Rogue Seasonality Detection in Supply Chains

Vinaya Shukla, Mohamed M. Naim, Nina F. Thornhill

PII: S0925-5273(12)00134-X
DOI: doi:10.1016/j.ijpe.2012.03.026
Reference: PROECO5075

To appear in: *Int. J. Production Economics*

Received date: 27 July 2009
Accepted date: 23 March 2012

Cite this article as: Vinaya Shukla, Mohamed M. Naim and Nina F. Thornhill, Rogue Seasonality Detection in Supply Chains, *Int. J. Production Economics,* doi:10.1016/j.ijpe.2012.03.026

Title: Rogue Seasonality Detection in Supply Chains

Corresponding and First Author: Dr Vinaya Shukla
Department of Business and Management
Middlesex University Business School
The Burroughs, London NW4 4BT
Tel: +44(0)20 8411 4247
Fax: +44(0)20 8411 6011
Email: v.shukla@mdx.ac.uk

Other Authors:

Professor Mohamed M Naim
Cardiff Business School, UK
E-mail: naimmm@cardiff.ac.uk
Tel: +44 (0) 29 2087 4635/6083
Fax: +44 (0) 29 2087 4301

Professor Nina F Thornhill
Centre for Process Systems Engineering
Department of Chemical Engineering
Imperial College
London
E-mail: n.thornhill@imperial.ac.uk

Rogue Seasonality Detection in Supply Chains

Abstract

Rogue seasonality or unintended cyclic variability in order and other supply chain variables is an endogenous disturbance generated by a company's internal processes such as inventory and production control systems. The ability to automatically detect, diagnose and discriminate rogue seasonality from exogenous disturbances is of prime importance to decision makers. This paper compares the effectiveness of alternative time series techniques based on Fourier and discrete wavelet transforms, autocorrelation and cross correlation functions and autoregressive model in detecting rogue seasonality. Rogue seasonalities of various intensities were generated using different simulation designs and demand patterns to evaluate each of these techniques. An index for rogue seasonality, based on the clustering profile of the supply chain variables was defined and used in the evaluation. The Fourier transform technique was found to be the most effective for rogue seasonality detection, which was also subsequently validated using data from a steel supply network.

Keywords: Supply chain management; Rogue seasonality; Data mining; Time series; Simulation

1. Introduction

Supply chains routinely face disturbances in provision of efficient value delivery to customers. These could originate from changing demand patterns, problems at the supplier, from the company's own value added processes or from control practices that use information to link entities in the chain (Davis, 1993). Amongst these, control practices, specifically those related to inventory and production control, have been the subject of intense research as their inappropriate use leads to generation of endogenous disturbances called the 'bullwhip' effect (Lee et al., 1997a). Much research in this area has focused on trying to frame better rules to avoid or minimize this effect (Dejonckheere et al., 2003; Disney et al., 2004). The attempts to frame rules lead to the important question; for which value streams should these rules be framed or modified? Rapid identification of such value streams from among thousands in a typical supply chain system, and then focussing on them, could significantly improve operational effectiveness. However, this practical aspect of identification has been ignored by most previous studies on bullwhip. For example, Geary et al. (2006) discuss in detail about different approaches to mitigate the bullwhip effect such as OR theory, simulation and control theory but provide limited insight about how to identify or detect this problem. The importance of such detection based approaches was highlighted by Christopher and Lee (2004) recently, who suggested the need for intelligent alerts for out of control situations. The present study is focused on the detection of rogue seasonality, which is associated with demand amplification (Forrester, 1961) or 'bullwhip' effect and could similarly indicate inefficiency in supply chain operations. Rogue seasonality is observed in many real world supply chains (Lee et al., 1997b; Carranza and Villegas, 2006) as well as in role playing simulations (Sterman, 1989) and involves generation of cyclic patterns in order and other supply chain variables of a frequency different from exogenous demand. It has adverse implications for the supply chain partners in terms of increase in production on-costs and/or increase in stock holding costs, with these additional costs not recoverable from the customer as rogue cyclic variations are not customer driven. Metters (1997) showed that costs to the extent of 10-20% could be reduced by elimination of such seasonal variations. Moreover, rogue seasonality could propagate though the supply chain and cause such inefficiencies in multiple echelons. Its detection, specifically, discriminating supply chains with higher rogue seasonality intensity from others, is therefore important, so that management could focus on such cases and effect timely mitigative action, where feasible and necessary. Such an approach is especially useful in a complex environment, where many multiple echelon supply chains are being managed by a focal company.

Studies looking specifically at rogue seasonality include Kim and Springer (2008), who used an analytical system dynamics based approach to study it, where their focus was on identifying conditions causing its generation. A relevant detection oriented study, based on discriminating endogenous from exogenous seasonal disturbances was by Thornhill and Naim (2006), who demonstrated the effectiveness of the spectral principal component analysis technique (SPCA) using the context of a steel supply network. However, their exploratory study was limited to one real world case. Their findings therefore, need to be validated and extended through the use of multiple rogue seasonality contexts, alternative techniques for detection, and use of a more automated process. The objective of this study is to provide this broader context and validation. We use a simulated model of the supply chain to generate datasets with different rogue seasonality intensities and characteristics by changing the supply chain designs, demand patterns and lead times. Simulation is used as it enables identification of cause and effect relationships between the input disturbance and rogue seasonality effects generated by the system. Alternative time series features such as autoregressive modeling coefficients (AR model), autocorrelation function (ACF), cross correlation function (CCF), amplitudes of Fourier transform (FT) and wavelet coefficients from discrete wavelet transform (DWT) were applied on these datasets for rogue seasonality detection. These techniques and features have been applied in different domains extensively (Liao, 2005) and have certain advantages compared to the spectra (FT) based technique used by Thornhill and Naim (2006). For example, ACF provides a finer grained periodicity detection (Vlachos et al., 2006) while DWT (Li et al., 2002), with its time and frequency localization property, can more efficiently and accurately analyse non-stationary signals i.e. signals which have time varying frequency characteristics. Similarly, the AR model based approach enables higher dimensionality reduction and also clustering of time series with different lengths (Ting et al., 2003).

The supply chain variables represented in terms of different features are clustered to identify patterns associated with rogue seasonality. To facilitate automation and for greater objectivity in assessment, an index which captures this characteristic cluster pattern is proposed. Finally, the effectiveness of different transformation techniques and features is assessed using empirical data from the steel industry. Hence, to summarise, the specific contribution of this study are:

a) Determining a characteristic signature of rogue seasonality useful for detecting its presence.
b) Introducing a new measure for rogue seasonality intensity called the "rogue seasonality index" and demonstrating its accuracy and robustness as a measure.
c) Evaluating different time series transformation techniques and features and identifying the most appropriate one for rogue seasonality detection. Studies of this kind, which compare different time series techniques and features are quite limited across domains (Keogh and Kasetty, 2003) including supply chain management.

The rest of the paper is structured as follows. We review relevant studies on rogue seasonality and time series methods for detection in the next section. Simulation design is discussed in Section 3 while the findings from the simulation are considered in Section 4. In Section 5 we discuss the case study and its analysis. Section 6 explains the potential applicability of rogue seasonality detection in a management control system and we conclude in Section 7.

2. RELEVANT STUDIES ON ROGUE SEASONALITY AND TIME SERIES METHODS FOR DETECTION

First demonstrated by Forrester (1961), referred to as such by Towill (1997) and alternatively called endogenous cyclicality or volatility (Kim and Springer, 2008), rogue seasonality is a disturbance characterised by a cyclic pattern in order and other supply chain variables, and which is generated endogenously from the inventory and production control system used i.e. it is not present in exogenous demand. While Thornhill and Naim (2006, pp. 148) specified the periodicity of the cyclic pattern associated with rogue seasonality to be of a few months (possibly because they observed such a behaviour in the steel sector context which they used), this study assumes no such periodicity restriction which could be of the order of days,weeks or months depending on the context. This generic characterisation is more realistic, given that different sectors have different cyclic characteristics: for example, sectors with faster operating dynamics such as fast moving consumer goods (FMCG) and high technology have periodicities in the variable profiles of the order of weeks and days (Fok et al., 2007; Wu et al., 2010; Lee et al., 1997a,b; Torres and Moran, 2006), which was also highlighted by Thornhill and Naim (2006, pp. 160), although they did not consider this in their rogue seasonality definition. Relatively stable sectors such as steel on the other hand exhibit periodicities of the order of months as seen in Thornhill and Naim's study. Hence, a definition of rogue seasonality, which is not associated with a particular periodicity is more generic and is therefore used.

Few studies in the past have exclusively focused on rogue seasonality. It has mostly been studied in conjunction with the bullwhip effect (Forrester,1961; Dejonckheere et al., 2003; Jaksic and Rusjan, 2008). An exception to this is the recent study by Kim and Springer (2008), who used an analytical system dynamics based approach to determine the conditions under which rogue seasonality could arise in a supply chain. These studies are aimed at reducing bullwhip and rogue seasonality through better supply chain design, including design of forecasting and ordering policies. Though no doubt an important objective, such an approach on its own may be less successful in many environments which are dynamic, decentralized and dominated by players who do not behave rationally (Bendoly et al., 2006). It needs to be supplemented with approaches that detect and manage or sense and respond to unplanned or abnormal occurrences (Haeckel and Nolan, 1993). Similar Supply Chain Event Management (SCEM) systems are being used to detect discrepancies in transaction ordering and order fulfillment processes (Otto, 2003) while a sense and respond based approach to logistics is being implemented by the military (Tripp et al., 2006). Use of this approach in the context of rogue seasonality was explored by Thornhill and Naim (2006) using monthly time series data of a steel supply network and the spectra principal component analysis (SPCA) technique. Though this study was able to identify and discriminate rogue from exogenous seasonality, it had limitations such as non-consideration of weekly/daily manifestations of rogue seasonality and non-evaluation of other techniques of detection. These limitations were highlighted in that study itself and suggested as areas for further research. Moreover, the nature of rogue seasonality identification/detection was binary i.e. assessing whether it was present or absent. In reality, rogue seasonality is present in all supply chains, but with different intensities (Kim and Springer, 2008) and therefore, an indication of its intensity rather than presence/absence is more relevant. In this study we extend the findings of Thornhill and Naim (2006) using different rogue seasonality contexts which involve use of daily data, alternative techniques of detection and a more automated process of detection based on a measure of rogue seasonality intensity. The last objective enables detection and assessment of rogue seasonality in a multiple supply chain analysis context, so that appropriate and timely management action could be taken against those with the highest intensities, as these have the maximum negative impact on supply chain operations (Metters, 1997).

Abnormal occurrences or disturbances are detected using either change point detection methods which rely on a change in the probability distribution/distribution parameters of univariate/multivariate data (Basseville and Nikiforov, 1993), or signature based methods which compare data profiles with known disturbance signatures/profiles (Han and Kamber, 2006). Signature based methods are gaining in popularity in various domains (Cortes and Pregibon, 2001; Sy, 2005) in view of their flexibility and greater availability of tools/techniques for application. A significant potential therefore exists of applying these methods in the supply chain management domain. This study proposes to do so using the context of rogue seasonality. In signature based methods, signatures are usually derived on the basis of clustering or finding similarity/dissimilarity relationships among the time series profiles of the operating variables. The following three approaches are generally used for time series clustering (Liao, 2005): a) raw data based (time series data used for clustering), b) feature based (features extracted from the time series are used for clustering) and, c) model based (time series is converted into a model and the model parameters are used for clustering). Clustering approaches based on time series data, though easier to interpret, have problems in computation (large, high dimensional data sets) as well as lower accuracy because of interference from noise. Feature based and model based approaches are therefore preferred (Liao, 2005; Keogh and Kasetty, 2003). Some of the commonly used features include amplitudes of Fourier transform (FT) (Agrawal et al., 1993; Caiado et al., 2006; Thornhill and Naim, 2006), coefficients from an autoregressive (AR) model (Kalpakis et al., 2001; Maharaj, 2000; Ting et al., 2003), wavelet coefficients of discrete wavelet transform (DWT) (Chan and Fu, 1999; Zhang et al., 2005), autocorrelation function (ACF) (Wang and Wang, 2000; Vlachos et al., 2006) and cross correlation function (CCF) which is a feature of pairs of time series (Bohte et al., 1980; Baragona, 2001).

A key aspect of most of the studies referenced above, as well as other such studies is their data bias (Keogh and Kasetty, 2003). This refers to the use of primarily synthetic, non contextual data for demonstrating the effectiveness of different techniques/features in clustering. Empirical data has been used in only a few studies as a supplement, but this data has also been mostly on stock price/finance domain. Since the results and effectiveness of different features and similarity measures is domain specific, and none of these studies has used any supply chain related data, their applicability in the supply chain context is limited. Secondly, few studies have compared the performance of their methods with other rival methods, and those that have, have used only one or two methods in the comparison (Keogh and Kasetty, 2003; Liao, 2005). This study, by using both supply chain contextual data (simulation and empirical) as well as multiple time series transformation methods attempts to fill these gaps. Results from this study would not only enlighten us about the superior method for rogue seasonality detection, but would also give pointers about the effectiveness of different transformation methods and features for clustering supply chain data.

3. SIMULATION DESIGN

Simulation design was used as it provides controlled experimentation so that rogue seasonality of different intensities and characteristics could be generated in the system variables. The simulated data could subsequently be transformed using different transformation techniques to extract features, which could be individually assessed in terms of their effectiveness in rogue seasonality detection.

