UNIVERSIDAD DE OVIEDO

Programa de Doctorado en Administración de Empresas

Área de Organización de Empresas

DOCTORAL THESIS

BULLWHIP EFFECT REDUCTION THROUGH ARTIFICIAL INTELLIGENCE-BASED TECHNIQUES

PhD Candidate:

BORJA PONTE BLANCO

Supervised by:

DAVID ALFONSO DE LA FUENTE GARCÍA

and

RAÚL PINO DÍEZ

Oviedo, 2016

RESUMEN DEL CONTENIDO DE TESIS DOCTORAL

1.- Título de la Tesis	
Español: **REDUCCIÓN DEL EFECTO BULLWHIP MEDIANTE TÉCNICAS BASADAS EN INTELIGENCIA ARTIFICIAL**	Inglés: **BULLWHIP EFFECT REDUCTION THROUGH ARTIFICIAL INTELLIGENCE-BASED TECHNIQUES**

2.- Autor	
Nombre: **BORJA PONTE BLANCO**	DNI/Pasaporte/NIE:
Programa de Doctorado: **PROGRAMA DE DOCTORADO EN ADMINISTRACIÓN DE EMPRESAS**	
Órgano responsable: **DEPARTAMENTO DE ADMINISTRACIÓN DE EMPRESAS**	

RESUMEN (en español)

La globalización ha revolucionado el contexto empresarial. El incremento de la oferta en bienes y servicios, los constantes cambios en los gustos de los consumidores y la expansión geográfica de las redes de distribución, entre otros factores, han trazado un nuevo entorno competitivo marcado por la intensidad, la complejidad y el dinamismo. Éste ha enfatizado el concepto de cadena de suministro. En los procesos, en las relaciones y en las interdependencias de la cadena de suministro se esconde una fuente clave de ventajas competitivas para las organizaciones que, sin embargo, es muy compleja de captar. Una de las razones de ello es la generación del denominado Efecto Bullwhip, que ha de entenderse como una fuente clave de ineficiencias en la cadena de suministro. Este fenómeno se refiere a la amplificación de la variabilidad de las órdenes transmitidas a lo largo del sistema.

Los capítulos 1 a 3 del presente trabajo exploran el papel de la inteligencia artificial en el desarrollo de mecanismos de previsión orientados a mejorar la gestión de la cadena de suministro. Se han utilizado redes neuronales artificiales (artificial neural networks, ANNs), bajo arquitecturas del tipo perceptron multi-capa (multi-layer perceptron, MLP) y funciones de base radial (RBF), junto a métodos estadístico dentro de una estructura multi-agente. Ante demandas con tendencia y estacionalidad, el sistema —que escoge en cada momento la previsión más adecuada— obtiene un gran rendimiento en la reducción del Efecto Bullwhip desde una perspectiva local. Asimismo, se muestra cómo este sistema se podría integrar con facilidad en un sistema de mayor alcance, lo cual representa una de las principales ventajas de esta aproximación.

Los capítulos 4 a 6, que representan la principal línea de investigación dentro de este trabajo, tratan esta problemática desde una perspectiva sistémica. En este sentido, se pretende contribuir al despliegue de esta perspectiva dentro de las cadenas de suministro; el cual entendemos como el gran reto de las cadenas de suministro en el siglo XXI. Con este objetivo, se desarrolla un marco integrador para la gestión colaborativa de sistemas de producción y distribución basado en el Modelo de los Sistemas Viables de Beer (Viable System Model, VSM) y la Teoría de las Restricciones de Goldratt (Theory of Constraints, TOC). Sobre este marco, se explora la implementación de la solución mediante herramientas de modelado y simulación. Más en concreto, se utiliza la metodología Drum-Buffer-Rope (DBR) para proponer un motor operativo para la cadena de suministro y demuestra su eficacia, en comparación con alternativas tradicionales basadas en la producción en masa, tanto en términos operacionales (donde se engloba el Efecto Bullwhip) como en términos económicos.

