenomena; such as:

- stability and chaos (Larsen et al., 1999; Laugesen and Mosekilde, 2006),

- the impact of capacity and batching constraints (e.g. Paik and Bagchi 2007; Cannella et al. 2008; Juntunen and Juga 2009; Hamdouch 2011; Ivanov et al. 2014), inaccuracies in inventory (Cannella et al., 2015), reverse logistics (Turrisi et al., 2013) and collaborative strategies (e.g. Cannella and Ciancimino 2010; Spiegler and Naim 2014) on system dynamics and supply chain performance,

- bullwhip effect in service supply chains Akkermans and Voss (2013)

- shipment planning (Shukla et al., 2009; Mula et al., 2013) and

- the effects of psychological pressure, misperceptions and misjudgement in work environments (Sterman, 1989; Syntetos et al., 2011; Bruccoleri et al., 2014).
Table 1: Summary of methods to analyse nonlinear supply chain systems

<table>
<thead>
<tr>
<th>General Approach</th>
<th>Method of Analysis</th>
<th>Types of Nonlinearities</th>
<th>Considerations/ Assumptions</th>
<th>Supply Chain Applications</th>
<th>Our contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linearisation methods</td>
<td>Small Perturbation Theory with Taylor series expansion</td>
<td>Continuous Single-valued</td>
<td>Assumption that the amplitude of the excitation signal is small Local stability analysis only</td>
<td>Limited application in supply chain context (<a href="#">Jeong et al., 2006</a>) Recommends but does not apply for supply chain design (<a href="#">Saleh et al., 2010</a>)</td>
<td>Fuller application for supply chain design and understanding of the impact of time-varying parameters Full application for understanding the impact of discontinuous, single- and multi-valued nonlinearities present in supply chain models</td>
</tr>
<tr>
<td>Describing Function</td>
<td>Continuous Discontinuous Single-valued Multi-valued</td>
<td></td>
<td>None identified</td>
<td>None identified</td>
<td></td>
</tr>
<tr>
<td>Small Perturbation Theory with Volterra/Wiener series expansion</td>
<td>Continuous Multi-valued</td>
<td></td>
<td>None identified</td>
<td>None identified</td>
<td></td>
</tr>
<tr>
<td>Averaging and best-fit line approximations</td>
<td>Continuous Discontinuous Single-valued Multi-valued</td>
<td></td>
<td>Testing supply chain re-engineering strategies (<a href="#">Wikner et al., 1992</a>) Identifying analogies between seemingly different decision rules (<a href="#">Naim et al., 2012</a>)</td>
<td>None identified</td>
<td></td>
</tr>
<tr>
<td>Graphical and numerical methods</td>
<td>Phase Plane and Graphical Solutions</td>
<td>Continuous Single-valued Multi-valued</td>
<td>Limitied to 1st and 2nd order systems</td>
<td>None identified</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Point transformation method</td>
<td>Discontinuous Single-valued Multi-valued</td>
<td>Piecewise linear systems only For high order systems, automated numerical methods must be employed</td>
<td>Exploring nonlinear behaviour of inventory systems (<a href="#">Wang et al., 2014</a>)</td>
<td></td>
</tr>
<tr>
<td>Exact solutions</td>
<td>Direct Solution</td>
<td>Continuous Single-valued</td>
<td>Limited to a finite number of equations</td>
<td>None identified</td>
<td></td>
</tr>
<tr>
<td>Complex and detailed methods</td>
<td>Lyapunov-based stability analysis for piecewise-linear systems</td>
<td>Discontinuous Only single-valued exceptions were found</td>
<td>Piecewise linear systems only Computation can be complex depending on the system</td>
<td>Stability analysis of inventory systems (<a href="#">Wang and Disney, 2012</a>)</td>
<td></td>
</tr>
<tr>
<td>Simulation</td>
<td>Computational and Simulation solutions</td>
<td>Continuous Discontinuous Single-valued Multi-valued</td>
<td>Trial and error approach wide spectrum of control parameters Can be time consuming Dependent on computer and software calculations capacity</td>
<td>the effects of psychological pressure, misperceptions and miscalculations in work environments (<a href="#">Sternan, 1989</a>; Syntetos et al., 2011; Bruccoleri et al., 2014); Shaping stability regions of nonlinear systems (<a href="#">Larsen et al., 1999</a>; Langsen and Monskilde, 2006); Robust control of supply chain decision rules (<a href="#">Jeong et al., 2000</a>) Bullwhip and backlash ([Shukla et al., 2009; Akkermans and Voss, 2013; Mula et al., 2013] Effect of capacity and batching ([Park and Dozzi, 2007; Cannella et al., 2008; Hanseleit, 2011], inaccuracies of inventory (<a href="#">Cannella et al., 2015</a>)); reverse logistics (<a href="#">Turrisi et al., 2013</a>); Improve nonlinear systems performance (<a href="#">Spiegler and Naim, 2014</a>)</td>
<td>Application for cross-checking results obtained via analytical methods</td>
</tr>
</tbody>
</table>

In contrast, there is limited research on the use of analytical methods. Many of the analytical studies on nonlinear system dynamics were undertaken in the same decade Forrester launched the World Dynamics model ([Forrester, 1971](#)), which is a simpler model when
comparing to the production-distribution model. Cuypers (1973) used averaging techniques for linearising discontinuous nonlinearities in the World Dynamics model. One year later, numerical perturbation techniques and model simplification, involving the removal of variables with little variation, were also explored (Cuypers and Rademaker, 1974). Ratnatunga and Sharp (1976) proposed the use of numerical analysis to linearise and reduce orders of system assuming that nonlinear associations can be approximated to a first order function. Mobapatra (1980) identified and categorised different types of nonlinearities in business system dynamics research. Although his work recommends a number of techniques to deal with nonlinearities, there is no implementation of such methods in the paper.

Within a supply chain management context, Wikner et al. (1992) undertook in-depth analysis of the complex Forrester production-distribution model (1961). By using averaging techniques and block diagram manipulation, they linearised and simplified the original model and provided more qualitative analytical insights. For example, they highlighted the lack of feedback information fed into the manufacturing rate and the separation of ‘real’ and ‘safety’ orders. By following the same simplification and linearisation steps, Naim et al. (2012) achieved the same result for the discrete z-domain model. In contrast, instead of using an averaging technique, Jeong et al. (2000) used small perturbation theory to linearise the continuous nonlinearities in the Forrester model and Matsubara’s time delay theorem to obtain a first-order delay approximation to represent an upstream echelon.

Another advocate of the use of analytical methods rather than just the use of exhaustive repeated simulation are Saleh et al. (2010). They suggest the use of small perturbation theory to perform structured policy analysis. However, their analytical framework disregards the presence of discontinuous nonlinearities. On the other hand, Wang and Disney (2012) and Wang et al. (2014) used graphical and eigenvalue methods to explore stability boundaries of a piecewise linear inventory control system and to identify a set of behaviours in the unstable region. Their work is limited to the analysis of a single-valued discontinuous nonlinearity given by a non-negative constraint on the replenishment order.

Our paper specifically addresses the gap in the use of describing functions in supply chain dynamics modelling to analyse both single- and multi-valued discontinuities. Moreover, our work focuses on understanding behaviour within the system’s stable region. In analysing the full set of nonlinearities we find in the Forrester model we couple describing functions with
small perturbation theory using Taylor series expansion. Although Saleh et al. (2010) recommend the use of small perturbation theory, application of this method has only been found in Jeong et al. (2000) to linearise the continuous, single-valued nonlinearities in the Forrester model. Despite their efforts to linearise part of the model, they use solely simulation methods to analyse the effect of different capacity levels, or discontinuous nonlinearities, in the factory’s production rate. Due to their lack of explicit research objectives and methodological structure, their contribution to supply chain dynamics research is unclear and perhaps explains the lack of citation to their paper. As we will see, the small perturbation theory will overcome the accuracy disadvantages seen in the use of averaging techniques previously utilised in supply chain modelling by Wikner et al. (1992); Naim et al. (2012).