3.1 Time series Data Generation

Time series data was generated using the automatic pipeline, inventory and order based production control system (APIOBPCS) (John et al., 1994), which is a versatile model that has been applied by decision-makers in the supply chain (Sterman, 1989), in industrial practice (Towill et al., 1997) and can be adapted to represent order-up-to as well as lean and agile scheduling policies (Dejonckheere et al., 2003). The ordering policy in the APIOBPCS may be described by; "the order placed is equal to the average sales rate plus a fraction ($1/T_i$) of the inventory error plus a fraction ($1/T_w$) of the work-in-process (WIP) error" where T_i is termed the "time to adjust inventory" and T_w is termed the "time to adjust WIP". The average sales rate is calculated using exponential smoothing and is dependent on a parameter T_a related to the exponential smoothing parameter α. While T_p is a physical parameter, T_i , T_w and T_a are decision parameters that are set according to performance criteria such as the minimisation of order variance, inventory availability and the speed of response to changes in demand. The equations defining the behaviour of different variables in APIOBPCS as per John et al. (1994) are given in Figure 1.

Take in Figure 1

The variables considered in the model are consumption/sales rate (CONS), forecast demand/sales rate (FORDMD), order rate (ORATE), work in process level (WIP), desired work in process level (DWIP), completion rate or receipts into inventory (COMRATE), actual inventory level (AINV), error between desired and actual inventory level (EINV) and error between desired and actual work in progress level (EWIP). Both, single echelon and three echelon cases are considered.

The dynamics of these system variables is dependent on the parameters T_i , T_w and T_a , whose 'optimum' values have been suggested by some authors from the perspective that these will dampen the amplification or bullwhip effect, while ensuring retention of customer service levels through sufficient on hand inventory. Examples of suggested 'optima' are:

- John et al. (1994): T_a=2T_p, T_i= T_p and T_w =2T_p. This design represents a rational rule based on `hard' engineering systems.
- Sterman (1989): T_a=0.5T_p, T_i=T_p and T_w=T_p. This design was based on minimizing total inventory costs which compensates the WIP.
- Shukla et al. (2009): T_a=2T_p, T_i=T_p and T_w=0.67T_p. This is a variant of the first design and is characterized by an increased WIP feedback by decreasing T_w.

Other authors have also explored the combinatorial parameter settings of multi-echelon supply chains predominantly from a perspective that attempts to filter out random or rogue seasonality while ensuring that the real demand signal is processed (Towill and del Vecchio, 1994). For example, Mason-Jones et al (1997) suggest an 'optimum' design for a four echelon model is to ensure that at each echelon T_a=2T_p, T_i= T_p and T_i= T_p.

A two stage process was followed in the simulation. In the first stage, a single echelon was analysed using standard exogenous disturbances for CONS (sales rate) such as step, pulse and random (Gaussian) which is typical of such studies (John, et al., 1994) as it facilitates easier understanding of the dynamics. Make to order (MTO) and make to stock (MTS) strategies were considered, as these are more commonly found in industry than a mixture of MTO and MTS systems (Buxey, 1995). Also, as these structures incorporate different degrees of CONS or sales information in their orders, variation was expected in their rogue seasonality generation. The parameters chosen were based on Naim et al. (2007) as MTO**:** T_a = 0, T_i = T_w = ∞; MTS: T_a = T_w = ∞, T_i = T_p. These parameter settings, far away from suggested 'optima', represent two extreme cases of the APIOBCPS order policy which will lead to two distinct system dynamics. In the MTS case we expect significant rogue seasonality and bullwhip, while in the MTO situation, we expect no rogue seasonality although inventory errors will be high. Hence, we have clear discrete benchmarks by which to test our rogue seasonality detection techniques. It is also important to clarify that a linear time invariant system, similar to that used by Towill and del Vecchio (1994), Mason-Jones et al. (1997) and Naim et al. (2007) is considered; backlogs, a non-linearity is avoided by keeping the desired inventory at an appropriately high level.

The methods described later in Sections 3.2 and 3.3 led to understanding of how rogue seasonality could be associated with a signature. The insights motivated a second stage of simulation in which detailed simulation was done as per Table 1.

Take in Table 1

Here it is important to clarify that simulation is not used to replicate the analysis in Thornhill and Naim (2006) but instead to complement it. While Thornhill and Naim used monthly data, and which was highlighted as a weakness in their study, here daily data is being simulated as such faster operating dynamics is observed in many sectors such as FMCG and high technology (Fok et al., 2007; Wu et al., 2010; Thornhill and Naim, 2006, pp. 160). One year data consisting of 250 daily data points (total days in a year less 104 data points pertaining to weekends, because as per the experience of one of the authors these are generally not recorded in practice, less another 10 data points corresponding to holidays and shutdowns) were generated for all variables for each simulated case. However, in order to assess if number of data points had any impact, two year data (500 daily data points) was also simulated and analysed for select cases.

Delay or lead time between ordering and receipt of goods which, in this case is only the production time, was varied as it plays a critical role in supply chain dynamics (Forrester, 1961; Chen et al., 2000). Lead times of half week, one week and two weeks are considered. Both the average delay as well as the order of delay, the latter reflecting the distribution of the output around the average delay, are important and hence were independently varied. Demand for many goods follow autoregressive (AR) and moving average (MA) processes of different orders (Chopra and Meindl, 2001). AR and MA processes of orders 1 and 2 besides random (Gaussian) were therefore considered. Within each process, parameter values such as ρ for AR (1) and θ for MA (1) were varied as these are known to affect the dynamics of the operating variables (Gilbert, 2005). Finally, exogenous seasonality of different periodicities was also added to the demand pattern to make the analysis more realistic. Given that daily data is simulated, realistic periodicity options for exogenous seasonality are weekly (frequency = 0.2) as per Fok et al. (2007) and monthly (frequency = 0.05), both of which were separately added to the stochastic demand processes and analysed. One hundred independent replications were used for each demand process and parameter combination using common random numbers. The data generation was done using Excel© and Matlab©.

3.2 Time series transformation techniques and features used in rogue seasonality detection

Alternative time series transformation techniques and features were evaluated to determine which of them gave the most discriminating signature for rogue seasonality. The signature referred to here, is derived from the clustering profile of variables in the transformed or feature space. All the commonly used techniques and features as given in Section 2 were applied and evaluated. Untransformed data i.e. data in the time domain was also clustered to serve as a baseline for comparison. Matlab© was used for all transformations, after normalizing the data to eliminate mean and amplitude scale differences. Details of the transformations and features used are given in Table 2.

Take in Table 2

All the transformations are aimed at extracting the key features of the time series and then using these for determining the similarity/dissimilarity relationships and subsequent clustering of this data. Fourier transform (FT) uses amplitudes of constituent frequencies in the time series as features, on the basis of equivalence in similarity relationships between the time and frequency domains (Agrawal et al., 1993). The features for the AR model method are the coefficients of a fitted model, in which the current values are a function of past values (Box and Jenkins, 1976). For ACF, correlation of observations across various time intervals or lags are the basis for comparison, while in CCF, a similar correlation between pairs of time series is used. Two measures of similarity based on CCF have been used in the literature. One is a composite measure derived from all the cross correlation values at different lags (Bohte et al., 1980) while the other one uses the maximum correlation value within the specified maximum lags (Baragona, 2001). We used the latter because of its intuitive simplicity. Wavelet coefficients, which essentially represent the correlation between the original and the local oscillation represented by the wavelet basis function (Percival and Walden, 2000) are used for clustering in the DWT based method. The basis functions are generated by translation and dilation of the mother wavelet. Haar wavelet was used as the mother wavelet in this study in view of its extensive use in time series data mining studies. Details of these transformations can be found in Chatfield (2004), Percival and Walden (2000), Orfanidis (2002) and Last et al. (2004).

The number of features extracted for each transformation was varied to assess the sensitivity of the results to the same. For FT, this corresponds to the number of frequency channels used. Options tried were first 7, first 28, all and number corresponding to 80 percent energy in the signal, referred to as FT 7, FT 28, FT Total and FT 80 % Energy respectively. Different numbers of features were similarly extracted for the other transformations as shown in Table 2. Parameter free approaches such as AIC (Akaike, 1981) for AR model and Shannon entropy function (Shannon and Weaver, 1964) for DWT were also considered. The alternative numbers of features used are referred to as parameters of that feature, while together they are referred to as a feature parameter combination such as FT 7 or FT 28.

3.3 Clustering and rogue seasonality index

The profiles of different variables, either in the time or transformed domain, are required to be clustered for identifying patterns/relationships between them. A characteristic profile of clusters found to be associated with rogue seasonality could then become the latter's signature. We used agglomerative hierarchical clustering in this study in view of its great visualisation power and the fact that it does not require a priori specification of number of clusters (Han and Kamber, 2006). Clusters in this method are graphically displayed as a hierarchical tree called a dendogram, which gives the order in which the cluster-sub cluster relationships are formed and merged. Clustering requires three decision choices: a) defining the similarity measure, b) deciding the basis for merging two clusters and, c) deciding where to cut the dendogram to specify the number of clusters. The similarity measure used for all except the CCF method was the Euclidean distance, as it is commonly used in time series clustering (Agrawal et. al., 1993; Liao, 2005) and gives results of comparable accuracy vis-à-vis other similarity measures (Keogh and Kasetty, 2002). Euclidean distance assesses the similarity between two profiles or sequences by making point to point comparisons. Maximum absolute cross correlation (between specified maximum lags) is used as the similarity measure for CCF (Baragona, 2001). For merging two clusters, the Ward's algorithm, again a popular method was used (Halkidi et. al., 2001). As this algorithms works only with the Euclidean distance, for the CCF method, we used the complete linkage method instead. Finally, the decision about where to cut the dendogram and identify clusters depends on its tree structure and pattern of branches. For dendograms, in which branches are separated and join far apart, distinct clusters of the variables attached to the branches can be identified. However, for others, identifying clusters and hence pattern of relationships between the variables is subjective and difficult. This is further complicated by issues such as the dissimilarity scale used in cluster visualisation. Hence, an alternative approach based on direct use of similarity/dissimilarity values between the variables was considered for rogue seasonality detection.

This alternative approach involves defining an index based on the similarity/dissimilarity values between the supply chain variables. The index could be used as a measure of rogue seasonality intensity and works on the hypothesis that rogue seasonality generation and propagation induces significant variation between the profiles of variables. With rogue seasonality, the other variables excluding CONS (or the exogenous demand pattern) such as AINV and WIP get a cyclical profile and thus become quite similar to each other, while becoming dissimilar to CONS. Therefore, the dissimilarity between CONS and the other variables is compared with the dissimilarity among the other variables themselves. If the ratio of these two quantities is large, it indicates rogue seasonality of greater intensity for the reason given above. This hypothesis was tested and found to be true in our initial single echelon simulation analysis with standard test inputs for CONS. It was therefore subsequently used in a detailed simulation analysis as per Tables 1 and 2. The index called the Rogue Seasonality index is expressed as:

$$\text{Rogue Seasonality Index} = \left(\frac{\text{Minimum dissimilarity between CONS and the other variables}}{\text{Average dissimilarity between all variables except CONS}}\right) \quad (1)$$

We also compared the above with alternative definitions such as

$$\begin{array}{l}\text{Rogue Seasonality Index} \\ \text{(Alternate 1)}\end{array} = \left(\frac{\text{Average dissimilarity between CONS and the other variables}}{\text{Average dissimilarity between all variables except CONS}}\right) \quad (2)$$

$$\begin{array}{l}\text{Rogue Seasonality Index} \\ \text{(Alternate 2)}\end{array} = \left(\frac{\begin{array}{c}\text{Minimum dissimilarity between CONS and the other variables} \\ \text{- Average dissimilarity between all variables except CONS}\end{array}}{\text{Standard deviation of dissimilarity between all variables except CONS}}\right) \quad (3)$$

Dissimilarity was measured in terms of the Euclidean distance between the variables (time or feature domain) for all techniques except CCF, where it is defined as *1 – max correlation* between two time series within maximum specified lags. Rogue seasonality index is alternatively referred to as index in the subsequent sections.

4. Simulation output analysis

The simulation analysis has been split into three sub sections to reflect the order in which the results were obtained and assimilated. Section 4.1 highlights rogue seasonality generation in a single echelon, its signature based on the clustering profiles of variables and the effectiveness of the index in capturing this signature. Section 4.2 extends this to a three echelon structure, with the final sub section focussing on the numerical analysis of the index, based on features from different techniques, to identify the most appropriate feature parameter combination for rogue seasonality detection.

4.1 SINGLE ECHELON ANALYSIS

A single echelon supply chain was used as the starting point as it provides easier understanding of the variable profiles and their relation to system structure, which could be used to determine the distinguishing signature of rogue seasonality. Figure 2 depicts the profiles of the system variables for MTO (make-to-order) / MTS (make-to-stock) and first order/infinite order production delay combinations excited with a Gaussian CONS (sales rate).

Take in Figure 2

Though this analysis was also done for other CONS profiles such as step and pulse, and which yielded similar results, their plots are not depicted here due to space constraints. For the same reason, outputs from different features are also selectively depicted. Time domain and FT feature based outputs are presented in this section and those based on other features such as ACF, CCF, DWT and AR model in the subsequent section. The clusters are depicted using dendograms whose vertical axis is a measure of dissimilarity. Any variables close together at the end of a long branch have very similar profiles, while others that are dissimilar are on different branches.