No obstante, el trabajo subraya que la integración de procesos es sólo una de las áreas clave para el diseño de soluciones colaborativas. La transparencia en la información relevante, la sincronización y distribución en la toma de decisiones, y el diseño de un sistema de rendimiento global han de entenderse igualmente como condiciones sine qua non para la implementación exitosa de la colaboración en las cadenas de suministro. La alineación de incentivos también es esencial. Los riesgos y los beneficios han de ser compartidos adecuadamente con el objetivo de reducir la amenaza de comportamientos oportunistas. Los cinco campos mencionados se han considerado en la propuesta de una solución colaborativa viable y beneficiosa para todos los miembros; dado que este esquema nos permite comprender por qué solo un pequeño porcentaje de las cadenas de suministro reales son capaces de crear valor a través de la colaboración.

Esta Tesis Doctoral también pretende resaltar las técnicas de modelado y simulación como poderosos laboratorios de ensayo para el estudio de grandes problemas organizacionales que serían complejos de estudiar de otra forma. Este hecho subraya el enorme potencial del desarrollo de prototipos como metodología para el apoyo a la toma de decisiones y la transformación empresarial, especialmente en torno al complejo proceso de transición de una aproximación reduccionista (basada en la optimización local) a una holista (basada en la optimización global) en la cadena de suministro.

RESUMEN (en Inglés)

Globalization has utterly changed the economic landscape. The increase in the supply of goods and services, the constant evolution in customer preferences, and the geographical expansion of distribution networks, among other factors, have set up a new competitive environment—marked by intensity, complexity, and dynamism—that has put a greater emphasis on the concept of supply chain. Supply chain processes, relationships, and interdependencies can be a key source of competitive advantages. However, these advantages are difficult to capture. One of the reasons behind it is the generation of the so-called Bullwhip Effect, a major source of inefficiencies within supply chains. It refers to the amplification of the variability of orders throughout the system.

Chapters 1 to 3 in the present dissertation explore the role of artificial intelligence in the development of forecasting mechanisms that improve the management of the supply chain. We employ artificial neural networks (ANNs), both under multi-layer perceptron (MLP) and radial basis function (RBF) architectures, together with statistical methods within a multi-agent structure. Facing demand series with trend and seasonality, the system—that selects the most suitable forecast for every moment—greatly mitigates the generation of the Bullwhip Effect from a local perspective. In addition, we show how this system could be easily integrated in a system with a larger scope, which represents one of the main benefits of this approach.

Chapters 4 to 6, which represent the main research stream of this research work, analyze this issue from a systemic perspective. In this sense, we aim to add to the deployment of this view throughout supply chains; which we understand as a major challenge for 21st-century supply chains. To this end, we develop an integrative framework for the collaborative management of production and distribution systems based on the Beer's Viable System Model (VMS) and on Goldratt's Theory of Constraints (TOC). Building upon this framework, we investigate the implementation of this solution though modelling and simulation techniques. Specifically, we design a Drum-Buffer-Rope (DBR) mechanism to act as the operational engine for the supply chain. We show its effectiveness in comparison with traditional alternatives based on the mass production paradigm both in operational (including the Bullwhip Effect) and financial terms.

Notwithstanding the foregoing, the present dissertation also underscores that process integration is only one of the key fields within the development of collaborative solutions for supply chains. Transparency in the relevant information, synchronization and allocation in the decision making must also be understood as conditions sine qua non for the successful implementation of collaboration across the system. Aligning incentives is also essential. In this regard, risks and benefits must be shared appropriately to reduce the menace of opportunistic behaviors. We carefully take into consideration all these fields in order to make the collaborative solution viable and profitable for every node, since we believe that this five-edge scheme makes it easier to understand why only a small percentage of real supply chains are capable of adding value through collaboration.

In this research, modelling and simulation techniques appear as powerful laboratories for the study of large organizational problems that would be difficult to study otherwise. This fact emphasizes the great potential of prototype development as a methodology for the support of decision making and business transformation, especially around the complex transition process from reductionism (based on local optimization) to holism (based on global optimization) in supply chains.