Although system design is not within the scope of this paper, the system’s transfer functions, damping ratio and natural frequencies will be presented for future evaluation of the impact of each control parameter on the supply chain performance. After linearisation is performed, the reader can refer to Saleh et al. (2010) and other relevant works (Disney et al., 2006; Ouyang and Daganzo, 2008; Zhou et al., 2010) for robust analysis and design of linear models.

3. Previous simplification and linearisation of Forrester’s production-distribution model

As previously mentioned, Wikner et al. (1992) endeavoured to gain more insights into the Forrester model by a two-stage linearisation and simplification approach. More importantly, their work translated the Forrester DYNAMO equations into Laplace domain block diagram format to improve visibility of the model’s system structure. Figure 1 illustrates the block diagram representation of a single echelon, namely the factory, of the Forrester model and the steps taken by Wikner et al. (1992) in simplifying the original model. Appendix A in the supplement file contains the listing and explanation of constants, variables and equations. As is commonly found in such research a single-echelon as the unit of analysis suffices in determining the structural causes of the bullwhip effect (e.g. Disney et al. 2006; Ouyang and Daganzo 2008).

The first step taken by Wikner et al. (1992) was to translate the original DYNAMO equations into a control engineering block diagram representation and to identify any non-
Figure 1: Wikner at al. (1992) approach for simplifying Forrester model linearities as shown in Figure 1(a). Then they removed the discontinuous nonlinearities, represented by the CLIP functions ( ), by assuming that the capacity limitations in manufacturing (AL) and shipping (IA) are never attained under ‘normal’ operating conditions, so that the wanted manufacturing rate is equal to the actual one (MW=MD) and shipments
tried are equal to shipments sent \((ST=SS)\) at all times. Note that while \(AL\) is a fixed capacity constraint \(IA\) is variable.

The continuous nonlinearity, caused by the nonlinear comparator ratio \((\Theta)\) in defining the delay in filling orders \((DF)\), was considered the average value of this delay and then kept fixed as shown in Figure 1(b), Step 2. Lastly in Step 3, by block diagram manipulation Wikner et al. (1992) collected constants and eliminated redundancies in the original model, resulting in the final block diagram of Figure 1(c). Step 3 is further explained in Section 4.1.1.

The resulting model highlights the lack of feedback information fed into the manufacturing rate \((MD)\) and reveals a separation between ‘real’ and ‘safety’ orders. Hence, Wikner et al. (1992) showed that the so-called ‘Forrester effect’, in which orders are amplified from sink to source, is not due to linear feedback control but due to a first-order derivative term, \(s\), in the feedforward path. As their study was focused on the bullwhip effect, the main problem with their model is that while accuracy is kept for analysing the manufacturing rate, their linearised and simplified model is less reliable for analysing inventory \((IA)\) and shipments \((SS)\) responses due to the use of averaging techniques for linearisation.

In this work we will show that it is possible to use Taylor series expansion to represent some of the nonlinearities present in the original model and linearise it with small perturbation theory to get better accuracy. Moreover, we do not neglect the presence of CLIP functions and we investigate their effect in the overall response of the model with the describing function method.

4. Our approach to simplification and linearisation

Wikner et al.’s (1992) approach consisted of conducting linearisation before simplification, as also seen in Jeong et al. (2000); Naim et al. (2012). However, causal relationships between certain variables may be lost during the linearisation process. For instance when cross-referencing Figures 1(a) and 1(b) of Wikner et al.’s (1992) model, the effect of \(IA\) and \(AI\) (constant for inventory) on \(SS\) has been omitted after linearisation. Hence, we propose that models should be simplified first, by eliminating all redundancies whenever possible. Then later, having clearer visibility of the inherent model structure, analysis and synthesis of the nonlinear elements are undertaken.
To summarise, Figure 2 illustrates the steps taken in this paper to simplify and linearise the original Forrester model in comparison to the approach used by Wikner et al. (1992). We start by taking the Forrester model equations and representing them in block diagram form. Then we take two steps to simplify the model: block diagram manipulation to remove redundancies found in the original model and low order modelling to reduce the order of the differential equation. Next, the linearisation process is initiated by identifying the types of nonlinearities present in the simplified model. Later, we use small perturbation theory to linearise continuous nonlinearities and describing functions to analyse discontinuous nonlinearities. Finally, we compare our simplified, low order and linearised model with the original Forrester and Wikner et al.’s (1992) models.

Figure 2: Research method

4.1. Simplification

In this section, we explain in detail how block diagram manipulation and low order modelling techniques can help simplify the Forrester model.
4.1.1. Block diagram manipulation

The original block diagram of Figure 1(a) can be manipulated so that redundancies are removed and constants are collected as given in Figure 3(a). The sequence of steps taken were:

**Redundancy 1:** calculation of actual pipeline orders in transit (LA). The figure shows that the information about manufacturing orders (MO) is being added and reduced at the same time when calculating LA; therefore the information about MO is redundant in the determination of LA.

**Collecting Constants:** gathering constants from the output variable in the summing comparator $SUM_1$. In the $SUM_1$ expression, all the constants that multiply the variable smoothed requisition orders (RS) were combined together and called K, which is equal to $DC + DP - DH - DU + AI$.

**Redundancy 2:** calculation of the output variable in the summing comparator $SUM_2$. In the calculation of $SUM_2$ both shipment received (SR) and shipment sent (SS) information is found to be redundant; therefore, they can be removed at that summation point.

After removing all the redundancies and combining the constants in K, the block diagram in Figure 3(b) is then presented. Note that the resulting simplified non-linear model in Figure 3(a) provides exactly the same responses as in the original model. No variable interactions were lost in this simplification process. Moreover, it can now be seen that the Forrester model contains a feedback loop in the manufacturing order rate (MD) but only to provide information regarding the manufacturing capacity. If the system cannot manufacture the amount wanted in a particular time period, this information is then fed back so that these orders can be produced later. However, if the capacity limitation is never reached, this feedback information is not needed and hence it can be ignored. As also evidenced by Wikner et al. (1992) inventory information is not fed back into manufacturing orders.

4.1.2. Low order modelling

A high-order control system often contains poles that produce little effect on the transient response. In the pipeline of Figure 3(b), represented by the sixth-order transfer function
(a) Forrester model: Original - removing redundancies and collecting constants

\[
\begin{align*}
S_{MM} &= \frac{1}{(1+SR-SA) + K \cdot RS - \frac{1}{(SR-SA) - \frac{1}{(MD-SA)}}} \\
S_{MM} &= \frac{1}{(1+SR-SA) + K \cdot RS - \frac{1}{(SR-SA) - \frac{1}{(MD-SA)}}}
\end{align*}
\]

(b) Forrester model: simplified and nonlinear

\[
\left(\frac{1}{1+DR_s}\right)^3 \cdot \left(\frac{1}{1+DR_s}\right)^3
\]

with selected parameter settings \((DC = 1\) and \(DP = 6\)), the delay \(DC\) has little impact on the transient response of the pipeline. This is due to the position of the poles in the s-plane. The reader can refer to Appendix B in the supplement file to review the methods suggested by Towill (1981) and Kuo and Golnaraghi (2003). The former method is extended from the time delay theorem developed by Matsubara (1965). This method attempts to determine a low-order model based on the system unit step response.