The MTO supply chain is shown in the first two rows of Figure 2. The time trends of CONS, FORDMD, ORATE, COMRATE, AINV, EINV, WIP, DWIP and EWIP (terms defined earlier in Section 3.1) are shown in the left hand panel and the spectra calculated from the Fourier transform (FT) in the right hand panel. An inspection of these plots for both the first order as well as infinite order lags shows a close similarity between the CONS, ORATE, FORDMD and DWIP profiles, in both the time as well as spectral representations. Based on the causal structure of APIOBPCS, such a nature of profiles and clustering is expected. With no demand smoothing and no pipeline and inventory feedback, CONS, FORDMD and ORATE are equivalent. Their profile is similar to DWIP as the latter is proportional to FORDMD and normalised data is used. The behaviour of COMRATE however, changes with the nature of lags. For a first order lag, production, or delivery, pipeline acts as a low pass filter which results in COMRATE having a different profile than CONS and ORATE and hence being clustered separately. For the pure delay representation however, we do find that COMRATE is clustered with CONS and ORATE. This is because in a pure delay COMRATE is equal to ORATE shifted by production lead time T_p , so that spectra, which is invariant to phase shifts in the time domain and looks at frequencies (or profiles) only, recognises them as similar. The plots also highlight the disadvantages of time domain analysis. For example, AINV and EINV, which are inverted versions of each other (as DINV is constant) and thus similar, are actually clustered separately in the time domain. This is however, accurately represented in the frequency domain where they are clustered together. From the perspective of rogue seasonality, since the CONS profile is transmitted to the other variables without distortion, i.e. other variables do not have frequencies different from those in CONS, there is no rogue seasonality generation. This is accurately reflected in the 0 index value in both the time as well as the spectra domain (FT Total).

The MTS supply chain results are shown in the bottom two rows of Figure 2. All the variables except CONS are seen to oscillate at a similar frequency of around 0.02 (50 days cycle, with the simulation being in days) in both MTS systems, which is especially clear in case of the delay order infinity system. This again arises from the system structure and control policies used. ORATE in a MTS system depends upon FORDMD and EINV. FORDMD is based on a long term forecast and is not updated regularly based on CONS while EINV is reduced by a factor of T_i (time to adjust inventory), when used to calculate ORATE. Hence, ORATE is not directly influenced by CONS but only indirectly through a smoothed EINV. This is what causes dissimilarity between ORATE and CONS which gets passed on to dissimilarity between other variables and CONS also. In terms of order of delay, variables for order of delay infinity show more consistent cyclicality as compared to order of delay one, which is also reflected in their spectral plots as higher peaks. This is because for the same average delay, the effective delay in an infinite order delay system is much greater than for delay order 1. For example, in a system with infinite order delay and average delay T_p, no output exits the system till time T_p has elapsed from the time of input. In contrast, in a delay order 1 system, partial outputs start exiting the system immediately after inputs are put into the system. The higher the effective delay in the system, the higher is the amplification. This effect has been studied by many researchers (Forrester, 1961; Chen et al., 2000; Gilbert, 2005) and can also be seen through use of analytical expressions by Disney et al. (2006). It is important to highlight here that the single echelon structure is only used to demonstrate and explain the origin of rogue seasonality. This objective is effectively met in the above analysis, despite the cyclicality in the variable profiles not being absolutely consistent, as evident in the plots for MTS delay order 1. Detailed analysis and investigations are however all based on the three echelon structure, where the variable profiles have consistent, akin to the real world, cycles (discussed in the next section).

The cluster profiles in Figure 2 clearly demonstrate the nature of clustering of CONS as a discriminating factor between presence/absence of rogue seasonality. Comparison of rows 1 and 2 with rows 3 and 4 in Figure 2 shows that, while CONS is clustered with the other variables in the MTO supply chain system where no rogue seasonality is present, it is clustered separately from the other variables in the MTS system associated with rogue seasonality. The other variables in the MTS system are all seen to be oscillating at a frequency of around 0.02 or 50 days cycle and are therefore clustered together, which is especially clear in the spectra representation for delay order infinity system. The signature for rogue seasonality can therefore be simply expressed in terms of whether internal system variables cluster or do not cluster with CONS. When the latter is true, rogue seasonality is considered present, otherwise not. Also, the extent of CONS dissimilarity indicates the strength of rogue seasonality. This is evident when we compare MTS delay order 1 and delay order infinity clustering profiles in rows 3 and 4 of Figure 2. Our definition of rogue seasonality index in Section 3.3, which is anchored around dissimilarity of CONS with the other variables, is therefore validated. Looking specifically at the index values we see that it captures the presence as well as intensity of rogue seasonality. In the time domain the index values for MTO (order 1 and infinity), MTS order 1 and MTS order infinity systems are 0, 0.88 and 0.91 respectively which accurately indicate absence, low intensity and marginally higher intensity of rogue seasonality. The corresponding values for spectra (FT Total) are 0, 3.98 and 7.04, which demonstrates its greater discriminatory power in rogue seasonality detection. However, this finding is based on analysis of a single instance and needs to be validated through multiple replications.

Finally we assessed, whether use of different number of time periods of data had any impact on the findings. While in Figure 2, the analysis uses 250 time periods of data (1 year daily data), the same analysis (clustering and index computation) was done using 500 time periods of data. No significant difference was observed between the cluster plots for the two cases for any of MTO/MTS, delay order one/infinity systems. The index values (based on spectra or FT Total) were also not found to be significantly different: while the index values based on 500 time period data for MTO delay order one, MTO delay order infinity, MTS delay order 1 and MTS delay order infinity were found to be 0, 0, 3.74 and 6.78 respectively, the same for 250 time period data were 0, 0. 3.98 and 7.04. The similar nature of cluster profiles and the insignificant difference in index values between the 250 and 500 time period data cases is not surprising. This is because the signature and index computation both involve relative profile comparisons (of variables); they are indifferent to the length of the time series as long as these are able to capture the key profile characteristics (cyclicality in this case). The subsequent analysis is therefore all based on 250 time period data only.

4.2 THREE ECHELON SUPPLY CHAIN ANALYSIS

The logic of rogue seasonality signature and index having been established, this now needed to be tested on realistic multiple echelon systems. Three echelon systems, created by coupling together of the corresponding single echelon systems were therefore simulated, with the same ordering parameters (T_a , T_i and T_w) and delay (T_p) used for each echelon. Coupling was done by making ORATE of the downstream echelon equivalent to the CONS of the upstream echelon. Such an approach to analyse the dynamic behaviour of multiple echelon systems has been used in Mason-Jones et al (1997) and Shukla et al (2009).

The behaviour of variables in the three echelon system was seen to be similar to that for the single echelon system. In case of the MTO system, the ORATE variable of each echelon is clustered together with the exogenous consumer demand (CONS1), highlighting that the frequency characteristics of the demand are maintained across echelons and no rogue seasonality is generated. On the other hand for the MTS supply chain, the rogue seasonality generated in echelon 1 is transmitted to upstream echelons 2 and 3, as seen in the nature of clustering of CONS1 relative to other variables in figure 3.

Take in Figure 3

In contrast to the single echelon system, cycles in the profiles of variables for the corresponding three echelon system are very consistent. This is clearly evident in the ACF plot where all the variables except CONS1 have very similar cycles (half cycle of around 27 days ~ frequency of 0.02) and are clustered together. This consistency arises from the Forrester effect (Forrester, 1961) where the cyclicality generated in one echelon gets amplified when transmitted to upstream echelons. Cyclicality is also visible in the plot of AR model coefficients for all the operating variables except CONS1. CONS1 being clustered separately from the other variables for the ACF, AR model and CCF features highlights the effectiveness of these techniques in rogue seasonality detection, as per the discussion in the previous section. However, this is not evident for the DWT feature. Also, the plots in Figure 3, especially for ACF, AR model and CCF clearly support the rationality of the rogue seasonality index. The operating variables excluding CONS1 show cyclicality, and therefore have similar profiles and low dissimilarity with each other resulting in a small value for the denominator in the index. CONS1 does not show cyclical behaviour and is therefore quite dissimilar from the other variables so that the numerator of the index is a large number. Hence, when rogue seasonality is present and its intensity is high, the index should be a large number which is seen for all the three features (index values for ACF 28, AR 28 and CCF 28 are 4.2, 2.0 and 10.4 respectively). The magnitude of the index however, varies with the features used, with DWT features (index value of 0.6) being the weakest in terms of differentiating presence/absence of rogue seasonality. However, valid conclusions cannot be drawn from one instance of a stochastic process, and therefore, a detailed simulation experiment was conducted.

4.3 ANALYSIS OF ROGUE SEASONALITY INDEX OUTPUT FROM SIMULATION

Detailed simulation as per Tables 1 and 2 was carried out to increase the validity of the findings about the rogue seasonality index and the features to be used. Only the MTS system was simulated as it alone showed rogue seasonality characteristics. 250 time periods of data, corresponding to one year of daily data, was simulated in each case. The output from the simulation was generated and organised as follows. For example, for an MTS system with average delay 3 and order of delay 1, excited with 1 instance of AR (1) demand process with ρ 0.1, rogue seasonality index was calculated based on each of raw time, FT Total, FT 7, FT 28, FT 80 % Energy, AR 7, AR 28, AR AIC, ACF 7, ACF 28, CCF 7, CCF 28, DWT (Level 5) and DWT (Level Shannon) feature parameter combinations. This was replicated 100 times to calculate the average as well as coefficient of variation (Standard Deviation/Average) values for each combination. These outputs were also produced for the two alternative rogue seasonality indexes mentioned in Section 3.3. We found high correlation (0.93-0.99) between the three indices for most of the feature parameter combinations, with marginally lower correlation (0.80-0.84) for only a couple of combinations. This was not unexpected, as all the three measures are closely related. As the three indices were essentially measuring the same thing, analysis considering alternative indices was considered unnecessary and therefore not done.

4.3.1 KEY OBSERVATIONS

The index values for all feature parameter combinations were seen to be greater than zero. A zero value is associated with a system such as MTO that does not exhibit any rogue seasonality. Since only MTS systems were used in the simulation, and all such systems with non zero lead times exhibit rogue seasonality, this signifies that all combinations could potentially be used for rogue seasonality detection. However, the index values for these combinations show a significant variance, with some combinations yielding relatively low index values.

Index value from the time domain was seen to be less than 1 for all demand-delay combinations. The magnitude of change in the average index value, from changes in demand parameter/order of delay in this domain was also seen to be small. A specific case can be seen in Figure 2 where the index values based on the time domain for MTS delay order 1 and delay order infinity systems are 0.88 and 0.91. This is not surprising as data in the time domain is unable to identify similarities in leading/lagged profiles, which makes the average dissimilarity among the other variables (excl. CONS1) much larger than that between CONS1 and these variables. Use of DWT coefficients was also seen to give a similar low index value.

The index values based on AR model features had reasonable magnitudes, though the proportion of values less than 1 was higher in comparison to other techniques. The highest index values were observed for features based on the FT, ACF and CCF techniques, thereby indicating their greater discriminating power in assessing the presence/absence of rogue seasonality. However, the index values varied with the parameters used with the feature. For example, in case of the FT feature, significant volatility in the average index values was seen for the FT 7 and FT 28 parameter options compared to FT Total and FT 80% Energy. Overall, the wide range of responses obtained from simulation made it difficult to determine the best feature parameter combination for rogue seasonality detection. A systematic and objective basis to collate and compare these responses is therefore required.

4.3.2 CRITERIA FOR COMPARISON OF DIFFERENT FEATURE PARAMETER COMBINATIONS

We defined three criteria to compare the effectiveness of different feature parameter combinations for index determination, which in turn translates into better rogue seasonality detection. These are:

a) **Consistency** – A superior feature parameter combination is that which generates index values, consistent with change in demand parameters and order of delays. By consistency, it is implied that demand parameters or order of delays associated with actual increase/decrease in rogue seasonality intensity, would have a similar impact (increase/decrease) on index values. For example, for the AR (1) demand process, as ρ, the autoregressive parameter changes from -0.8 to -0.5 to 0.1 the rogue seasonality intensity increases. This is because spectral energy at lower frequencies is greater for ρ of 0.1 than for -0.8 (Chatfield, 2004) and these frequencies correspond to those which get amplified by the MTS system used in the simulation. A consistent feature parameter combination from a demand perspective would therefore be that, which gives increasing index values as the demand parameter ρ changes from -0.8 to -0.5 to 0.1. The parameters for each of the other demand processes such as AR (2), MA (1) and MA (2) given in row 5 of Table 1 were also similarly chosen so as to give contrasting (high/low) spectral energies at low frequencies. In terms of consistency based on order of delay, a system with an infinite order of delay produces rogue seasonality of greater intensity than that which has an order of delay 1 as discussed earlier. A consistent feature parameter combination from an order of delay perspective would therefore be that, which gives index values that vary similarly with order of delay; a lower index value for order 1 and a higher index value for order infinity.

b) **Magnitude of index (Discrimination)** – A superior feature parameter combination is that, which yields index values of greater magnitude than other feature parameter combinations for the same system. This increases the discrimination between presence/absence of rogue seasonality as well as between rogue seasonalities of different intensities. A ranking system was used, where for each demand parameter, delay and order of delay combination, the feature parameter combinations were ranked based on their average index values. The sum of ranks for all the parameter, delay and order of delay combinations was used as the basis for classification; the lower the sum of ranks, the better was that feature parameter combination.

c) **Coefficient of variation of index (Stochastic robustness)** – A superior feature parameter combination is that, which yields index values which have a smaller coefficient of variation, and which therefore, is more robust to stochastic variations than other combinations. A ranking system similar to magnitude of index was used, with the only difference being that, a higher rank was given to the feature parameter combination with a lower value of coefficient of variation.

4.3.3 Comparison of feature parameter combinations based on defined criteria

The different feature parameter combinations were analysed based on the above criteria and its summary is given in Tables 5, 6, 7. In this analysis exogenous seasonality was not present in the demand process. The same has been considered in the next section (Section 4.3.4). Gradual build up of complexity in this way improves understanding, and was therefore used.