SR. DIRECTOR DE DEPARTAMENTO DE ADMINISTRACIÓN DE EMPRESAS
SR. PRESIDENTE DE LA COMISIÓN ACADÉMICA DEL PROGRAMA DE DOCTORADO EN ADMINISTRACIÓN DE EMPRESAS

UNIVERSIDAD DE OVIEDO

Programa de Doctorado en Administración de Empresas

Área de Organización de Empresas

DOCTORAL THESIS

BULLWHIP EFFECT REDUCTION THROUGH ARTIFICIAL INTELLIGENCE-BASED TECHNIQUES

PhD Candidate:

BORJA PONTE BLANCO

Supervised by:

DAVID ALFONSO DE LA FUENTE GARCÍA

and

RAÚL PINO DÍEZ

Oviedo, 2016

UNIVERSIDAD DE OVIEDO

Programa de Doctorado en Administración de Empresas

Área de Organización de Empresas

TESIS DOCTORAL

REDUCCIÓN DEL EFECTO BULLWHIP MEDIANTE TÉCNICAS BASADAS EN INTELIGENCIA ARTIFICIAL

Alumno:

BORJA PONTE BLANCO

Dirigida por:

DAVID ALFONSO DE LA FUENTE GARCÍA

y

RAÚL PINO DÍEZ

Oviedo, 2016

Agradecimientos

Quiero agradecer en las primeras líneas de esta Tesis Doctoral a todas aquellas personas e instituciones que han contribuido, de una forma u otra, al desarrollo del presente trabajo durante este largo y complejo, aunque apasionante, camino que empezó en septiembre de 2012.

En primer lugar, a mis Directores de Tesis. Al Dr. David de la Fuente le quiero agradecer la confianza que ha depositado en mí desde que me ofreció la oportunidad de unirme al Grupo de Ingeniería de Organización cuando finalicé mis estudios de Ingeniería Industrial. Su conocimiento, su saber hacer y sus valiosas sugerencias me han ayudado mucho a la hora de afrontar los distintos retos que han ido apareciendo en estos cuatro años. Al Dr. Raúl Pino también le quiero agradecer tanto su confianza como su contribución técnica al desarrollo de la presente Tesis Doctoral, especialmente en lo relacionado con la primera línea de investigación (Capítulos 1, 2 y 3), donde su aportación como experto en previsión de series temporales ha incrementado significativamente la calidad del trabajo realizado.

Quiero extender este primer agradecimiento al resto de miembros del Grupo de Ingeniería de Organización de la Universidad de Oviedo: Dr. Alberto Gómez; Dr. Javier Puente; Dr. Jesús Lozano; Dr. José Parreño; Dr. Nazario García; Dr. Paolo Priore; y, especialmente, a la Dra. Isabel Fernández por su constante predisposición a compartir su valioso conocimiento en el campo de la logística, y al Dr. Rafael Rosillo por sus siempre realistas y sinceros consejos. También quiero expresar mi gratitud hacia el resto de compañeros del Departamento de Administración de Empresas de la Universidad de Oviedo con los que he trabajado durante estos años. Siempre recordaré con cariño haber formado parte de este departamento.

En segundo lugar, a mis coautores José Costas y el Dr. Julio Puche. Con ellos he compartido decenas de interminables (en el buen sentido) charlas por Skype durante estos cuatro años en torno a los trabajos desarrollados en la segunda línea de investigación (Capítulos 4, 5 y 6). En esas conversaciones, guiadas magistralmente por José, he aprendido innumerables y muy preciadas lecciones en el campo de la gestión de la cadena de suministro, así como también muchas otras que lo trascienden. Sin duda, permanecerán en mi memoria como momentos esenciales en el desarrollo de esta Tesis Doctoral.

En tercer lugar, a mis supervisores en The Cardiff Business School (2015) y The Open University Business School (2016), donde he realizado mis estancias de investigación. Por un lado, abrir una línea de trabajo junto al Profesor Stephen Disney, probablemente el mayor experto mundial en el Efecto Bullwhip, ha sido y está siendo una oportunidad fascinante. Estoy enormemente agradecido por su cálida acogida y por su constante ayuda, así como la del resto de la sección de Logistics and Operations Management. Por otro lado, he tenido también la gran suerte de que el Dr. Björn Claes me abriese las puertas del Department for People and Organisations de The Open University Business School, permitiéndome colaborar en distintos proyectos. Su experiencia industrial ha enriquecido significativamente mi visión de la cadena de suministro y es un honor haber colaborado y seguir colaborando con él y su equipo.