Figure 3: Simplification process of Forrester model
and has already been used by Jeong et al. (2000) to approximate high-order delays in the Forrester model. Kuo and Golnaraghi (2003) recommend a method proposed by Hsia (1972) that approximates a high-order system to a low-order model by approaching their frequency responses.

We compared the Matsubara and Hsia methods to find the best low order model of the Forrester pipeline, and we found that Matsubara method provides better results for both step and sinusoidal inputs. For this reason, we use the Matsubara method for simplifying the Forrester model further. Hence, in the following analysis we consider the Matsubara first-order plus dummy pole pipeline approximation in the form of $T_M(s)$, which is:

$$T_M(s) = \frac{1}{(1 + \frac{DP}{3}s)(1 + (DC + \frac{2}{3}DP)s)}$$

(1)

The derivation of Equation 1 can be found in subsection a) of Appendix B in the supplement file.

4.2. Linearisation and quasi-linearisation

In this section, analysis of the nonlinearities in the Forrester production-distribution model will be undertaken. We analyse the continuous and discontinuous nonlinearities separately since each of them requires different linearisation methods.

4.2.1. Analysis of continuous nonlinearities

We start our analysis by temporarily assuming that the CLIP functions are not active (Figure 4). In other words, the manufacturing rate decision will be equal to the manufacturing rate wanted, MD=MW, and the shipment sent will be the same as the shipment tried, SS=ST, independent of actual inventory levels.

Hence, the system in Figure 4 can be described by the system of differential equations (Equations 2-8), where $\dot{r}s, \dot{r}s_s, \dot{r}md, \dot{s}r, \dot{u}o$ and $\dot{i}a$ are the state variables, $\dot{x} = f(x, u)$, of the system and $rr$ is the system input, $u$. Note that, since the six-order pipeline was replaced by the lower-order equation in Equation 1, four other states have been excluded. The state variables $rs_s$ and $md_d$, representing dummy variables $RS_S$ and $MD_D$ respectively, have been added to help derive the state variable equations below:
\begin{align*}
\dot{r}s &= f_1(x, u) = \frac{rr - rs}{DR} \\
\dot{r}s_s &= f_2(x, u) = \frac{rs - rs_s}{DI} \\
\dot{md}_d &= f_3(x, u) = \frac{3K.rs}{DI.DP} - \frac{3K.rs_s}{DI.DP} - \frac{3md_d}{DP} + \frac{3rr}{DP} \\
\dot{sr} &= f_4(x, u) = \frac{3(md_d - sr)}{3DC + 2DP} \\
\dot{uo} &= f_5(x, u) = rr - ss = rr - \frac{uo.ia}{AI.DU.rs + DH.ia} \\
\dot{ia} &= f_6(x, u) = sr - ss = sr - \frac{uo.ia}{AI.DU.rs + DH.ia} \\
rs(0) &= a(0) = b(0) = sr(0) = rr(0), ia(0) = AI.rr(0), \\
uo(0) &= (DH + DU)rr(0)
\end{align*}

The outputs, \( y = g(x, u) \), of interest are the manufacturing rate, \( MD \), the actual inventory levels, \( IA \), and shipment sent, \( SS \). In addition to these outputs, it is interesting to know how the time-varying parameter \( DF \) will be affected after the linearisation.
\[ md = g_1(x, u) = rr + \frac{K_{rs}}{DI} - \frac{K_{rs}}{DI} \]

\[ ss = g_2(x, u) = \frac{uo.ia}{AI.DU.rs + DH.ia} \]

\[ ia = g_3(x, u) = ia \]

\[ df = g_4(x, u) = \frac{AI.rs}{ia}.DU + DH \]

The mathematical model above is nonlinear due to the presence of nonlinear algebraic differential equations, which in the block diagram are represented by the symbol \( \oplus \). The overall model can be linearised about a nominal operating state space \( x^* \) and for a given input \( u^* \) by using small perturbation with Taylor series expansion. The first order Taylor series approximation of the nonlinear state derivatives leads to the following linearised function:

\[ \Delta \dot{x} = A\Delta x + B\Delta u \]

\[ \Delta y = C\Delta x + D\Delta u \]

where \( \Delta x = x - x^*, \Delta y = y - y^*, \Delta u = u - u^* \). The equilibrium or resting points \((x^*, u^*)\) can be determined by the final value theorem and A, B, C, D can be found through the partial derivatives of the state space equations, which result in the matrix given by Equation 15. Equation 15 can then be converted back to a block diagram representation as in Figure 5. Note that, in the resulting matrix only \( DF \) is input-dependent, and hence it could not be represented in the block diagram of Figure 5.

\[
\begin{pmatrix}
\frac{1}{\text{DR}} & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{\text{DR}} & \frac{1}{\text{DR}} & 0 & 0 & 0 & 0 \\
\frac{1}{\text{DR}} & \frac{1}{\text{DR}} & \frac{1}{\text{DR}} & 0 & 0 & 0 \\
0 & 0 & \frac{1}{\text{DC} + \text{2DP}} & \frac{-3}{\text{DC} + \text{2DP}} & 0 & 0 \\
0 & 0 & \frac{1}{\text{DC} + \text{2DP}} & \frac{-3}{\text{DC} + \text{2DP}} & 0 & 1 \\
0 & 0 & 0 & 1 & \frac{1}{\text{DC} + \text{2DP}} & \frac{1}{\text{DC} + \text{2DP}} \\
0 & 0 & 0 & 1 & \frac{1}{\text{DC} + \text{2DP}} & \frac{1}{\text{DC} + \text{2DP}} \\
0 & 0 & 0 & 0 & 0 & 1 \\
-\frac{1}{\text{DR}} & -\frac{K}{\text{DR}} & 0 & 0 & 0 & 0 \\
-\frac{1}{\text{DR}} & 0 & 0 & 1 & \frac{1}{\text{DC} + \text{2DP}} & \frac{1}{\text{DC} + \text{2DP}} \\
-\frac{1}{\text{DR}} & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

\[ (A \mid B) = \]

\[ (A \mid C) = \]

\[ (B \mid D) = \]
When comparing Figures 4 and 5, it can be seen that after linearisation the product functions ($\Pi$) are replaced by summing comparators ($\Sigma$) after linearisation. Moreover, the system transfer functions can now be determined and linear control theory can be used for analysing the impact of each control parameter and improving system design (please refer to Appendix C in the supplement file). The separation between ‘real’ and ‘safety’ orders is well evidenced in the MD transfer function in Equation C1. Hence, bullwhip does not occur due to feedback control, but due a derivative term in the numerator of the transfer function.

![Figure 5: Forrester model: Simplified and linearised with small perturbation theory](image)

Figure 5: Forrester model: Simplified and linearised with small perturbation theory

Figure 6 illustrates unit step and sinusoidal responses in changes in manufacturing rate ($\Delta MD$), inventory ($\Delta IA$) and shipment sent ($\Delta SS$) from their initial states and the delay in filling orders ($DF$), comparing the output responses of the original model with averaging (Wikner et al., 1992) and small perturbation linearisation techniques. Figure 6 demonstrates the power of small perturbation theory for the analysis of continuous nonlinearities. Accuracy is increased as the amplitude of the input signals is decreased although in all cases the small perturbation theory approach gives a better approximation than the averaging method.

4.2.2. Analysis of discontinuous nonlinearities

Now we re-insert the CLIP functions into the linearised model resulting in Figure 7. Note that when re-inserting the CLIP functions, two feedback loops have also been re-instated. We then analyse each of the discontinuous nonlinearities separately.