Take in Tables 5, 6, 7

Table 5 gives the consistency of the index with respect to demand parameters and order of delays for different feature parameter combinations. The shaded cells indicate the maximum values for that row. The maximum possible consistency values used in the tables are derived as follows. For example, for AR (1), the maximum consistency for demand parameters is 12, because use of three demand parameters (ρ = -0.8, -0.5 and 0.1) means 2 comparisons for each of 3 delays (3, 7, 14) and 2 order of delays (order 1 and infinity). Similarly, its maximum consistency for order of delays is 9 as there are two orders of delays implying 1 comparison for each of the 3 demand parameters and 3 delays. These have been similarly arrived at for the other demand processes. Looking at the table, we find that the FT and ACF based features yield the most consistent index values from both demand parameter as well as order of delay perspectives. This is primarily on account of their ability to ignore leads/lags between the data profiles. FT and ACF are well established for analysing univariate data with seasonality, but this study shows that these are also effective for analysing multiple data series contexts having varying seasonal characteristics.

The features however differ in terms of their parameter sensitivity. While ACF features show a marginal change in consistency with change in parameters from 7 to 28, the change is significant for the FT features. FT 7 and FT 28 with first seven and twenty eight frequencies show poor consistency, which improves when more frequencies are used such as in the FFT 80 % energy option, with the highest consistency being for the FT Total option where all frequencies are used in the index calculation. The behaviour of FT 7 and FT 28 is on account of non inclusion of relevant rogue frequencies within the first 7 and 28 frequencies used in these options. As regards other features such as CCF and the time domain, these were found to yield slightly lower consistencies than ACF and FT, which is understandable, as both are based on direct comparisons of time series profiles, though CCF has some advantages as it looks at lags/leads in relationships also. The DWT and the AR model features, however, showed the least consistency. For DWT, this is because of the inability of the discontinuous Haar wavelet to capture the smooth variable profiles. The resulting low feature values have an adverse impact on consistency. However, for the AR model the result is surprising, as this approach has been shown to be effective by Ting et al. (2003) and Caiado (2006). This could be due to the larger number of variables used here as well as the significant non-stationarity induced in the data from rogue seasonality.

The sum of ranks based on average index magnitudes and coefficient of variation are given in Tables 6 and 7 respectively. As discussed earlier, the best transformation technique and feature is that which yields index values with the largest magnitudes and which have the lowest coefficient of variation. However, many feature parameter combinations are seen to be not similarly good or bad on both these criteria. The sum of ranks for both criteria have therefore been added together and presented at the bottom of Table 7. FT is seen to be balanced in meeting both criteria for some parameters such as FT Total and FT 80% energy. This is evident in their total sum of ranks being the lowest, though not having the lowest sum of ranks in either criteria. However, other FT parameters such as FT 7 and FT 28 appear to be less adequate. CCF meets both criteria effectively, while ACF, which showed good consistency, was found to be deficient in the coefficient of variation criteria. DWT and AR model based techniques, which were behind other features on the consistency criteria, were similarly placed on these criteria as well. Finally, time based index are seen to be a less valid measure of rogue seasonality intensity as well as a less effective discriminator of rogue seasonalities of different intensities. As discussed earlier, this is because of lack of alignment of the cyclic profiles of variables in the time domain. Lower validity is evident in Table 5 which shows the consistency (measure of validity) of the time domain index to be relatively low for most demand process and order of delay combinations. The lower discrimination effectiveness is evident in Table 6 with the sum of ranks of the time based index values being relatively higher (indicating a lower index value) for most combinations. Index based on the time domain is therefore less appropriate in comparison to other techniques despite having advantages of ease in computation and interpretation.

Based on all the three criteria, we found FT, specifically the FT Total feature parameter combination to be the best for rogue seasonality detection. Using fewer numbers of parameters as in FT 7 and FT 28, though increasing efficiency in computation, was seen to reduce the accuracy and effectiveness of detection significantly. Intermediate approaches such as FT 80 % Energy therefore seem a reasonable compromise. Index values based on ACF features showed good consistency to varying demand and delay characteristics as well as good discrimination ability in view of the large index magnitudes. However, coefficient of variation of index values based on ACF was greater compared to other features, indicating lesser robustness to stochastic variability. With regard to CCF, index values calculated from it were not very consistent, though it performed adequately on the other two criteria. Finally, AR model and DWT features were found to be the worst in each of the three criteria. In fact, the performance on the criteria for some AR model and DWT parameters was even worse than that of the time domain.

Finally, we further assessed the effectiveness of FT Total (in relation to ACF and time) for a demand process with greater autocorrelation: ρ values of 0.5 and 0.7 were considered for the AR (1) demand process as these are occasionally observed in practice. The supply chain system considered was the same as in the previous analysis - three echelon MTS supply chain with delays of 3, 7, 14 and order of delays 1 and infinity. Given that higher ρ values (from 0.1 through 0.5 to 0.7) and delay order infinity (in comparison to delay order 1) should mean greater rogue seasonality intensity and therefore higher index values (refer section 4.3.2), it was assessed whether this was true or not. The analysis revealed the index values based on FT Total and ACF 28 to have a similar high consistency as in the earlier analysis (19/21), with the time domain showing a lower consistency at 13/21 (refer to the results for AR(1) demand process in Table 5 for comparison). To highlight a specific example, the index values based on FT Total for the delay 7, delay order one system were found to be 3.0, 3.2 and 3.4 for ρ values of 0.1, 0.5 and 0.7 respectively; the corresponding values for the delay order infinity system were 3.3. 3.4 and 3.7 respectively. Although FT Total and ACF 28 were similar in consistency terms, index values based on the latter showed lower stochastic robustness as in the earlier analysis. Therefore, from the overall analysis, FT, and especially FT Total, is found to be the best method for index computation and therefore rogue seasonality detection.

4.3.4 Exogenous seasonality impact on rogue seasonality index

Seasonality in demand is observed in many products and therefore, an analysis involving stochastic demand with exogenous seasonality was considered. The objective was to assess the impact of such a demand process on the profiles of system variables and if the logic of rogue seasonality index as defined in Section 3.3 remained valid. Since the simulation involved daily data (as specified in Section 3.1), exogenous seasonality of a week (frequency = 0.2) as per Fok et al. (2007) and month (frequency = 0.05) were considered. This meant that sinusoidal cyclicality of unit amplitude and frequencies 0.2 and 0.05 were added to the stochastic demand processes given in Table 1 and analysed.

Analysis of the frequency plots of variables for MTS systems with delay order one/infinity, delays of 3/7/14 and demand process (stochastic with cyclicality of frequency 0.2/0.05) revealed transmission of the exogenous seasonality to other variables at the same frequency. However, the nature of transmission was different: in systems with delays of 7 and 14, the upstream echelon variables showed attenuation at both 0.2 and 0.05 frequencies, while in the system with delay 3, these variables showed amplification for frequency 0.05 and attenuation for frequency 0.2. This behaviour is consistent with the previous theoretical work on such systems by Towill and del Vecchio (1994) and Dejonckheere et al. (2003). Frequencies in exogenous demand (CONS1) within a certain range, which is defined by the system structure are amplified, while those beyond that range are attenuated. In terms of rogue seasonality, though its generation was evident in all cases, it was significantly masked for cases with amplification. Thus, with the variable profiles showing variation based on the frequency of exogenous seasonality and the characteristic amplification-attenuation frequency range of the system, rogue seasonality index as defined in Section 3.3 may not be an accurate measure of rogue seasonality intensity for such cases. This would apply to many real world cases as demand profiles for most products show seasonality and system structure is generally unknown.

One way to ensure the continued validity of the index is to filter out the exogenous seasonality from all variables before calculating the rogue seasonality index. This is possible since exogenous seasonality is transmitted upstream to many variables, and can therefore by identified on that basis. Therefore, for a test system, if many variables across multi echelons have high signal energy at a particular frequency, this could be manifesting from exogenous seasonality. Using the notch filter, a well known filtering technique in signal processing (Orfanidis,2002), this common frequency could be filtered out from all the variables, where after the index could be calculated in the usual way based on any feature parameter combination. To test the effectiveness of this approach, simulated data (which includes exogenous frequencies of 0.2 or 0.05) as per Table 1 was generated, and index values based on different feature parameter combinations computed, post notch filtering.

The index values for contexts with exogenous seasonality, and which had been notch filtered, were compared with the same with no exogenous seasonality. The difference in values between the two cases was seen to be small. For example, the index value based on FT Total after notch filtering, for an MTS three echelon system with delay 3 and delay order 1, excited with AR (1), ρ -0.8 and exogenous frequency 0.05 demand process, was 1.90. The same for exogenous frequency 0.2 and no exogenous frequency was 1.80 and 1.71 respectively. A similar small difference was seen with other ρ values in the AR (1) demand process: for ρ -0.5, the index values were 2.37, 2.25 and 2.34 for exogenous seasonality 0.05, 0.2 and no exogenous seasonality respectively, while the same for ρ 0.1 were 2.97, 2.90 and 2.77. Notch filtering therefore, seems to ensure the integrity of the index as a measure for contexts with exogenous seasonality. However, its effectiveness needs to be more comprehensively assessed. For this, an approach similar to that discussed in Section 4.3.3 is used: index values are collated in the form of tables for consistency, discrimination ability and stochastic robustness (similar to Tables 5, 6, 7) for each of 0.05 and 0.2 exogenous frequency analyses separately. These tables are not depicted here due to space constraints and only their salient aspects are discussed.

Comparison of the consistency tables for exogenous frequencies 0.05, 0.2, and no exogenous seasonality, show the difference in their table values to be minimal. This is not surprising, given that the difference in index values between these three cases is itself small, as seen in the previous paragraph. The highest consistencies in case of both exogenous frequency 0.05 and 0.2 analyses are seen for FT Total, ACF7 and ACF28 at 64/69, 63/69, 65/69 and 68/69, 69/69, 69/69 respectively, which are similar to those for no exogenous seasonality (refer to Table 5). In terms of discrimination ability and stochastic robustness also, the relative position of each feature parameter combination is the same in all three cases. For example, ACF7 is the best and FT Total, the sixth best in terms of discrimination ability, although their total ranks (refer to Table 6 for context) in each case are slightly different: total ranks for ACF7 are 103, 90 and 93 for exogenous frequency 0.05, 0.2 and no exogenous seasonality respectively, while the same for FT Total are 347, 361 and 354. Similarly, with regards to stochastic robustness, time domain is seen as the best and FT Total the second best in each of the three cases. Overall, incl. both the discrimination ability and stochastic robustness criteria, FT Total is seen to outperform the other feature parameter combinations in case of both exogenous frequency 0.05 and 0.2, as was seen previously for no exogenous seasonality (refer Table 7). From this comprehensive analysis, it can therefore be inferred that for contexts with exogenous seasonality, FT Total based index after notch filtering, is an effective measure for rogue seasonality intensity given its high consistency, discrimination ability and stochastic robustness in measurement.

5. Case Study Analysis

In order to strengthen the validity of our approach, we derived the rogue seasonality index using real data and the same set of transformation techniques and features as used in simulation. The data is the same as that used in the steel case study by Thornhill and Naim (2006) and pertains to a supply network consisting of four autonomous business units; steel works, section mill, bar mill and rod & wire mill. The steel works supplies the raw material (billets) to each of the three mills to form a diverging network. It was identified in the study that the section and bar mills operate on a make to stock (MTS) basis and generate rogue frequencies in some of their operating variables. These were therefore used for the analysis, with each of the two dyads, steel works-section mill and steel works-bar mill, considered separately. Variables used in the analysis consist of all available operating variables for entities in the dyad. The variables used for the Steel Works-Section Mill dyad are: Section mill (Total orders, Production, Total despatches, Receipts from FG, Total billet stocks, Total stocks and Order book) and Steel works (FG billets total stocks, FG billets despatches to section mill, FG billets production). The Steel works-Bar mill dyad variables are similar, except that the variables for the Section mill are replaced with those for the Bar mill. Six years of monthly time series data (72 data points) were available for each variable and used in the analysis. Figure 4 shows the relevant supply chain structure (a single figure is used to represent both the dyads as their structures are similar). Since the data used here is the same as in Thornhill and Naim (2006), it is important to first discuss the difference between the two analyses.