También quiero resaltar el relevante papel del Gobierno del Principado de Asturias, a través del Programa de Becas Predoctorales Severo Ochoa, el cual ha contribuido significativamente en mi desarrollo tanto investigador como docente. Igualmente, deseo mostrar mi gratitud con la Fundación Banco Sabadell, que apostó por la presente Tesis Doctoral dentro de la convocatoria 2014 de Ayudas a la Investigación.

Finalmente, pero no por ello menos importante, quiero dar las gracias a mi familia por su apoyo incondicional en la carrera académica que estoy comenzando, así como en tantas otras cosas que serían complicadas de resumir en un solo párrafo. Principalmente a mis padres, de quienes heredé la vocación por esta profesión, a mi hermano Álvaro, quien siempre me ayuda a ver las cosas de otra manera en los momentos más complicados, a mis tíos y a mis primas. Quiero acordarme también de forma especial de mi abuelo Benigno, quien desgraciadamente no podrá acompañarme en la defensa de esta Tesis Doctoral.

En el apartado familiar también se engloba, oficialmente desde hace muy poquito, Raquel, lo cual me hace inmensamente feliz. Tanto en el plano personal como en el profesional, le quiero agradecer que comparta conmigo de forma desinteresada sus grandes conocimientos y habilidades, que haga siempre un esfuerzo por entender mis preocupaciones y retos y ayudarme a superarlos, y que su forma de hacer las cosas sea un ejemplo excelente. Aún me queda mucho que aprender. Me faltan las palabras, también, para agradecer cómo me ha acogido su familia.

ABSTRACT

Globalization has utterly changed the economic landscape. The increase in the supply of goods and services, the constant evolution in customer preferences, and the geographical expansion of distribution networks, among other factors, have set up a new competitive environment—marked by intensity, complexity, and dynamism—that has put a greater emphasis on the concept of supply chain. Supply chain processes, relationships, and interdependencies can be a key source of competitive advantages. However, these advantages are difficult to capture. One of the reasons behind it is the generation of the so-called Bullwhip Effect, a major source of inefficiencies within supply chains. It refers to the amplification of the variability of orders throughout the system.

Chapters 1 to 3 in the present dissertation explore the role of artificial intelligence in the development of forecasting mechanisms that improve the management of the supply chain. We employ artificial neural networks (ANNs), both under multi-layer perceptron (MLP) and radial basis function (RBF) architectures, together with statistical methods within a multi-agent structure. Facing demand series with trend and seasonality, the system—that selects the most suitable forecast for every moment—greatly mitigates the generation of the Bullwhip Effect from a local perspective. In addition, we show how this system could be easily integrated in a system with a larger scope, which represents one of the main benefits of this approach.

Chapters 4 to 6, which represent the main research stream of this research work, analyze this issue from a systemic perspective. In this sense, we aim to add to the deployment of this view throughout supply chains; which we understand as a major challenge for 21st-century supply chains. To this end, we develop an integrative framework for the collaborative management of production and distribution systems based on the Beer's Viable System Model (VMS) and on Goldratt's Theory of Constraints (TOC). Building upon this framework, we investigate the implementation of this solution though modelling and simulation techniques. Specifically, we design a Drum-Buffer-Rope (DBR) mechanism to act as the operational engine for the supply chain. We show its effectiveness in comparison with traditional alternatives based on the mass production paradigm both in operational (including the Bullwhip Effect) and financial terms.