**Manufacturing constraint**: if a sinusoidal input, MW

$$MW(t) = A.\cos(\omega t) + B$$  \hspace{1cm} (16)

with amplitude A, mean B and angular frequency, $\omega = 2\pi/T$, is the input into the nonlinear-
Figure 6: Comparing Forrester, Wikner et al.’s (1992) and linearised model using small perturbation theory.

Manufacturing rate (MD)

Actual inventory (IA)

Shipment sent (SS)

Delay in filling order (DF)

Forrester Model
Linearised
Wikner et al. Model

Forrester, Wikner et al.’s (1992) and linearised model using small perturbation theory.

ity, which consists of a saturation function with a maximum limit value AL, an output MD will result. MD does not depend on the past values of the input MW, but it varies according to the actual state of MW. Although the function is nonlinear, it can be represented by two
Figure 7: Forrester model: Simplified, containing only discontinuous nonlinearities

piecewise linear equations:

$$MD(t) = \begin{cases} MW(t) & \text{if } MW < AL \\ AL & \text{if } MW \geq AL \end{cases}$$ (17)

By investigating the block diagram again in Figure 7, it is clear that if the output MD differs from demand RR, this error will be accumulated due to the presence of an integrator $1/s$ in the feedback loop. Hence, the manufacturing rate, MD, will only align with the demand if the manufacturing capacity AL is at least equal to the average demand. If manufacturing capacity is less than the required demand, RR, then the manufacturing wanted, MW, will increase exponentially and the system will never stabilise.

In order to investigate the impact of discontinuous nonlinearities in feedback systems, we can use the describing function method, which is a quasi-linear representation for a nonlinear element subjected to a sinusoidal input.

For asymmetric nonlinearities at least two terms of the describing function are needed: one that describes the change in the output amplitude ($N_A$) as the input amplitude increases or the saturation value decreases, and another term that determines the change in the output mean ($N_B$). We can also investigate the change in phase angle ($\phi$) of the output response in relation to its input. Given MW as input, the output MD can be approximated to:

$$MD(t) = N_A.A.cos(\omega t + \phi) + N_B.B$$ (18)

In order to determine the terms of the describing function ($N_A$, $N_B$ and $\phi$) the series
have to be expanded and its first harmonic coefficients must be determined. The Fourier series expansion method is used to represent the output $MD$ as a series such as:

$$MD(t) \approx b_0 + a_1 \cos(\omega t) + b_1 \sin(\omega t) + a_2 \cos(2\omega t) + b_2 \sin(2\omega t) + \cdots =$$

$$\approx b_0 + \sum_{k=1}^{\infty} [a_k \cos(k.\omega t) + b_k \sin(k.\omega t)]$$

where $a_k$, $b_k$ and $b_0$ are the Fourier coefficients and $MD$ is the piecewise linear function given in Equation 17.

If we approximate the piecewise linear output MD to the first harmonic, we have:

$$MD(t) = b_0 + a_1 \cos(\omega t) + b_1 \sin(\omega t) = b_0 + \sqrt{a_1^2 + b_1^2} \cdot \cos(\omega t + \phi)$$

where, $\phi = \arctan \left( \frac{b_1}{a_1} \right)$

In this way the two terms of the describing function can be determined as:

$$N_A = \frac{\sqrt{a_1^2 + b_1^2}}{A}$$

$$N_B = \frac{b_0}{B}$$

For single-valued nonlinearities the coefficient $b_1$, the imaginary part of the describing function, will be equal to zero and therefore the phase angle will be also zero. Hence, for the asymmetric saturation in the Forrester system it is found that:

$$N_A = \frac{-\gamma + \pi + \cos\gamma \cdot \sin\gamma}{\pi}$$

$$N_B = \frac{B \cdot \pi + A \cdot \gamma \cdot \cos\gamma - A \cdot \sin\gamma}{B \cdot \pi}$$

where $\gamma = \cos^{-1} \left( \frac{AL-B}{A} \right)$.

Figure 8 illustrates how the coefficients of the describing function for the manufacturing capacity vary as the amplitude of manufacturing wanted rate, $A_{MW}$, increases. For amplitudes lower than the capacity, AL, the system behaves as linear and output MD will be equal to the input MW corresponding to a describing function gain equal to 1. However, when MW hits the capacity AL we have to consider that only a fraction of this rate will actually
be manufactured.

![Graph](image)

Figure 8: Terms of describing function for MD with asymmetric saturation

Figure 9 demonstrates the changes in the sinusoidal responses of the change in manufacturing orders ($\Delta MD$) comparing the original Forrester model with Wikner et al.’s (1992) model and the describing function method for different frequencies, noting the different scales in the x-axis. Again we have found a better linear approximation for the original model.

**Shipmenet constraint:** the CLIP function in the shipment system is used to avoid any shipments being made to customers if no inventory is actually available. Hence, shipments sent, SS, will be equal to shipment tried, ST, only if actual inventory, IA, is greater than ST.

This second nonlinearity is different from the discontinuity in the manufacturing system in that it is not only amplitude-dependent but also frequency-dependent. In Figure 10, we explore a set of system responses for inventory and shipments for different amplitude and frequencies, again noting the different scales in the x-axis. In the example in Figure 10, it seems that higher frequencies and lower amplitudes (Figure 10a) result in a seemingly linear response. Hence, SS will be equal to ST, corresponding to a describing function of 1. However, for lower frequencies and increased amplitude (Figure 10d), the shipment capacity is reduced and a complex nonlinear behaviour is observed. Figure 10 also illustrates that this nonlinearity is multi-valued, as may be noted by the insets on the top right corner of each figure, which shows the variations in the output amplitude, on the y-axis, for a given input amplitude, on the x-axis. For a given input ST the output SS can assume different values depending on the past states of ST.
Figure 9: Comparing Forrester, Wikner et al.’s (1992) and quasi-linearised model using a Describing Function for MD

Figure 10: Actual inventory (IA), shipment tried (ST) and sent (SS) responses to different amplitudes and frequencies
Since this discontinuous nonlinearity is frequency-dependent, there will be one describing function for each frequency. We used Matlab® combined with Simulink® to find the describing function corresponding to the amplitude gain and also to identify the phase shift, resulting in Figure 11.

Figure 11 confirms that the nonlinearity in the shipment process only occurs for very low frequencies and high amplitudes. Another important factor is regarding the target inventory, AI, as inventory levels decrease the nonlinearity becomes more pronounced. With the nominal values for AI in the Forrester model set between 4 to 8 weeks, then the Wikner et al. (1992) representation is reasonable because the CLIP function will only take effect at extremely low frequencies and high demand amplitudes. The system designer does not have to be concerned with the shipment constraint when demand has medium to high frequencies and low amplitudes.

![Figure 11: Describing Function amplitude gain and phase in relation to ST amplitude and frequency](image)

In Figure 12 we compare the change in shipment sent, ΔSS, due to sinusoidal inputs of two distinct frequencies, of the original Forrester model, the Wikner et al.'s (1992) representation and the describing function method when the system reaches steady state. The amplitude is fixed to 2 for both cases. Note that for this comparison we have set the value of
AL to a high enough value so that the manufacturing constraint is inactive and MD=MW. For both responses in Figure 12 the describing function method provides a better approximation to the Forrester model response than that achieved by Wikner et al. (1992). Although the differences between the three models are not significant, the describing function method provides a better understanding of the shipment constraint in relation to its effect on the output phase and amplitude shift for certain input frequencies and amplitudes.