Take in Figure 4

- - - - - - - - - - - - - - - -

Thornhill and Naim assessed exogenous/endogenous seasonalities and rogue seasonality generation using the following process: transforming the system variable to their spectra representation to eliminate lags/leads in relationships, using principal component analysis (PCA) to reduce each system variable to 3 dimensions, plotting these variables, manually identifying clusters and cluster characteristics, mapping clusters onto network diagram and looking for causal links between variables to identify exogenous and endogenous cyclical characteristics and finally compiling this information into a related look up table. Multiple steps are involved in this process, many of which require manual intervention, which makes its application in a multiple supply chain analysis context, difficult. Also, the spectra of variables need to be reducible to 3 dimensions during PCA, to enable their plotting and subsequent cluster profile assessment. Though possible for the steel data, this may not work in a generic sense. This study differs from Thornhill and Naim (2006) in terms of the process used: the signature and index based rogue seasonality detection process being proposed here is simpler, generic, requires minimal assumptions, and can be applied in an automated fashion for analysing multiple supply chains. The second difference is in terms of rogue seasonality assessment. Thornhill and Naim's look up table based process gives only a binary assessment (presence/absence) of rogue seasonality. However, rogue seasonality is present in all supply chains though with differing intensities (Kim and Springer, 2008) and therefore, getting an indication of its intensity rather than presence alone would be more useful, as it would indicate the extent of supply chain inefficiency (Metters, 1997). Such an indicator called the rogue seasonality index, which was seen to be effective with simulated data is being tested here on the Steel works-section mill and Steel works-bar mill contexts. Finally, Thornhill and Naim compared the effectiveness of their Spectra PCA (or SPCA) technique with only the time domain analysis. Here the spectra technique (same as FT used here) is being compared with other commonly used techniques such as Autocorrelation function (ACF), Autoregressive (AR) model, Cross correlation function (CCF), and Discrete wavelet transform (DWT) using each of the Steel works-section mill and Steel works-bar mill contexts. The sequence of steps followed in the analysis are as follows:

Data Pre-processing: This is a standard step in data mining (Han and Kamber, 2006) and involves preparing the data as per the requirements of the analysis. In order to do this automatically, appropriate rules need to be specified. The steps followed were:

- De-trending and normalising the data, the former to avoid its appearance in the zero frequency channel which could interfere with the analysis, and the latter, to eliminate scale differences between variables. Long term trends, which too appear in the low frequency range in the spectra, also need to be filtered. The important thing here is the selection of a threshold or cut-off frequency, which filters the long term trends while retaining all rogue frequencies. This selection has to be automatic, unlike in Thornhill and Naim (2006), who visually inspected the spectra plots to decide the same. One simple and generic option could be to use 0.08 (time period of 12 months) as the threshold frequency. This is because rogue seasonality in real world contexts is generally of the order of days, weeks or months and therefore, any frequency smaller than 0.08 could safely be classified as a long term trend. This was followed here and all frequencies lower than 0.08 were filtered out from all the variables for each dyad. Though the threshold of 12 months time period could be further refined to possibly time periods of 10/9/8 months by analysing the energy distribution (square of the amplitude post FT) across frequencies and variables, this increases the risk of filtering out rogue frequencies, and was therefore not done. The threshold frequency of 0.08 could be used in general also, although for cases where previous and expert knowledge about operating dynamics is available, the threshold could be more specific.

- In the second step, the exogenous seasonality in demand, and which is shared with the other variables, needs to be filtered out (refer Section 4.3.4). This was done as follows: First, FT was applied to the Total Orders variable (demand in this context) to compute the amplitudes at different frequencies; Next, the frequencies with the highest amplitudes (indicative of exogenous frequencies) were identified using a cut off of (Average + 2*Standard deviation) applied to the amplitude values; Finally, these frequencies were filtered out from the Total Orders variable as well as all other system variables. This cut off definition classifies 5% of the highest amplitude frequencies as exogenous. 5% rather than a tighter 1% or 2% was used as the cut-off to ensure that no exogenous frequency is missed. The frequencies identified as exogenous and filtered out automatically in this way from all the variables were: (being provided for the sake of completeness though this information is not required/used in the analysis):

Steel Works-section mill: Frequencies of 0.25 (4 month cycle) and 0.16 (6 month cycle)
Steel Works-bar mill: Frequencies of 0.25 (4 month cycle) and 0.40 (2 and half month cycle)

Data Analysis: After data pre-processing, each of the FT, ACF, CCF, AR model and DWT based features were applied to the data, with time domain used as the reference. The variables in terms of each feature representation were clustered and the resulting cluster profile assessed in relation to the rogue seasonality signature. Next, the rogue seasonality index was computed based on each feature and the relative effectiveness of the features assessed. Because of limited data, only one parameter per feature could be analysed. The feature parameter combinations used are FT Total, 36 lags for ACF, max lags 36 for CCF, order 36 for AR and level 2 transformation (18 coefficients) for DWT. Finally, indices based on alternative definitions (refer Section 3.3) were also calculated for each of these for comparison purposes. The cluster profiles based on FT amplitudes (FT Total) and index values based on different features for each dyad are given in Figure 4. Also, in Figure 4 are the spectra plots for the two dyads, but unlike in Thornhill and Naim (2006), these do not form a part of rogue seasonality assessment here.

Findings: The review of cluster profiles and index values in Figure 4 reveal the following:

- The cluster profiles show exogenous demand variables (9 and 23) to be clustered separately from the other system variables for both dyads, which is a characteristic signature for rogue seasonality (refer Section 4.1 page 13). Hence, rogue seasonality is successfully detected in the dyads, and this was possible without having to make any of the assumptions/manual intervention required in Thornhill and Naim (2006).
- Though presence of rogue seasonality is detected, what is more important to know is its intensity as per the earlier discussion. Looking at the index values, these are greater than zero for both dyads for all the features considered. Index values greater than zero indicate presence of rogue seasonality (refer Section 4.3.1) and therefore, rogue seasonality is successfully detected by the index based approach as well. However, the magnitude of the indices is low, especially in comparison to the simulation output. This indicates the rogue seasonality intensity for both dyads to be low. This finding is reasonable because unlike in simulation, where a pure make to stock or MTS system was used (upstream orders are based only on bridging the gap between desired and actual stock), the dyads use a hybrid MTO-MTS where actual customer demand is also an input into the ordering decision. Customer demand contributing to upstream order decision ensures lesser over/under ordering, and hence, lesser intensity of rogue seasonality in comparison to a system, where it is not at all an input (MTS). Use of a hybrid MTO-MTS rather than a pure MTS control system by the dyads is borne out by the order book variables (orders accepted by a company but not yet delivered to the customer -variables 29 and 30). It will be useful at this stage to also look at the spectra plots in Figure 4 to get a sense of the rogue seasonality frequencies and intensity (this may be required only for select supply chains as discussed in the last paragraph of this section). The plots indicate presence of rogue seasonality, as spectral peaks at certain frequencies which are shared across variables (frequency of 0.33 for the Steel works-Section mill dyad and 0.17 & 0.33 for the Steel works-Bar mill dyad). Rogue seasonality intensity on the other hand is reflected in the height of the spectral peak (higher the spectral peak, greater is the intensity, as it is indicative of all the spectral energy being concentrated in those frequencies) and the extent

to which these peaks are shared by variables in the supply chain (more the number of variables sharing the spectral peak, more is the intensity). Examining the spectra plots, it is evident that the intensity is low on account of both these factors for each dyad. This shows that assessment of the rogue seasonality intensity from index is possible even for real world contexts.

- The next objective was to find out the best feature for computing the rogue seasonality index. For the steel works-section mill dyad, the FT Total based index is seen to have the highest value, with those based on AR model and DWT the least and ACF and CCF based index values in the middle and marginally better than the time domain. This is in line with the findings from the simulation analysis. A higher index value based on a particular feature more accurately reflects the true rogue seasonality characteristics. This is apparent when we compare the simulation output summaries of FT and ACF with Time, AR Model and DWT, specifically Consistency (Table 5) with Rank analysis (Table 6), where a high correlation between the two criteria is evident. Higher index values also facilitate better discrimination between rogue seasonalities of different intensities. Next, examining the index values for the steel works-bar mill dyad, these appear to be marginally different from those for the steel works-section mill dyad. Though FT, ACF and CCF are seen to be similarly highly placed in terms of index values, and DWT and time domain near the bottom, what is surprising is the AR model yielding the highest index value. However, this was also observed for select cases in the simulation analysis, where the AR model based approach yielded index values that were greater than those from other features. This means that for some unique conditions, it is possible for otherwise inappropriate features to outperform others. However, on a generic basis, features whose robustness has been assessed with multiple data contexts will still be the most appropriate, and therefore, from this analysis also, FT Total turns out to be the best feature for rogue seasonality index computation.

In order to assess the robustness of the rogue seasonality index as a measure, alternative index definitions (refer to Section 3.3) were used to compute the index for each feature. The correlation between the index used and the alternative indices was found to be high, as in the case of simulation (0.83 and 0.90 for the steel works-section mill dyad and 0.81 and 0.93 for the steel works-bar mill dyad). This establishes the index definition to be robust and not very sensitive to the nature of dissimilarity measurement between exogenous demand and other variables.

Overall, the index based approach was able to establish in an automatic fashion, the presence of low intensity rogue seasonality in the two dyads. Given a context where many supply chains are being managed at the same time, such an approach can be quite useful: the controller can use the index to quickly discriminate supply chains with high rogue seasonality intensity from others, and focus their investigations on these. Investigation may involve plotting the spectra of system variables and uncovering the frequency/s of rogue seasonality and the reason/s for their generation. For example, in the case of the two dyads, their index values are not high, and this can be used to assign them a lower priority in terms of investigation. With Thornhill and Naim's (2006) process however, every supply chain irrespective of their rogue seasonality intensity are treated in the same way and investigated in detail, causing both its efficiency and effectiveness to be lower.

6. Managerial application of rogue seasonality index

The aim of the index is to give the managers a compact indication of the intensity of rogue seasonality in a

multiple supply/value chain environment so that they can then concentrate their efforts on those with the highest values or which show maximum deviation from historic trends. While Thornhill and Naim (2006) gave a conceptual framework for such an application, this paper strengthens it by providing a quantitative basis and making it more generic. Figure 5 gives the application context of the index. Examples of such centralized monitoring of supply chains by focal companies, as envisaged here for rogue seasonality, include Cisco's eHub initiative (Grosvenor and Austin, 2001) which uses real time intelligence on inventory, commitments, capacity and other operating information of thousands of suppliers in three tiers for effective supply chain execution. Similarly, Eaton Corporation (Supply & Demand Chain Executive, February 2006) and Honda (Ward's Autoworld, July 2006) operate supply chain wide systems to monitor supplier data for thousands of suppliers for advance notice/warning of potential problems.

Take in Figure 5

Successful application of the proposed index based approach requires large volumes of timely and accurate data with different sampling intervals and from different parts of the supply chain. For accurate assessment, the sampling frequency should be at least twice the highest frequency contained in the data variables because of the Nyquist sampling theorem (Chatfield, 2004). This means that, for contexts having seasonality of the order of days (2 days or more), such as in the high technology and FMCG sectors, daily sampled data will be needed. Other contexts, having seasonality of the order of months (2 months or more), would require at least monthly sampled data, which means that monthly/weekly/daily could all be used for assessment. The advent of RFID coupled with information and communication technologies makes this less of a technical challenge nowadays. Data analysis and reporting can also be rapidly performed. However, it is possible that data on only a few variables is provided by certain players due to trust and confidentiality issues. The proposed approach is robust and can still be applied as long as those variables show rogue seasonality characteristics. For example, use of only orders and stock data for the Steel works-Section mill and Steel works-Bar mill dyads yielded index values of 1.22 and 1.06 (FT Total), still accurately indicating presence of rogue seasonality. The second issue pertains to specifying appropriate thresholds for classification of rogue seasonality intensities for management action. We propose the following:

To start with, a common index threshold could be specified for all newly added supply/value chain which is kept low enough to classify relatively more supply chains as having high rogue seasonality and investigated. The investigation would involve plotting the spectra profiles of the variables in the supply chain, assessing the rogue seasonality frequencies, the causal factors responsible for the generation of the same and the extent to which those could be controlled (to minimise rogue seasonality intensity). For example, a low level of trust among the focal company and its partners could make it difficult to modify the ordering heuristics. Another stream of investigation could involve assessing the cost implications of rogue seasonality. For example, a supply chain with a high index value, could have surplus capacity and/or high flexibility, meaning less adverse cost impact from rogue seasonality and therefore limited need for management intervention. Another possibility is where rogue seasonality and its associated costs are already incorporated into the decision structure and optimised with respect to other ordering considerations as in Jaksic and Rusjan (2008). Once sufficient history of index values and related management interventions is available, specific maximum thresholds for each supply/value chain could be prescribed for regular monitoring and detection of significant deviations. The index is therefore proposed to be used not only for detecting rogue seasonality but also for assessing changes in its characteristics over time, akin to statistical process control (SPC) in a manufacturing environment. Use of multiple variables in the index's computation reduces its variability and therefore makes it easier to specify robust thresholds. Once an instance of high rogue seasonality is detected, and which has not been optimised within the ordering decision structure, actions such as those proposed by Kim and Springer (2008) could be effected. This would mean either decreasing the supply lead times (T_p) and/or increase the time to adjust inventory (T_i) and/or reduce the time to adjust WIP (T_w) all of which reduce the propensity of strong cyclicality.

7. DISCUSSION AND CONCLUSIONS

Rogue seasonality, characterised by endogenous generation of cyclicality in multiple supply chain variables, reduces the cost efficiency of supply chains because of unnecessary ramping up and down of production and/or excess stocking in a capacitated environment. It is observed in many real world supply chains, though with differing intensities, and arises principally from the nature of ordering processes used for supply chain control. In an environment where a large number of supply/value chains are being managed, identifying and prioritising those with high rogue seasonality as compared to others could increase both the timeliness as well as quality of management intervention. The intervention could involve either minimising the intensity of rogue seasonality (Kim and Springer, 2008) or optimising it with other ordering considerations (Jaksic and Rusjan, 2008) to improve cost efficiencies. This study focuses on the rogue seasonality detection or the identification and prioritisation aspect and uses data from a generic three echelon simulation model and a case study to: 1) identify the signature that characterises presence of rogue seasonality 2) determine an objective measure called the rogue seasonality index to quantify the signature and use it for rogue seasonality detection and, 3) evaluate alternative transformation techniques and features to identify the most appropriate one for rogue seasonality detection. The signature of rogue seasonality was determined on the basis of the clustering profile of the supply chain variables, specifically, whether exogenous demand did or did not get clustered with the other internal operating variables such as inventory and work in process. To overcome the subjectivity involved in assessing the signature from the profiles of clusters, an objective measure called the rogue seasonality index was defined. It is based on dissimilarities among the variables in comparison to dissimilarity between the variables and the exogenous demand and was found to be a valid and robust measure for rogue seasonality intensity. Effectiveness of different transformation techniques and features such as Fourier (FT), discrete wavelet (DWT), autocorrelation function (ACF), cross correlation function (CCF) and autoregressive model (AR model) was assessed, by computing the index values from each of them.