Notwithstanding the foregoing, the present dissertation also underscores that process integration is only one of the key fields within the development of collaborative solutions for supply chains. Transparency in the relevant information, synchronization and *allocation in the decision making must also be understood as conditions sine qua non for the successful implementation of collaboration across the system. Aligning incentives is also essential. In this regard, risks and benefits must be shared appropriately to reduce the menace of opportunistic behaviors. We carefully take into consideration all these fields in order to make the collaborative solution viable and profitable for every node, since we believe that this five-edge scheme makes it easier to understand why only a small percentage of real supply chains are capable of adding value through collaboration.*

In this research, modelling and simulation techniques appear as powerful laboratories for the study of large organizational problems that would be difficult to study otherwise. This fact emphasizes the great potential of prototype development as a methodology for the support of decision making and business transformation, especially around the complex transition process from reductionism (based on local optimization) to holism (based on global optimization) in supply chains.

RESUMEN

La globalización ha revolucionado el contexto empresarial. El incremento de la oferta en bienes y servicios, los constantes cambios en los gustos de los consumidores y la expansión geográfica de las redes de distribución, entre otros factores, han trazado un nuevo entorno competitivo marcado por la intensidad, la complejidad y el dinamismo. Éste ha enfatizado el concepto de cadena de suministro. En los procesos, en las relaciones y en las interdependencias de la cadena de suministro se esconde una fuente clave de ventajas competitivas para las organizaciones que, sin embargo, es muy compleja de captar. Una de las razones de ello es la generación del denominado Efecto Bullwhip, que ha de entenderse como una fuente clave de ineficiencias en la cadena de suministro. Este fenómeno se refiere a la amplificación de la variabilidad de las órdenes transmitidas a lo largo del sistema.

Los capítulos 1 a 3 del presente trabajo exploran el papel de la inteligencia artificial en el desarrollo de mecanismos de previsión orientados a mejorar la gestión de la cadena de suministro. Se han utilizado redes neuronales artificiales (artificial neural networks, ANNs), bajo arquitecturas del tipo perceptron multi-capa (multi-layer perceptron, MLP) y funciones de base radial (RBF), junto a métodos estadístico dentro de una estructura multi-agente. Ante demandas con tendencia y estacionalidad, el sistema —que escoge en cada momento la previsión más adecuada— obtiene un gran rendimiento en la reducción del Efecto Bullwhip desde una perspectiva local. Asimismo, se muestra cómo este sistema se podría integrar con facilidad en un sistema de mayor alcance, lo cual representa una de las principales ventajas de esta aproximación.

Los capítulos 4 a 6, que representan la principal línea de investigación dentro de este trabajo, tratan esta problemática desde una perspectiva sistémica. En este sentido, se pretende contribuir al despliegue de esta perspectiva dentro de las cadenas de suministro; el cual entendemos como el gran reto de las cadenas de suministro en el siglo XXI. Con este objetivo, se desarrolla un marco integrador para la gestión colaborativa de sistemas de producción y distribución basado en el Modelo de los Sistemas Viables de Beer (Viable System Model, VSM) y la Teoría de las Restricciones de Goldratt (Theory of Constraints, TOC). Sobre este marco, se explora la implementación de la solución mediante herramientas de modelado y simulación. Más en concreto, se utiliza la metodología Drum-Buffer-Rope (DBR) para proponer un motor operativo para la cadena de suministro y demuestra su eficacia, en comparación con alternativas tradicionales *basadas en la producción en masa, tanto en términos operacionales (donde se engloba el Efecto Bullwhip) como en términos económicos.*

No obstante, el trabajo subraya que la integración de procesos es sólo una de las áreas clave para el diseño de soluciones colaborativas. La transparencia en la información relevante, la sincronización y distribución en la toma de decisiones, y el diseño de un sistema de rendimiento global han de entenderse igualmente como condiciones sine qua non para la implementación exitosa de la colaboración en las cadenas de suministro. La alineación de incentivos también es esencial. Los riesgos y los beneficios han de ser compartidos adecuadamente con el objetivo de reducir la amenaza de comportamientos oportunistas. Los cinco campos mencionados se han considerado en la propuesta de una solución colaborativa viable y beneficiosa para todos los miembros; dado que este esquema nos permite comprender por qué solo un pequeño porcentaje de las cadenas de suministro reales son capaces de crear valor a través de la colaboración.