![Figure 12: Comparing Forrester, Wikner et al.’s (1992) and quasi-linearised models using a Describing Function for SS with AL such that MD=MW](image)

Finally, Figure 13 illustrates the change in inventory ($\Delta I_A$) when all discontinuous nonlinearities are in effect. The inventory response given by the Forrester model does not return to the target value, $\Delta\text{mean}=0$, due to the capacity constraints. Our model was able to track this behaviour, while the Wikner et al.’s (1992) model was incapable of replicating the original behaviour since they disregarded the CLIP functions. As the demand frequency decreases and amplitude increases the nonlinear responses in the original Forrester become more acute and hence the Wikner et al.’s (1992) linearised model becomes less accurate.

Further analysis can be undertaken using the describing function method to design supply chain systems. Appendix D in the supplement file explains how to use the root locus technique in combination with the describing function method to predict the effect of discontinuous nonlinearities on the system’s natural frequency ($\omega_n$) and damping ratio ($\zeta$). $\omega_n$ determines how fast the system oscillates during the transient response, while $\zeta$ describes how much the system oscillates as the response decays towards steady state.

We found that while both manufacturing and shipment capacity constraints equally decrease $\omega_n$, the manufacturing capacity has the most significant impact on $\zeta$ by damping the
system’s response and making it slower. Hence, the nonlinearity present in the shipment process is not much of a concern for the supply chain designer.

5. Conclusion

To date, simulation techniques have been mainly used to deal with the dynamics of complex, nonlinear supply chain systems. However, this research suggests a more rigorous approach that permits mathematical analysis of nonlinearities as precursor for simulation experiments and any subsequent robust system optimisation.

For this reason this research identified and categorised the different types of nonlinearities that may appear in supply chain dynamics models in order to suggest suitable analytical methods for investigating each type of nonlinearity. Moreover, simplification techniques have also been used to reduce model complexity and to assist in gaining system dynamics insights. Hence, another outcome of this research was the development of a methodological framework to obtain more accurate simplified linear representations of complex nonlinear supply chain models by using nonlinear control theory. Better accuracy of our proposed approach is a result of firstly simplifying the original model so that complexity can be removed (Figure 14). This step facilitated the application of advanced linearisation methods that improved the precision of the system’s responses and the understanding of the impact of each nonlinearity on the system’s behaviour.

The well-known Forrester model has been used as a benchmark supply chain system

![Figure 13: Comparing Forrester, Wikner et al.’s (1992) and quasi-linearised models for IA when all CLIP functions are active](image)

Figure 13: Comparing Forrester, Wikner et al.’s (1992) and quasi-linearised models for IA when all CLIP functions are active
to study nonlinear control structures and to experiment with the application of low order modelling, small perturbation and describing function methods. In this model, we have identified all types of nonlinearities, although not all combinations of nonlinearities. Continuous nonlinearities were used by Forrester to represent a time-varying parameter: delay in filling order (DF), which increases substantially as inventory approaches zero. Discontinuous nonlinearities were used to represent manufacturing and shipment constraints. Both types of nonlinearities can be represented by piecewise linear functions. Single-valued nonlinearities were used to represent fixed capacity values in the manufacturing while multi-valued nonlinearities emerged from variable capacities in shipments. Performances of the linearised models have been compared with numerical solutions of the original Forrester model. Moreover, these techniques have also been compared to simple averaging and best-fit line approximation advocated and applied by previous research (Wikner et al., 1992; Naim et al., 2012).

The use of this approach brought a number of insights to bear on the understanding of the system dynamics behaviour and how each nonlinearity affects responses. Table 2 summarises the analytical insights obtained with the use of nonlinear control theory, the suggestions for further simulation experiments and the possible implications of not conducting a mathematical analysis before simulation.

The simplification techniques, such as block diagram manipulation and low order modelling, have contributed to providing a better visualisation and understanding of the variable interactions in the model. Moreover, by reducing the number of equations and orders, these
Analytical insights | Proposed simulation experiments | Potential consequences if not undertaken
---|---|---
**• Better visualisation and understanding of the model’s relevant variables, constants and their relationship.** | **• Since system policies can be more easily identified, simulations can be carried out only for specific control parameters.** | **• Simulation would be a slow process since the relationships between variables were not well understood.**

**Example 1:** It has been identified that no inventory information is fed back to the ordering policy and that bullwhip is not caused by feedback loops.

**Example 2:** The only feedback information in the ordering process is due to the manufacturing capacity.

**Example 3:** It has been identified that in the Forrester model there is a target value for unfilled orders.

**• Reduction of equations, which enabled the application of small perturbation theory and describing functions techniques.** | **• Simulations were undertaken with original equations to compare results with analytical, simplified and linearised models.** | **• The linearisation process of the system would be more difficult given the number of equations.**

**Example 1:** Eight variables were eliminated from the original Forrester model without any loss of accuracy.

**Example 2:** The linearisation process of the system would be more difficult given the number of equations.

**Block diagram manipulation**

**• Better understanding of the delays involved in the shipment receipt** | **• Sensitivity analysis via repeated simulations can be carried out to check the impact of increased delays on system’s performance.** | **• Unnecessary simulations may be conducted.**

**Example 1:** It was found that both DC and DP have the same effect on the shipment receipt delay. Hence there is no need to simulate different values for each parameter separately.

**Example 2:** Small values of DI and DR will generally increase speed of responses. Simulation can be used for trade-off and sensitivity analysis.

**Example 3:** It is possible to find local stability boundaries. Simulations can be undertaken only within the pre-determined stability boundaries.

**Example 1:** Analysis shows that manufacturing constraints always provoke negative impact on system’s damping ratio. Simulation can be used to further investigate this effect.

**Example 2:** Analysis showed that shipment constraint decreases the system’s natural frequency but slightly improves the damping ratio. Simulation can be used to further investigate this effect.

**Low order modelling**

**• Possibility to find system transfer functions.** | **• Simulations can focus only on important parameters for achieving desired supply chain responses** | **• Greater effort would be needed to apply more rigorous analytical methods. High order models may demotivate the researcher who may opt for a trial and error approach.**

**Example 1:** Analysis of the system transfer functions reveals that the parameter AI provokes conflicting impact on inventory and unfilled order responses. Hence it should be investigated in greater depth in a simulation process.

**Example 2:** Small values of DI and DR will generally increase speed of responses. Simulations can be used for trade-off and sensitivity analysis.

**Small perturbation theory**

**• It is possible to find local stability boundaries** | **• Simulations can be undertaken only within the pre-determined stability boundaries** | **• Unnecessary simulations would be carried out**

**Example 1:** Analysis of the system transfer functions reveals that DR and DI should be positive control parameters in order to reach stability.

**Example 2:** Analysis shows that manufacturing constraints always provoke negative impact on system’s damping ratio. Simulation can be used to further investigate this effect.

**Describing functions**

**• Understanding the impact of different input frequencies on system’s behaviour** | **• Simulations can be undertaken for specific frequencies to confirm analytical insights.** | **• Several simulation experiments would be necessary to gain the same insights**

**Example 1:** Analysis shows that manufacturing constraints always provoke negative impact on system’s damping ratio. Simulation can be used to further investigate this effect.

**Example 2:** Analysis showed that shipment constraint decreases the system’s natural frequency but slightly improves the damping ratio. Simulation can be used to further investigate this effect.

**Example 3:** Analysis of the system transfer functions reveals that the parameter AI provokes conflicting impact on inventory and unfilled order responses. Hence it should be investigated in greater depth in a simulation process.

*Table 2: Table of insights and further analysis*
techniques supported the application of small perturbation theory and describing functions methods. Finally, these linearisation techniques provided further insights since they make possible the calculation of the system transfer functions and local stability boundaries and the understanding of how different capacity constraints impact on the system’s behaviour. While simulation generates confidence, analysis breeds insights.