Compared to other features, the FT feature was found to be the best for rogue seasonality detection, with index values derived from it showing high consistency, discrimination ability and stochastic stability in detection. It has been shown to be effective on a generic basis for individual time series by Agrawal et al. (1993), Caiado et al. (2006) and Vlachos et al. (2006). This study shows that it can also be applied to multivariate data as well as data with inconsistent periodicity and thus validates the findings of Thornhill and Naim (2006). Parameters used, which for FT is the number of frequency channels, was found to significantly impact the index values and hence effectiveness of detection. FT Total which uses features from all the frequency channels was found to perform better than options such as FT 7 and FT 28 which use lesser number of frequencies. Choice of parameters therefore constitutes an important factor during application. Relative superiority of FT also answers the question about appropriate features to be used for clustering time series in the supply chain domain. However, this requires further empirical work. Regarding other features, ACF was a close second to FT in terms of rogue seasonality detection. DWT and AR model features on the other hand were found to be the least effective. The findings for DWT are surprising, as the multi resolution property of wavelets was expected to yield better results than FT. With regard to the AR model based approach, it can be concluded that it is less effective for data with cyclical characteristics.

The index defined in this paper helps in assessing rogue seasonality in a supply chain system independent of bullwhip. Use of multiple variables in its computation makes it less susceptible to inconsistencies in some variable profiles. This is also consistent with the view of Disney and Towill (2003), who in the context of the measure for bullwhip show how using one variable, such as order, gives an incomplete and erroneous picture, and requires additional variables such as inventory to be included. The index is robust and maintains its validity even for a case with rogue seasonality where some variables do not exhibit seasonal characteristics and could be similar to CONS. This is because though the numerator in the index is reduced (refer to Equation 1), alignment of the remaining variables at common amplification frequencies reduces the denominator value significantly to still generate a high index value for accurate detection of rogue seasonality. Alternative indexes proposed such as Alternate 1 (refer to Equation 2) have this robustness built into their definition.

While this study suggested a new approach for rogue seasonality detection it also raised a few questions

such as: a) Would the index be effective in a supply network, and if not, how should it be modified for different network configurations, b) How to identify the main source/contributor of rogue seasonality based on index values at select points in the chain/network, c) Would the index be effective for systems characterised by significant batching and backlogs. The effectiveness of the index therefore needs to be investigated using more complex (and realistic) simulation contexts as well as in other empirical settings.

References:

Agrawal, R., Faloutsos, C., & Swami, A. 1993. Efficient similarity search in sequence databases, Proc. of the 4th International conference on foundations of data organization and algorithms, 69-84. Chicago, IL, USA.

Akaike, H. 1981. Likelihood of a model and information criteria. Journal of Econometrics, 16(1), 3-14.

Baragona, R. 2001. A simulation study on clustering time series with metaheuristic methods. Quaderni di Statistica, 3, 1-26

Basseville, M., & Nikiforov, I. V. 1993. Detection of abrupt changes: theory and application. Englewood Cliffs, N.J.: Prentice Hall.

Bendoly, E., Donohue, K., & Schultz, K. L. 2006. Behavior in operations management: Assessing recent findings and revisiting old assumptions. Journal of Operations Management, 24(6), 737-752.

Bohte, Z., Cepar, D., & Kosmelj, K. 1980. Clustering of time series, Proceedings in Computational Statistics: Physica-Verlag.

Box, G. E. P., & Jenkins, G. M. 1976. Time series analysis: forecasting and control. San Francisco: Holden-Day.

Buxey, G. 1995. A managerial perspective on aggregate planning. International Journal of Production Economics, 41(1-3), 127-133.

Caiado, J., Crato, N., & Pena, D. 2006. A periodogram-based metric for time series classification. Computational Statistics and Data Analysis, 50(10), 2668-2684.

Carranza Torres, O., & Villegas Moran, F. 2006. Managing the bullwhip effect through benchmarking. In O. Carranza Torres, & F. Villegas Moran (Eds.), The Bullwhip effect in supply chains - A review of methods, components and cases: 215-242. Basingstoke: Palgrave Macmillan.

Chan, K. P., & Fu, A. W. C. 1999. Efficient Time Series Matching by wavelets, Proceedings of The International Conference On Data Engineering: 126-135. Sydney, Australia.

Chatfield, C. 2004. The Analysis of Time Series: An Introduction. Boca Raton, Florida: CRC Press.

Chen, F., Drezner, Z., Ryan, J. K., & Simchi-Levi, D. 2000. Quantifying the bullwhip effect in a simple supply chain: the impact of forecasting, lead times, and information. Management Science, 46(3), 436-443.

Chopra, S., & Meindl, P. 2001. Supply chain management: strategy, planning, and operation. Upper Saddle River, N.J.: Prentice Hall.

Christopher, M., & Lee, H. 2004. Mitigating supply chain risk through improved confidence. International Journal of Physical Distribution & Logistics Management, 34(5), 388-396.

Cortes, C., & Pregibon, D. 2001. Signature-based methods for data streams. Data Mining and Knowledge Discovery, 5(3), 167-182.

Davis, T. 1993. Effective Supply Chain Management. Sloan Management Review, 34(4), 35-46.

Dejonckheere, J., Disney, S., Lambrecht, M., & Towill, D. R. 2003. Measuring and avoiding the bullwhip effect: A control theoretic approach. European Journal of Operational Research 147, 567-590.

Disney, S. M., & Towill, D. R. 2003. On the bullwhip and inventory variance produced by an ordering policy. Omega, 31(3), 157-167.

Disney, S. M., Towill, D. R., & Van de Velde, W. 2004. Variance amplification and the golden ratio in production and inventory control. International Journal of Production Economics, 90(3), 295-309.

Disney, S. M., Towill, D. R., & Warburton, R. D. H. 2006. On the equivalence of control theoretic, differential, and difference equation approaches to modeling supply chains. International Journal of Production Economics, 101(1), 194-208.

Fok, D., Franses, P.H., & Paap, R. 2007. Seasonality and non-linear effects in scanner-data-based market-response models. Journal of Econometrics, 138, 231-251

Forrester, J. W. 1961. Industrial Dynamics. Cambridge, Mass.: M.I.T. Press.

Geary, S., Disney, S. M., & Towill, D. R. 2006. On bullwhip in supply chains-historical review, present practice and expected future impact. International Journal of Production Economics, 101(1), 2-18.

Gilbert, K. 2005. An ARIMA supply chain model. Management Science, 51(2), 305-310.

Grosvenor, F. & Austin, T. A. 2001. Cisco's eHub Initiative. Supply Chain Management Review, 5(4), 28-35

Haeckel, S. H., & Nolan, R. L. 1993. Managing by wire. Harvard Business Review, 71(5), 122-132.

Halkidi, M., Batistakis, Y., & Vazirgiannis, M. 2001. On clustering validation techniques. Journal of Intelligent Information Systems, 17(2-3), 107-145.

Han, J., & Kamber, M. 2006. Data Mining: Concepts and Techniques. San Fransisco: Morgan Kaufmann.
Jaksic, M., & Rusjan, B. 2008. The effect of replenishment policies on the bullwhip effect: A transfer function approach. European Journal of Operational Research, 184(3), 946-961.
John, S., Naim, M. M., & Towill, D. R. 1994. Dynamic analysis of a WIP compensated decision support system. International Journal of Manufacturing System Design, 1(4), 283-297.
Kalpakis, K., Gada, D., & Puttagunta, V. 2001. Distance Measures for effective clustering of ARIMA Time-Series, Proceedings of the 2001 IEEE International Conference on Data Mining. San Jose, CA: IEEE Computer Society.
Keogh, E., & Kasetty, S. 2003. On the need for time series Data Mining Benchmarks: A Survey and Empirical Demonstration. Data Mining and Knowledge Discovery, 7(4), 349-371.
Kim, I., & Springer, M. 2008. Measuring endogenous supply chain volatility: Beyond the bullwhip effect. European Journal of Operational Research., 189(1), 172-193.
Last, M., Kandel,A.&Bunke, H.2004.Data Mining in Time Series Databases. Singapore.World Scientific Publishing
Lee, H. L., Padmanabhan, V., & Whang, S. 1997a. The Bullwhip effect in supply chains. Sloan Management Review, Spring, 93-102.
Lee, H. L., Padmanabhan, V., & Whang, S. 1997b. Information distortion in a supply chain: The bullwhip effect. Management Science 43(4), 546-558.
Li, T., Li, Q., Zhu, S., & Ogihara, M. 2002. A survey on wavelet applications in data mining. SIGKDD Exploration. Newsletter., 4(2), 49-68.
Liao, T. W. 2005. Clustering of time series data-a survey. Pattern Recognition. 38(11), 1857-1874.
Maharaj, E. A. 2000. Clusters of time series. Journal of Classification, 17(2), 297-314.
Mason-Jones, R., Naim, M. M. and Towill, D. R. 1997. The Impact of Pipeline Control on Supply Chain Dynamics", International Journal of Logistics Management, 8(2) pp 47 - 62
Metters, R. 1997. Quantifying the Bullwhip Effect in Supply Chains. Journal of Operations Management, 15(2), 89-100.
Naim, M. M., Wikner, J., & Grubbstrom, R. W. 2007. A net present value assessment of make-to-order and make-to-stock manufacturing systems. Omega, 35(5), 524-532.
Orfanidis, S. 2002. Introduction to signal processing. New Jersey: Prentice Hall
Otto, A. 2003. Supply Chain Event Management: Three Perspectives. International Journal of Logistics Management, 14(2), 1-13.
Percival, D. B., & Walden, A. T. 2000. Wavelet methods for time series analysis. Cambridge; New York: Cambridge University Press.
Shannon, C. E., & Weaver, W. 1964. The Mathematical Theory of Communication. University of Illinois Press
Shukla, V., Naim, M.M., and Yaseen, E.A., 2009. `Bullwhip' and `backlash' in supply pipelines. International Journal of Production Research, 47 (23), pp. 6477-6497.
Sterman, J. 1989. Modelling managerial behaviour: Misperceptions of feedback in a dynamic decision-making experiment. Management Science 35(3), 321–339.
Sy, B. K. 2005. Signature-based approach for intrusion detection, Lecture Notes in Computer Science. Vol. 3587 LNAI,526-536.
Thornhill, N. F., & Naim, M. M. 2006. An exploratory study to identify rogue seasonality in a steel company's supply network using spectral principal component analysis. European Journal of Operational Research, 172, 146-162.
Ting, J. K., Ng, M. K., Rong, H., & Huang, J. Z. 2003. Statistical models for time sequences data mining, I E E E International Conference on Computational Intelligence for Financial Engineering, 347-354. Hong Kong, China.
Torres, O. A. C., and Moran, F. A. V. 2006. Managing the Bullwhip effect through Benchmarking in The Bullwhip effect in supply chains: A review of methods, components and cases, O. A. C. Torres and F. A. V. Moran, eds., Palgrave Macmillan, New York, 215-242.
Towill, D. R., & del Vecchio, A. 1994. The application of filter theory to the study of supply chain dynamics Production Planning and Control, 5 (1), 82-96.
Towill, D.R. 1997. FORRIDGE-Principles of good practice in material flow. Production Planning & Control, 8(7), 622-632
Towill, D. R., Evans, G. N., & Cheema, P. 1997. Analysis and design of an adaptive minimum reasonable inventory control system. Production Planning and Control, 8(6), 545-557.

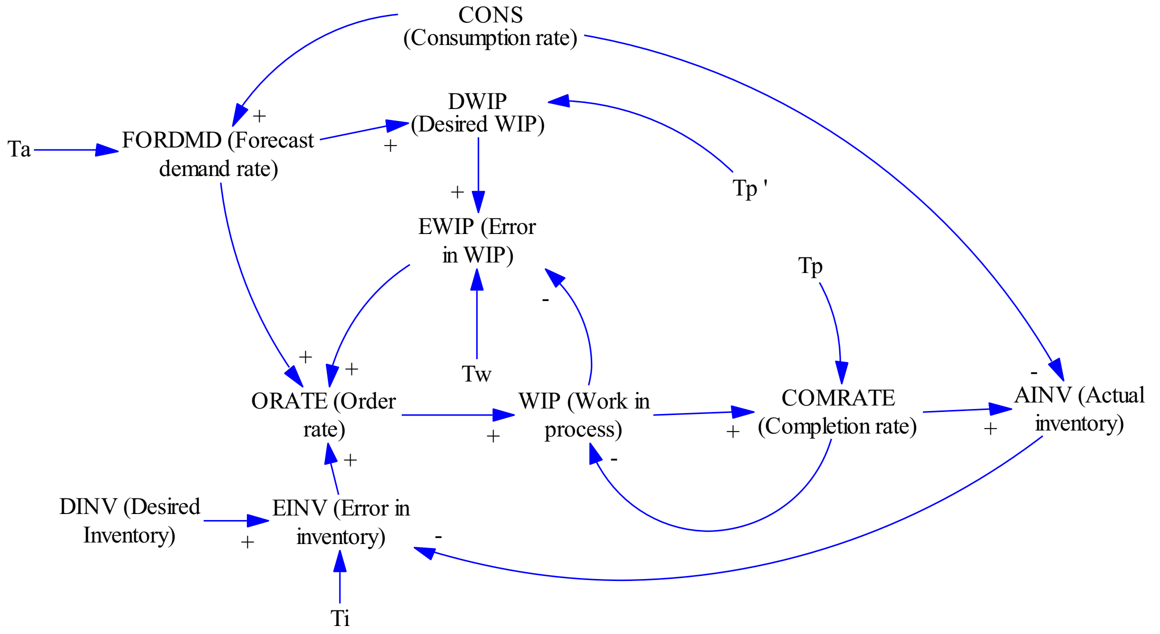
Tripp, R. S., Amouzegar, M. A., McGarvey, R. G., Bereit, R., George, D., & Cornuet, J. 2006. Sense and Respond Logistics Integrating Prediction, Responsiveness, and Control Capabilities: Rand.