Esta Tesis Doctoral también pretende resaltar las técnicas de modelado y simulación como poderosos laboratorios de ensayo para el estudio de grandes problemas organizacionales que serían complejos de estudiar de otra forma. Este hecho subraya el enorme potencial del desarrollo de prototipos como metodología para el apoyo a la toma de decisiones y la transformación empresarial, especialmente en torno al complejo proceso de transición de una aproximación reduccionista (basada en la optimización local) a una holista (basada en la optimización global) en la cadena de suministro.

INDEX

Index of Contents

CHAPTER 3
INTELLIGENT DECISION SUPPORT SYSTEM FOR REAL-TIME WATER DEMAND MANAGEMENT **81**

CHAPTER 4
SYSTEMIC APPROACH TO SUPPLY CHAIN MANAGEMENT THROUGH THE VIABLE SYSTEM MODEL AND THE THEORY OF CONSTRAINTS **113**

Index of Tables

CHAPTER 6
HOLISM VERSUS REDUCTIONISM IN SUPPLY CHAIN MANAGEMENT: AN ECONOMIC ANALYSIS

Index of Figures

INTRODUCTION

CONTEXT, OBJECTIVES, AND STRUCTURE

Context: The Bullwhip Effect in Supply Chains

The organizational environment dramatically evolved in the 1990s. The geopolitical restructuring that took place after the Cold War, the revolution of information and communication technologies, the decrease of transportation costs, and the liberalization of capital markets —among other reasons— have set up a modern global business scene. Globalization has drawn a new scenario of opportunities and threats, where competition has not only increased substantially but also become more complex and dynamic. In this sense, competition currently surpasses the firm level, covering the overall concept of supply chain. Under these circumstances, a premium has been placed upon supply chain management as a key source of competitive advantages.

This is a relatively new concept that encompasses managing all the relationships in the production and distribution system (Mentzer et al., 2001). Managers must now deal with distant —not only in geographical terms but also in cultural and administrative terms— suppliers, control convoluted worldwide supply networks with long and variable lead times, and be able to agilely react to the frequent changes in customer requirements. How they manage all these issues can, and does, make the difference. In this process, they will face a powerful enemy: the *Bullwhip Effect*.

This term —first coined by Procter & Gamble in the 1990s after discovering that the firm was suffering from it (Lee et al., 1997)[1]— refers to a dynamical phenomenon in supply chains that results in the tendency of the variability of orders to increase as they pass through the various echelons of a production and distribution system towards raw material suppliers (Disney and Lambrecht, 2008). The Bullwhip Effect, which has been shown to usually occur in almost every sector, is commonly measured through the ratio between the variance of —either production or purchase— orders issued and the variance of orders —demand— received (Wang and Disney, 2016). This counterintuitive phenomenon is illustrated in figure 1 for four real supply chains of different industries.

From that definition, it can be easily understood why the Bullwhip Effect has a strong negative impact on businesses (Towill et al., 2007). On the one hand, large swings in the

[1] Nonetheless, it should be noted that the first documentation of this intriguing phenomenon date back to the 1910s, also between Procter & Gamble and its wholesalers (Schisgall, 1981).

orders received seriously threaten the firms' ability to meet demand. Thus, companies have to invest in large safety stocks, which in turn amplifies other risks in the system, such as obsolescence. On the other hand, even larger swings in the orders issued create unstable production schedules, which trigger a wide range of unnecessary costs, such as extra capacity and overtime (labor) costs. Things are worse if we consider that Bullwhip also tends to increase lead times[2] (Disney and Lambrecht, 2008). Hence, this phenomenon undoubtedly reduces the financial performance of companies —Metters (1997) estimated that it entails an avoidable reduction between 10% and 30% on business profitability.

Under these circumstances, the strategic importance of the Bullwhip Effect has led to a large amount of research over the last two decades. However, this issue is still far from being solved. After analyzing a sample of 14,933 buyer-supplier dyad observations in different US industries, Isaksson and Seifert (2016) recently reported an average increase in the variability (measured through the coefficient of variation) of orders between echelons that equaled 90%, which illustrates the current prevalence of this phenomenon.