The methodological framework proposed by this work can guide researchers and practitioners in undertaking comprehensive analysis and design of nonlinear dynamics systems in a less time consuming and less resource intensive manner. Hence, research with the aim of acquiring only exploratory knowledge of system dynamics may opt to conduct a simulation-only approach. In the search for deep understanding, many of the simulation studies reviewed in Table 1 could benefit from simplification methods to reduce the complexity of their models when the number of variables and parameters is too large. For supply chain dynamics studies considering more than one echelon, analysis can be firstly undertaken for a single echelon to gain preliminary insights on the existing underlying mechanisms. These simplification techniques will then facilitate the application of linearisation methods that will provide more insights into understanding the systems behaviour and how nonlinearities affect system responses. Further investigation and results confirmation can then be more easily undertaken via simulation and consequently leading to a more complete system design.

This research is limited to the dynamics of single-echelon supply chain systems and focused on the analysis of each nonlinearity individually. Further research, motivated by the analytical research and due consideration of the literature review, could include in-depth analysis of system stability and limit cycles via describing function and identification and investigation of other combination of nonlinearities, such as continuous and multi-valued requiring the Volterra-Wiener series expansion method.

References


Appendix A. Forrester Industrial Dynamics Model: Equations and Parameters

Equations used in the DYNAMO program (factory echelon)

\[
RR.KL = \begin{cases} 
  RRI, & \text{if } t \leq 0 \\
  RRI + \text{STEP}, & \text{if } t > 0 
\end{cases} \quad (A.1)
\]

\[
UO.K = UO.J + (DT)(RR.JK - SS.JK) \quad (A.2)
\]

\[
IA.K = IA.J + (DT)(SR.JK - SS.JK) \quad (A.3)
\]

\[
ST.K = UO.K/DF.K \quad (A.4)
\]

\[
NI.K = IA.K/DT \quad (A.5)
\]

\[
SS.KL = \text{CLIP}(ST.K, NI.K, NI.K, ST.K) \quad (A.6)
\]

\[
DF.K = (ID.K/IA.K)(DU) + DH \quad (A.7)
\]

\[
ID.K = (AI)(RS.K) \quad (A.8)
\]

\[
RS.K = RS.J + (DT)(1/DR)(RR.JK - RS.J) \quad (A.9)
\]

\[
MW.K = RR.KL + (1/DI)(ID.K - IA.K + LD.K - LA.K + UO.K - UN.K) \quad (A.10)
\]

\[
MD.KL = \text{CLIP}(MW.K, AL, AL, MW.K) \quad (A.11)
\]

\[
LD.K = (RS.K)(DC + DP) \quad (A.12)
\]

\[
LA.K = CP.K + OP.K \quad (A.13)
\]

\[
UN.K = (RS.K)(DH + DU) \quad (A.14)
\]

\[
CP.K = CP.J + (DT)(MD.JK - MO.JK) \quad (A.15)
\]

\[
MO.KL = \text{DELAY3}(MD.KL, DC) \quad (A.16)
\]

\[
OP.K = OP.J + (DT)(MO.JK - SR.JK) \quad (A.17)
\]

\[
SR.KL = \text{DELAY3}(MO.KL, DP) \quad (A.18)
\]

Variables used in the DYNAMO program (factory echelon)

<table>
<thead>
<tr>
<th>CP</th>
<th>clerical in-process orders</th>
<th>NI</th>
<th>negative inventory limit rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>DF</td>
<td>delay (variable) in filling orders</td>
<td>OP</td>
<td>orders in production</td>
</tr>
<tr>
<td>IA</td>
<td>inventory actual</td>
<td>RR</td>
<td>requisition (orders) received</td>
</tr>
<tr>
<td>ID</td>
<td>inventory desired</td>
<td>RS</td>
<td>requisition (orders) smoothed</td>
</tr>
<tr>
<td>LA</td>
<td>pipeline orders actual in transit</td>
<td>SR</td>
<td>shipment received inventory</td>
</tr>
<tr>
<td>LD</td>
<td>pipeline orders desired in transit</td>
<td>SS</td>
<td>shipment sent</td>
</tr>
<tr>
<td>MD</td>
<td>manufacturing rate decision</td>
<td>ST</td>
<td>shipping rate tried</td>
</tr>
<tr>
<td>MO</td>
<td>manufacturing orders</td>
<td>UN</td>
<td>unfilled orders normal</td>
</tr>
<tr>
<td>MW</td>
<td>manufacturing rate wanted</td>
<td>UN</td>
<td>unfilled orders</td>
</tr>
</tbody>
</table>

Constants used in the DYNAMO programme (factory echelon)
AI=4 constant for inventory
AL=1000(RRI) constant specifying capacity limit *
DC=1 delay clerical
DH=1 delay due to minimum handling time
DI=4 delay in inventory/pipeline adjustment
DP=6 delay in production lead time
DR=8 delay in smoothing requisitions
DU=1 delay, average, in unfilled orders
DT=1 solution time interval
RRI=1000 initial value of demand *
STEP=100 requisition step change *

* Different values of AL were considered when evaluating the impact of manufacturing constraints.
**Author used standard unit step input in order to compare simulation and mathematical results.

Initial conditions in the DYNAMO programme (factory echelon)

SS=RR IA=AI.RR
MD=RR CP=DC.RRI
RS=RRI OP=DP.RRI
UO=RRI (DH+DU)
Appendix B. Low order modelling

Let a high-order system be represented by a transfer function in the following form:

\[ T(s) = \frac{1 + b_1 s + b_2 s^2 + \cdots + b_q s^q}{1 + a_1 s + a_2 s^2 + \cdots + a_n s^n} \] (B.1)

The low order model will then be:

\[ T_M(s) = \frac{1 + B_1 s + B_2 s^2 + \cdots + b_Q s^Q}{1 + A_1 s + A_2 s^2 + \cdots + A_N s^N} \] (B.2)

so that \( Q \leq q \) and \( N \) must be less than \( n \).

a) Matsubara time delay theorem for low order modelling

This method initially involves choosing the poles nearest to the imaginary axis to determine \( T_M(s) \). However, the Matsubara time delay theorem is also incorporated to compensate for inaccuracies in the low order model. This gives us the following model:

\[ T_M(s) = e^{-\tau s} \left( \frac{1 + B_1 s + B_2 s^2 + \cdots + b_Q s^Q}{1 + A_1 s + A_2 s^2 + \cdots + A_N s^N} \right) \] (B.3)

where \( \tau \) is a time delay in the response which is determined by matching the system and model step responses according to the integral of error from time zero to infinity. In other words, the area \( D \), between the input and output lines in the system, \( T(s) \), should match the respective area in the low order model, \( T_M(s) \). This area can be found by calculating the integral of error from time zero to infinity as:

\[ D = \int_0^{\infty} [input(t) - output(t)] dt \] (B.4)

For a unit step input, the Laplace transform of the above equation we have

\[ D = \frac{1}{s} \left[ \frac{1}{s} - \frac{T(s)}{s} \right] \] (B.5)

where \( T(s) \) is the transfer function of the high order system in Equation B.1. Replacing this general form of transfer function results:

\[ D = \frac{1}{s^2} \left[ \frac{(a_1 - b_1)s + (a_2 - b_2)s^2 + \cdots + b_q s^q}{1 + a_1 s + a_2 s^2 + \cdots + a_n s^n} \right] \] (B.6)

Using the final value theorem, it is obtained that the area between the input and output lines, \( D \), in the system is simply equal to \((a_1 - b_1)\). Analogously, the corresponding area in the low order model will be \((A_1 - B_1)\), which is normally smaller than the area, \((a_1 - b_1)\), in the system. Hence, adding the time delay proposed by Matsubara, the relation \((a_1 - b_1) = (A_1 - B_1) + \tau\) is obtained.