Vlachos, M., Yu, P. S., Castelli, V., & Meek, C. 2006. Structural Periodic Measures for Time-Series Data. Data Mining and Knowledge Discovery, 12(1), 1-28.

Wang, C., & Wang, X. S. 2000. Supporting content-based searches on time series via approximation. Proceedings of the 12th International Conference on Scientific and Statistical Database Management.

Wu, D.S., Kempf, K.G., Atan, M. O., Aytac B., Shirodkar, S.A., & Mishra, A. 2010. Improving New-Product Forecasting at Intel Corporation. Interfaces, 40(5), 385-396

Zhang, H., Ho, T., & Huang, W. 2005. Blind Feature Extraction for Time-Series Classification Using Haar Wavelet Transform. In J. Wang, X. Liao, & Z. Yi (Eds.), Advances in Neural Networks – ISNN 2005, 605-610. Berlin Heidelberg: Springer-Verlag.



Forecasted Demand Rate: FORDMD (t) = FORDMD (t –1) + α [CONS (t) – FORDMD (t –1)] (1)
where; $\alpha = 1/(1+ T_a/\Delta t)$, CONS (t) is the demand at time t, and for a linear, time invariant system is a surrogate for sales, T_a is the time to average demand, Δ t is our simulation time increment set at 1

Factory Order Rate: ORATE (t) = FORDMD (t) + [EINV (t –1) / T_i] + [EWIP (t –1) / T_w] (2)
where; $1/T_i$ is the fraction of inventory feedback and $1/T_w$ if the fraction of work in process or WIP feedback

Work in process level: WIP (t) = WIP (t-1) + ORATE (t) – COMRATE (t-1) (3)

Completion Rate/Receipts into inventory: COMRATE (t) = WIP (t) / T_p ------ *First order lag* (4a)
COMRATE (t) = ORATE (t –T_p) -- *Pure delay lag* (4b)
where T_p is the production lead time

Actual Inventory level: AINV (t) = AINV (t-1) + COMRATE (t) – CONS (t) (5)

Error in Inventory: EINV (t) = DINV – AINV (t) (6)
where; DINV is a constant inventory target level

Error in Work in Process or EWIP: EWIP (t) = DWIP (t) – WIP (t) (7a)
where; DWIP (t) = T_p’ * FORDMD (t); (7b)
T_p’ is the estimated production lead time and is assumed to be equal to T_p
DWIP is the desired work in process level

Figure 1. Causal loop diagram and simulation equations for APIOBPCS

Table 1

Simulation Design – Time series data generation

Structure and number of echelons	Linear, Three Echelons
Supply chain strategy	Make to Order (MTO) and Make to Stock (MTS)
Delay order	Order 1, Order infinity/pipeline
Delay (time)	3,7,14
Demand process and parameters	AR (1): ρ = -0.8,-0.5, 0.1 MA (1): θ = 0.7, 0.4, -0.2 AR (2): ρ_1 = 0.1 ρ_2 = -0.8, ρ_1 = 0.7 ρ_2 = -0.2 MA (2): θ_1 = 0.7 θ_2 = -0.2, θ_1 = 0.1 θ_2 = -0.8 Gaussian
Exogenous seasonality	Absent, Present (Amplitude: 1 frequency: 0.2, frequency: 0.05)
100 replications of each based on common random numbers	

Table 2

Time series transformation techniques and features used

Fourier transform (FT)	Amplitudes used as features. Number of frequencies considered: a) Total (FT Total) b) First 7 (FT 7) c) First 28 (FT 28) d) Number with 80% of total energy in the data series (FT 80% Energy)
Autocorrelation function (ACF)	Autocorrelation upto specified lags used as features. Lags considered: a) 7 (ACF 7) b) 28 (ACF 28)
Cross correlation function (CCF)	Maximum Cross correlation between pairs of variables within maximum lags specified, used as the feature/similarity measure. Maximum lags considered: a) 7 (CCF 7) b) 28 (CCF 28)
Autoregressive model (AR)	Coefficients of the fitted AR model, upto specified orders used as features. Different order of AR models considered: a) 7 (AR 7) b) 28 (AR 28) c) Minimum AIC (Akaike, 1981) (AR AIC)
Discrete wavelet transform (DWT) – Haar wavelet	Wavelet coefficients from DWT at different levels used as features. Levels considered: a) 5 (DWT 5) b) Based on minimum Shannon entropy (Shannon and Weaver, 1964) (DWT Shannon)
Time series data without any transformation was used as the default for comparison (Raw Time)	

Bracketed text are feature parameter combination as referred in the document

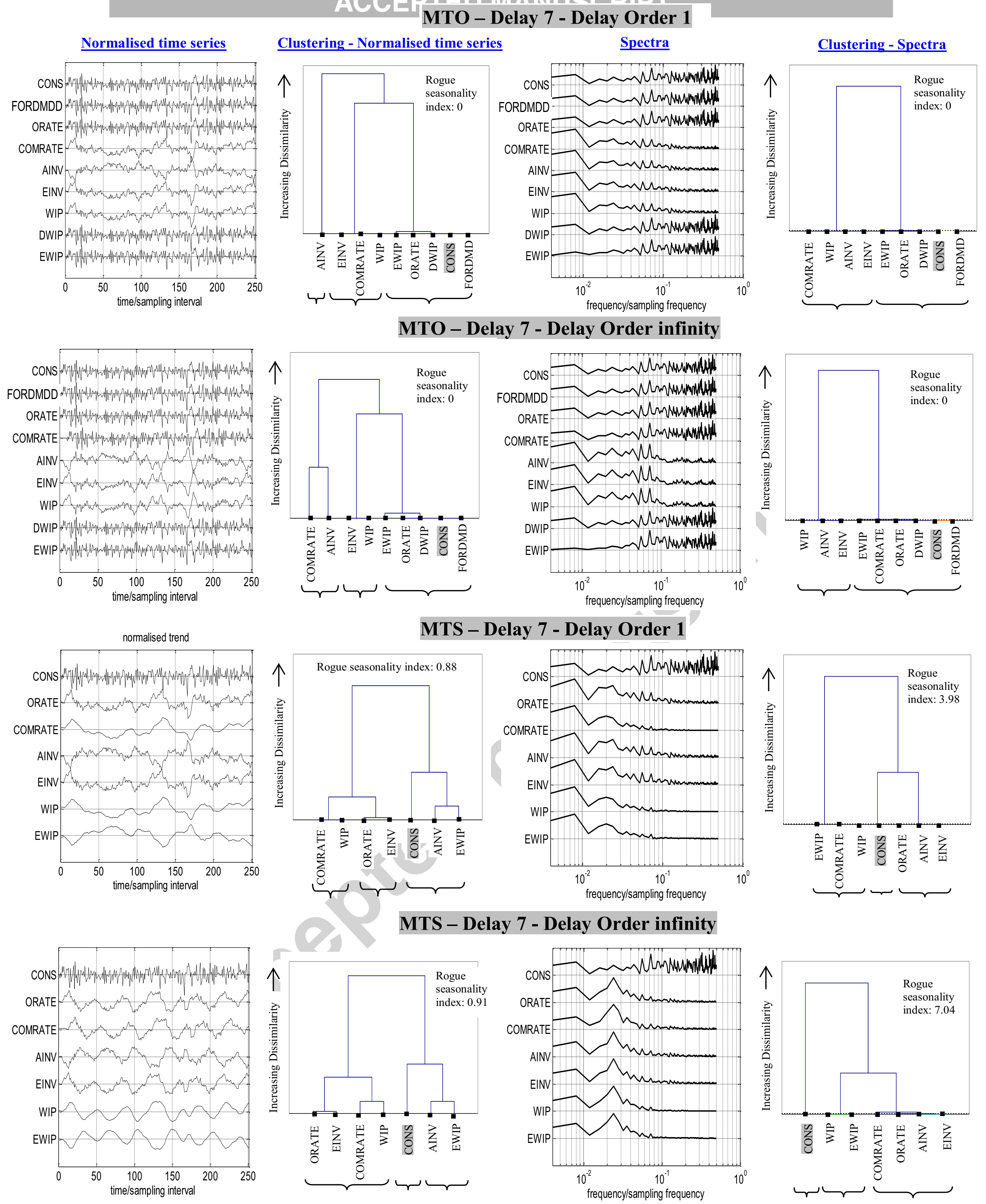
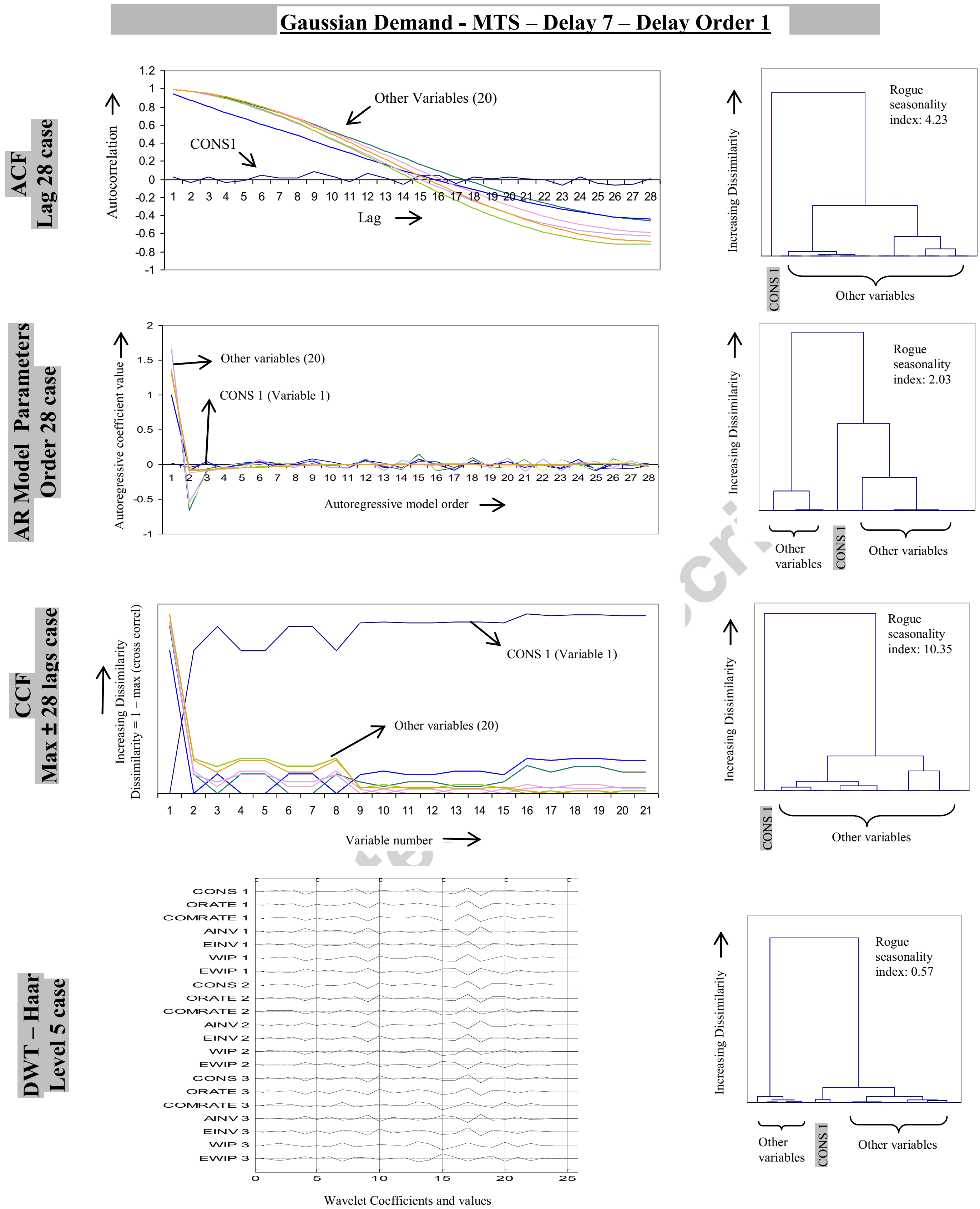


Figure 2. Time series, spectra and their clustering for a single echelon system excited by Gaussian demand (CONS)



PS: Other variables: APIOBPCS variables excl. CONS 1 (Consumption or demand for Echelon 1)

Figure 3. Plots and clusters of different features for a three echelon system excited by Gaussian demand (CONS1)

Table 5

Consistency of Rogue Seasonality index for different demand parameters and order of delays (Number of cases)

Based on average rogue seasonality index value

Nature of Demand Process	Consistency Basis	Raw Time	FT Total	FT 7	FT 28	FT 80% Energy	AR 7	AR 28	AR AIC	ACF 7	ACF 28	CCF 7	CCF 28	DWT Haar (Level 5)	DWT Haar (Level Shannon)
AR 1	Demand parameters	12/12	11/12	3/12	3/12	6/12	11/12	11/12	11/12	11/12	12/12	12/12	12/12	1/12	6/12
	Order of delay	7/9	9/9	0/9	4/9	9/9	1/9	0/9	0/9	9/9	9/9	2/9	4/9	1/9	7/9
MA 1	Demand parameters	12/12	12/12	4/12	6/12	7/12	6/12	6/12	6/12	11/12	12/12	11/12	11/12	4/12	10/12
	Order of delay	7/9	9/9	0/9	4/9	9/9	1/9	0/9	0/9	9/9	9/9	3/9	5/9	5/9	7/9
AR 2	Demand parameters	5/6	6/6	3/6	1/6	5/6	3/6	2/6	2/6	6/6	6/6	5/6	6/6	1/6	6/6
	Order of delay	4/6	6/6	1/6	3/6	6/6	2/6	1/6	1/6	6/6	6/6	2/6	3/6	3/6	6/6
MA 2	Demand parameters	6/6	6/6	2/6	0/6	3/6	1/6	1/6	1/6	6/6	6/6	6/6	6/6	0/6	0/6
	Order of delay	4/6	6/6	0/6	3/6	6/6	0/6	0/6	0/6	6/6	6/6	1/6	2/6	1/6	5/6
Gaussian	Order of delay	2/3	3/3	1/3	1/3	3/3	0/3	0/3	0/3	3/3	3/3	0/3	1/3	1/3	2/3
	Overall Consistency - Demand and Order of delay	**59/69**	**68/69**	**14/69**	**25/69**	**54/69**	**25/69**	**21/69**	**21/69**	**67/69**	**69/69**	**42/69**	**50/69**	**17/69**	**49/69**

Table 6

Rank analysis based on value of Rogue Seasonality index

(For assessment of discrimination ability of Rogue Seasonality index based on different transformations and features)

Larger the rogue seasonality index higher the rank is used as basis

	Raw Time	FT Total	FT 7	FT 28	FT 80% Energy	AR 7	AR 28	AR AIC	ACF 7	ACF 28	CCF 7	CCF 28	DWT Haar (Level 5)	DWT Haar (Level Shannon)
Sum of Ranks - AR (1) Minimum / best possible 18	203	98	195	128	74	144	183	159	22	89	89	39	243	224
Sum of Ranks - MA (1) Minimum / best possible 18	204	102	191	125	89	133	173	151	28	104	90	34	242	224
Sum of Ranks - AR (2) Minimum / best possible 12	133	59	120	80	71	97	121	112	18	56	57	23	163	150
Sum of Ranks - MA (2) Minimum / best possible 12	136	64	130	84	49	99	118	113	17	58	57	25	161	149
Sum of Ranks - Gaussian Minimum / best possible 6	65	31	71	50	25	45	61	54	8	28	25	12	81	74
Sum of Ranks - TOTAL Minimum / best possible 66	**741**	**354**	**707**	**467**	**308**	**518**	**656**	**589**	**93**	**335**	**318**	**133**	**890**	**821**

For each demand parameter, delay magnitude and order combination the techniques are ranked with the highest rank given to that which has the highest rogue seasonality magnitude. These ranks have been summed up for each demand process and technique; The lower the value (sum of ranks), the better is the technique as rogue seasonality index derived from it would give better discrimination

Table 7

Rank analysis based on Coefficient of Variation of Rogue Seasonality index

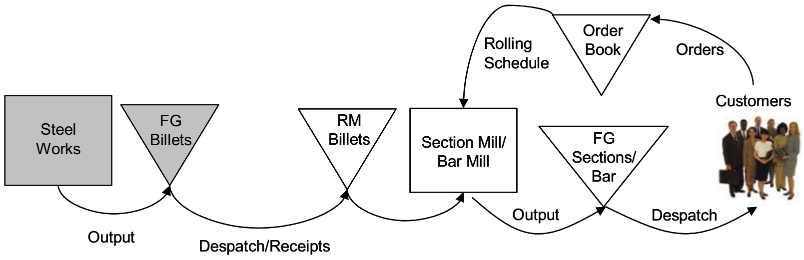
(For assessment of stochastic robustness of Rogue Seasonality index based on different transformations and features)

Lower the coefficient of variation the higher the rank is used as basis

	Raw Time	FT Total	FT 7	FT 28	FT 80% Energy	AR 7	AR 28	AR AIC	ACF 7	ACF 28	CCF 7	CCF 28	DWT Haar (Level 5)	DWT Haar (Level Shannon)
Sum of Ranks - AR (1) Minimum / best possible 18	21	63	215	95	109	202	187	211	186	191	106	153	83	68
Sum of Ranks - MA (1) Minimum / best possible 18	18	64	213	103	98	219	192	217	181	190	90	134	90	81
Sum of Ranks - AR (2) Minimum / best possible 12	12	43	129	67	73	145	141	155	116	120	57	88	61	53
Sum of Ranks - MA (2) Minimum / best possible 12	12	43	150	70	68	132	128	134	125	136	65	96	54	47
Sum of Ranks - Gaussian Minimum / best possible 6	6	19	76	32	27	75	70	77	56	58	26	41	27	40
Sum of Ranks - TOTAL Minimum / best possible 66	**69**	**232**	**783**	**367**	**375**	**773**	**718**	**794**	**664**	**695**	**344**	**512**	**315**	**289**

Coefficient of variation is same as SD/Avg; Their values are based on 100 replications for each

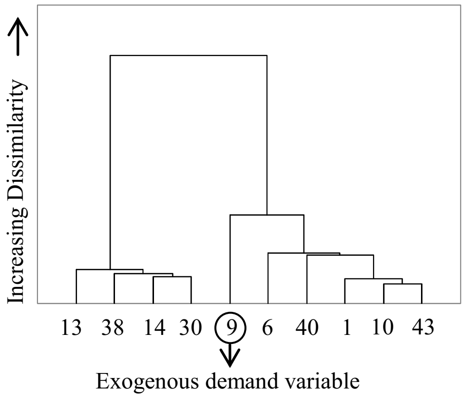
	Raw Time	FT Total	FT 7	FT 28	FT 80% Energy	AR 7	AR 28	AR AIC	ACF 7	ACF 28	CCF 7	CCF 28	DWT Haar (Level 5)	DWT Haar (Level Shannon)
Sum of Ranks - TOTAL based on Value and Coefficient of Variation of Rogue seasonality index Minimum / best possible 132	**810**	**586**	**1490**	**834**	**683**	**1291**	**1374**	**1383**	**757**	**1030**	**662**	**645**	**1205**	**1110**



Steel Works – Section/Bar Mill supply dyads

Steel Works - Section Mill Dyad

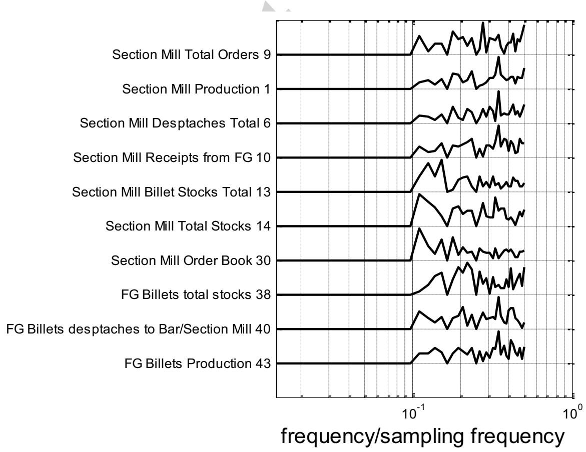
Clustering of FT amplitudes



Rogue Seasonality Index values

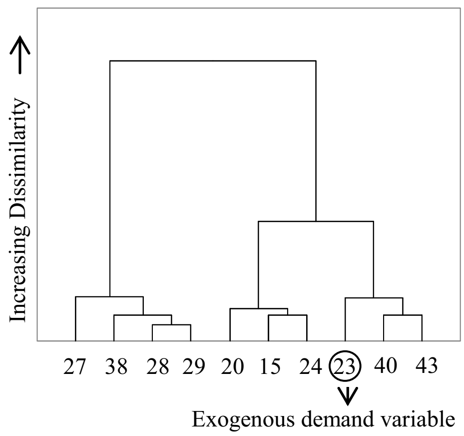
Techniques	Values
FT Total (All frequencies)	0.94
Time Domain	0.7
ACF (Lag 36)	0.74
CCF (Max Lags ± 36)	0.78
AR Model (Order 36)	0.6
DWT Haar (Level 2)	0.5

Spectra plot



Steel Works - Bar Mill Dyad

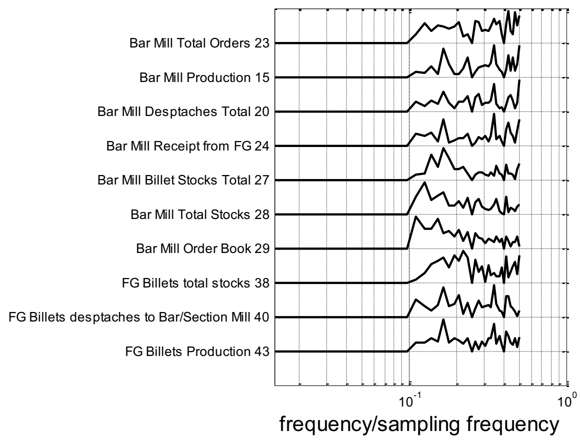
Clustering of FT amplitudes



Rogue Seasonality Index values

Techniques	Values
FT Total (All frequencies)	0.93
Time Domain	0.77
ACF (Lag 36)	0.98
CCF (Max Lags ± 36)	1
AR Model (Order 36)	1.08
DWT Haar (Level 2)	0.7

Spectra plot



Note: Variable number descriptions for each dyad are given in their respective spectra plots

Figure 4. Rogue seasonality signature and index for steel case data

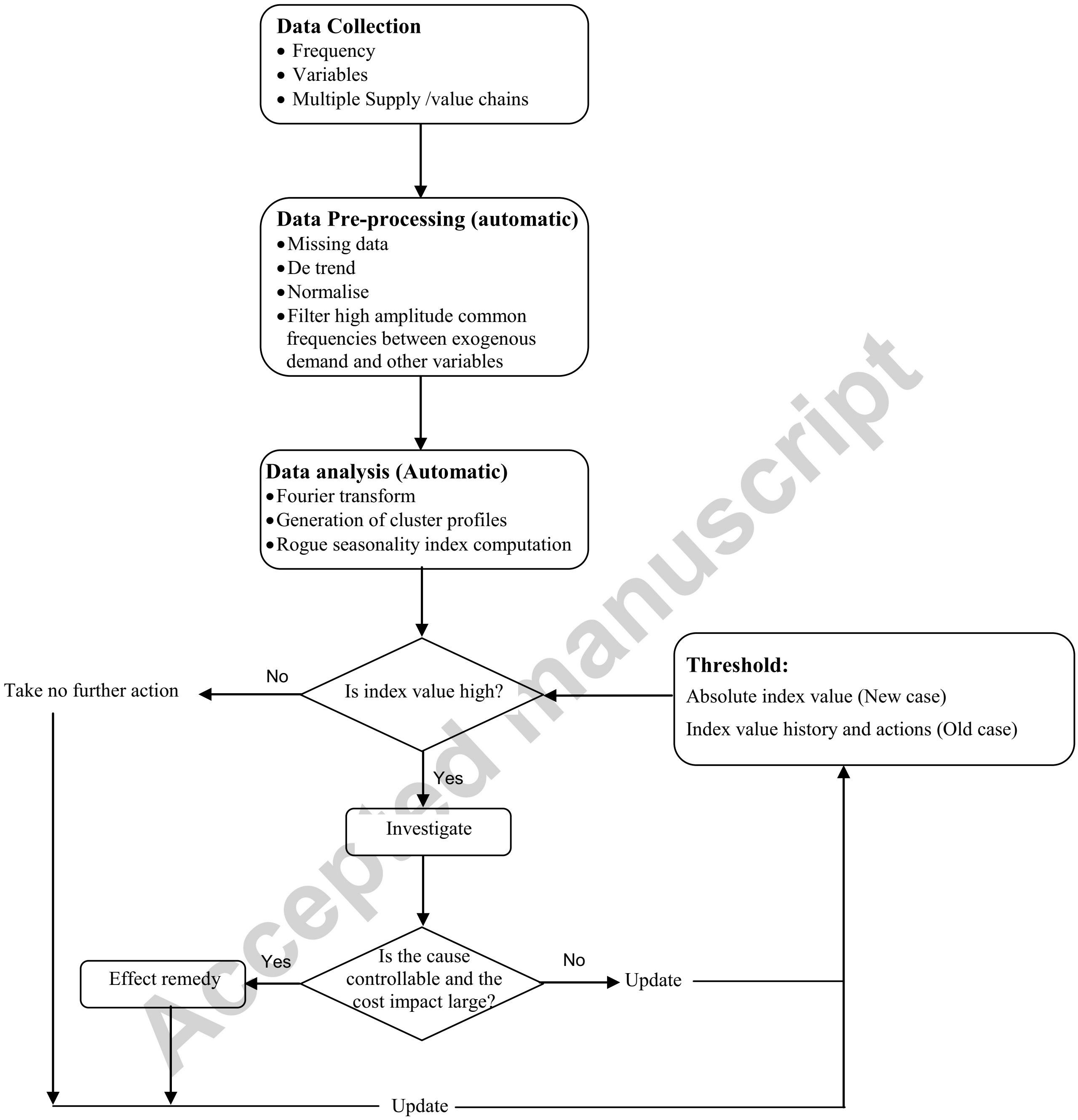


Figure 5. Flow chart for rogue seasonality index application