In the Forrester pipeline we have:
\[ T(s) = \frac{SR}{MD} = \left( \frac{1}{1 + \frac{DC}{3}s} \right)^3 \cdot \left( \frac{1}{1 + \frac{DP}{3}s} \right)^3 = \]

\[ = 1 \left[ 1 + \frac{1}{1 + \left( DC + DP \right)s + \left( \frac{DC^2}{3} + DC \cdot DP + \frac{DP^2}{3} \right)s^2 + \left( \frac{DC^3}{27} + \frac{DC^2DP}{3} + \frac{DCDP^2}{3} + \frac{DP^3}{27} \right)s^3 + \right. \]

\[ + \left. \left( \frac{DC^3DP}{27} + \frac{DC^2DP^2}{9} + \frac{DCDP^3}{27} \right)s^4 + \left( \frac{DC^3DP^2}{81} + \frac{DC^2DP^3}{81} \right)s^5 + \left( \frac{DC^3DP^3}{729} \right)s^6 \right] \quad (B.7) \]

Hence, \( b_1 = 0 \) and \( a_1 = DC + DP \). In order to approach the system above to a first order model, we have to choose one pole. This is normally the pole with the least magnitude. Assuming that \( DP \) is always greater than \( DC \) we would choose then \( \frac{3}{DP} \) as our initial pole. Hence the low order model will be:

\[ T_{M'}(s) = \frac{1}{1 + \frac{DP}{3}s} \quad (B.8) \]

\( B_1 = 0 \) and \( A_1 = \frac{DP}{3} \). Since \( \tau = (a_1 - b_1) - (A_1 - B_1) \), we can then find that the time delay will be:

\( \tau = DC + \frac{2DP}{3} \). Hence, the system can be better approached by

\[ T_{M\tau}(s) = e^{-\left( DC + \frac{2}{3}DP \right)s} \left[ \frac{1}{1 + \frac{DP}{3}s} \right] \quad (B.9) \]

If for any reason it is necessary to avoid the time delay model \( T_{M\tau}(s) \) and a low-order model in the form \( T_M(s) \) is preferable, it is possible to adjust the model coefficients by placing a ‘dummy’ pole so that \( (A_1 - B_1) = (a_1 - b_1) \) (Towill, 1981). However, by placing this ‘dummy’ pole, a minimum of a second order model will be necessary. Hence, the first order low-order time delay model will become the following second low-order model:

\[ T_M(s) = \frac{1}{(1 + \frac{DP}{3}s) \cdot [1 + \left( DC + \frac{2}{3}DP \right)s]} \quad (B.10) \]

b) Hsia Method for low order modelling

The approximation method proposed by Hsia (1972) is based on selecting \( A_i \) and \( B_i \) (see Equation ??), in such a way that \( T_M(s) \) has a frequency response very close to that of \( T(s) \). In other words, the magnitude of the frequency function \( \frac{T_M(j\omega)}{T_M(j\omega)} \) is required to deviate the least amount from unity for various frequencies. Hence, the following relation should be satisfied:

\[ \frac{|T(i\omega)|^2}{|T_M(i\omega)|^2} = 1, \text{ for } 0 \leq \omega \leq \infty \quad (B.11) \]

The ratio \( T(s)/T_M(s) \) can also be written as:
\[
\frac{T(s)}{T_M(s)} = \frac{(1 + b_1 s + b_2 s^2 + \cdots + b_q s^q)(1 + A_1 s + A_2 s^2 + \cdots + A_N s^N)}{(1 + a_1 s + a_2 s^2 + \cdots + a_n s^n)(1 + B_1 s + B_2 s^2 + \cdots + b_Q s^Q)}
\]

\[
= \frac{(1 + m_1 s + m_2 s^2 + \cdots + m_u s^u)}{(1 + l_1 s + l_2 s^2 + \cdots + l_v s^v)}
\]

where \(u = q + N\) and \(v = n + Q\). Equation B.11 can be re-written as

\[
\frac{|T(i\omega)|^2}{|T_M(i\omega)|^2} = \left| \frac{T(s)T(-s)}{T_M(s)T_M(-s)} \right|_{s=i\omega}
\]

\[
= \frac{(1 + m_1 s + m_2 s^2 + \cdots + m_u s^u)(1 - m_1 s + m_2 s^2 + \cdots + (-1)^u m_u s^u)}{(1 + l_1 s + l_2 s^2 + \cdots + l_v s^v)(1 - l_1 s + l_2 s^2 + \cdots + (-1)^v l_v s^v)}
\]

This can be re-written in the form of

\[
\frac{|T(i\omega)|^2}{|T_M(i\omega)|^2} = 1 + \left( \frac{e_2 - f_2 s^2 + (e_4 - f_4)s^4 + \cdots + (e_{2u} - f_{2u})s^{2u}}{1 + f_2 s^2 + f_4 s^4 + \cdots + f_{2v} s^{2v}} \right)_{s=i\omega}, \text{ if } u=v
\]

Then, to satisfy the condition of Equation B.11, \(e_2 = f_2, e_4 = f_4, \cdots, e_{2u} = f_{2u}\) should hold true. However, if \(u < v\), which it is in most practical cases, then we have that this formula will imply an error. Given the conditions that \(e_2 = f_2, e_4 = f_4, \cdots, e_{2u} = f_{2u}\) and Equation B.13, the unknown coefficients for determining \(T_M(s)\) once \(T(s)\) is given can be calculated by solving the system of non-linear equations.

When applying this method to find a first order approximation of Forrester’s pipeline, the value of coefficient \(A_1\) of the following low-order model is to be found.

\[
T_M^{(1)}(s) = \frac{1}{1 + A_1 s}
\]

Then, the next step is to find the ratio \(\frac{T(s)}{T_M(s)}\):

\[
\frac{T(s)}{T_M(s)} = \frac{1 + m_1 s}{1 + l_1 s + l_2 s^2 + l_3 s^3 + l_4 s^4 + l_5 s^5 + l_6 s^6}
\]

where the coefficients \(l_i\) correspond to the coefficients \(a_i\) of the system \(T(s)\) of Equation B.7 and \(m_1\) is equal to the coefficient \(A_1\) of the low-order model in Equation B.15. Hence,

\[
m_1 = A_1
\]

\[
l_1 = DC + DP
\]

\[
l_2 = \frac{DC^2}{3} + DC.DP + \frac{DP^2}{3}
\]

36
\[ l_3 = \frac{DC^3}{27} + \frac{DC^2DP}{3} + \frac{DC.DP^2}{3} + \frac{DP^3}{27} \]  
\( \text{(B.20)} \)

\[ l_4 = \frac{DC^3DP}{27} + \frac{DC^2DP^2}{9} + \frac{DC.DP^3}{27} \]  
\( \text{(B.21)} \)

\[ l_5 = \frac{DC^3DP^2}{81} + \frac{DC^2DP^3}{81} \]  
\( \text{(B.22)} \)

\[ l_6 = \frac{DC^3DP^3}{729} \]  
\( \text{(B.23)} \)

The magnitude ratio between the system and the model will then be:

\[
\frac{|T(j\omega)|^2}{|T_M(j\omega)|^2} = \frac{T(s)T(-s)}{T_M(s)T_M(-s)} = \frac{1 + m_1s}{1 + l_1s + l_2s^2 + l_3s^3 + l_4s^4 + l_5s^5 + l_6s^6} \cdot \frac{1 - m_1s}{1 - l_1s + l_2s^2 - l_3s^3 + l_4s^4 - l_5s^5 + l_6s^6} = \frac{1 + e_2s^2}{1 + f_2s^2 + f_4s^4 + f_6s^6 + f_8s^8 + f_{10}s^{10} + f_{12}s^{12}}
\]  
\( \text{(B.24)} \)

where,

\[
\begin{cases}
e_2 = f_2 = -m_1^2 = -A_1^2 \\
f_2 = 2l_2 - l_1 = \frac{1}{3}(-DC^2 - DP^2)
\end{cases}
\]  
\( \text{(B.25)} \)

By replacing the second equation in the first one, the first order model can be determined as:

\[ T_M(s) = \frac{1}{1 + \frac{\sqrt{DC^2 + DP^2}}{\sqrt{3}}} \]  
\( \text{(B.26)} \)

Figure B.15 compares Matsubara and Hsia method for both step and sinusoidal inputs of Forrester pipeline.
Figure B.15: Comparing step and frequency responses of Hsia and Matsubara methods of low order modelling

Appendix C. Transfer functions obtained via Small Perturbation Theory

From the linearised model represented in the block diagram of Figure 5, it is possible to determine the system transfer functions in relation to the input demand or requisitions (RR). In order to simplify the equations, the author has decided to separate the control parameters from the parameters that the supply chain designer cannot select or control, such as physical parameters like delivery and manufacturing lead-times. Other system dynamics researchers have done the same when equations become large and therefore difficult to interpret (Towill, 1992; Wikner et al., 1992; Jeong et al., 2000). In the Forrester model, the delay in smoothing requisitions (DR), the delay in inventory/pipeline adjustment (DI) and the constant for
inventory (AI) are the control parameters and all the other parameters occur due to physical conditions.

By replacing the physical parameters with actual values given in Appendix A, the following transfer functions can be determined:

\[
\frac{MD}{RR} = 1 + \frac{(5 + AI)s}{(1 + DI_s)(1 + DR_s)} \tag{C.1}
\]

\[
\frac{SR}{RR} = \frac{1 + (5 + AI + DI + DR)s + DIDRs^2}{(1 + 2s)(1 + 5s)(1 + DI_s)(1 + DR_s)} \tag{C.2}
\]

\[
\frac{SS}{RR} = \left(1 + AI + (5 + 7AI + DI + AIDI + DR + AIDR)s + (3AI + 6AIDI + 7AIDI + DIDR + AIDIDR)s^2 + (-10AI + 3AIDI + 10AIDR + 7AIDIDR)s^3 + (-10AIDI + 10AIDIDR)s^4 \right) / \left((1 + 2s)(1 + 5s)(1 + AI + 2AI_s)(1 + DI_s)(1 + DR_s)\right) \tag{C.3}
\]

\[
\frac{UO}{RR} = \left(2 + 2AI + (10 + 7DI + 7DR + AI(21 + 3DI + 2DR))s + (10DI + 10DR + 7DIDR + AI(30 + 21DI + 14DR + 2DIDR))s^2 + (10DIDR + 2AI(15DI + 10DR + 7DIDR))s^3 + 20AIDIDRs^4 \right) / \left((1 + 2s)(1 + 5s)(1 + AI + 2AI_s)(1 + DI_s)(1 + DR_s)\right) \tag{C.4}
\]

\[
\frac{IA}{RR} = \frac{AI(1 + AI) + AI(5 - 4DI - 5DR)s - 5AIDI(DR - 1)s^2}{(1 + 5s)(1 + AI + 2AI_s)(1 + DI_s)(1 + DR_s)} \tag{C.5}
\]
Appendix D. Analyses via Describing Function

In Figure 7, the system transfer functions can be found by replacing the CLIP functions with their respective describing function gains. Let \( N_{A(MW)} \) and \( N_{A(ST)} \) be the gains caused by the discontinuities in the manufacturing and shipment processes respectively. Therefore the new system characteristic equation can be determined as:

\[
(1 + 2s)(1 + 5s)(1 + DRs)((1 + AI)N_{A(ST)} + 2AI)s(N_{A(MW)} + DI) = 0
\] (D.1)

Note that only two poles of the characteristic equation are affected by the discontinuous nonlinearities. Since many high-order systems can be represented by a series of second and first order transfer functions (Srivastava et al., 2009), the characteristic equation in Equation D.1 can be re-arranged as:

\[
(1 + 2s)(1 + 5s)(1 + DRs)\left(N_{A(MW)}N_{A(ST)} + AI.N_{A(MW)}N_{A(ST)} + (2AI.N_{A(MW)} + DI.N_{A(ST)} + AI.DI.N_{A(ST)})s + 2AI.DIs^2\right) = 0
\] (D.2)

In this way, it is possible to determine the damping ratio, \( \zeta \), and the natural frequency, \( \omega_n \) of the second order term in Equation D.2 as:

\[
\omega_n = \sqrt{\frac{(1 + AI)N_{A(MW)}N_{A(ST)}}{2AI.DI}}, \quad \zeta = \frac{2AI.N_{A(MW)} + (1 + AI)DI.N_{A(ST)}}{4AI.DI\sqrt{(1 + AI)N_{A(MW)}N_{A(ST)}}} \] (D.3)

By keeping the values of DI and AI fixed (\( DI = 4 \) and \( AI = 4 \), as given by Forrester), Table D.3 illustrates the values of \( \omega_n \) and \( \zeta \) as \( A_{MW} \) and \( A_{ST} \), increase. Figures 8 and 11 can be referred in order to check on the values of the describing functions for a given input amplitude. Table D.3 shows us that the value of \( \omega_n \) decreases as both \( A_{MW} \) and \( A_{ST} \) increase. Regarding \( \zeta \), the system is slightly overdamped, \( \zeta = 1.107 \), when linear \( (N_{A(MW)} = 1 \) and \( N_{A(ST)} = 1) \). As the CLIP function becomes active and the gain in the manufacturing constraint, \( N_{A(MW)} \), decreases the system becomes even more overdamped. On the other hand, as the input amplitude in the shipment constraint \( A_{ST} \) increases, the system in practice becomes critically damped with \( \zeta = 1.006 \).

Table D.3: Effect of the CLIP functions on system’s natural frequency and damping ratio

<table>
<thead>
<tr>
<th>( A_{ST} )</th>
<th>( A_{MW} )</th>
<th>( N_{A(MW)} )</th>
<th>( N_{A(ST)} )</th>
<th>( \omega_n )</th>
<th>( \zeta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \leq AL )</td>
<td>1</td>
<td>0.395</td>
<td>1.107</td>
<td>0.375</td>
<td>0.280</td>
</tr>
<tr>
<td>1.45AL</td>
<td>0.9</td>
<td>0.375</td>
<td>1.133</td>
<td>0.356</td>
<td>0.265</td>
</tr>
<tr>
<td>2.03AL</td>
<td>0.8</td>
<td>0.354</td>
<td>1.167</td>
<td>0.335</td>
<td>0.250</td>
</tr>
<tr>
<td>3.12AL</td>
<td>0.7</td>
<td>0.331</td>
<td>1.209</td>
<td>0.314</td>
<td>0.234</td>
</tr>
<tr>
<td>6.29AL</td>
<td>0.6</td>
<td>0.306</td>
<td>1.266</td>
<td>0.290</td>
<td>0.217</td>
</tr>
<tr>
<td>( \infty )</td>
<td>0.5</td>
<td>0.280</td>
<td>1.342</td>
<td>0.265</td>
<td>0.198</td>
</tr>
</tbody>
</table>

* Based on input frequency of 0.1Hz

** Imaginary parts cause little impact on the poles position and will be disregarded for further calculations