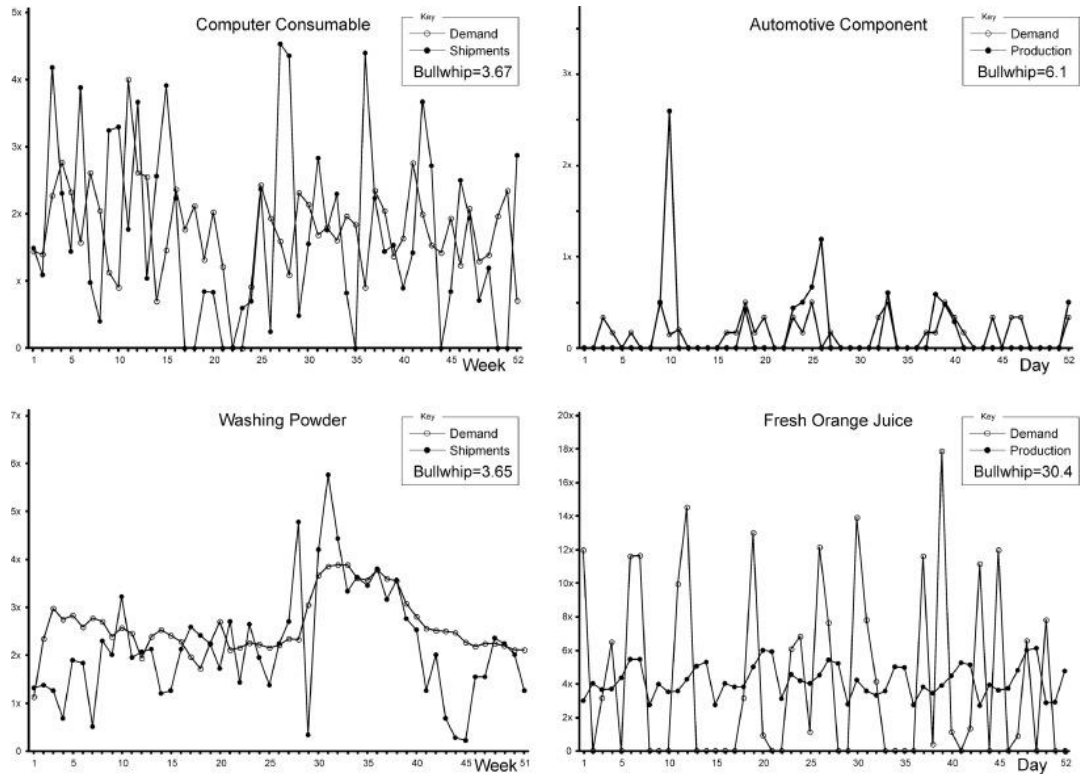


Figure 1. Empirical evidence of the Bullwhip phenomenon. Source: Wang and Disney (2016).

---

[2] It is widely accepted the injurious impact of long and variable lead times in organizations. Indeed, the innovative philosophy of Lean Production defines lean time reduction as one of the firms' key objectives.

To tackle this issue, the key question is, of course, why it occurs. Lee et al. (1997) pointed out to the information distortion throughout the supply chain as the underlying root cause of the Bullwhip Effect[3]. From this point on, several operational causes —understood as sources of information distortion in the system— can be identified, such as the processing of the demand signal (Forrester Effect)[4], order batching (Burbidge Effect)[5], the rationing and shortage gaming (Houlihan Effect)[6], and price fluctuations (Promotion Effect). It should be underlined that these operational causes interact with the so-called behavioral causes (Croson and Donohue, 2006), which also contribute to the information distortion. Decision makers repeatedly underweight the supply chain when making order decisions, which adds to the Bullwhip generation.

Together with these operational and behavioral causes, there are a large number of contextual factors that may significantly foster the amplification of the variability of orders in the supply chain. They can therefore be understood as effective action points to mitigate the Bullwhip phenomenon. In this regard, how the lead time —both in mean and in variance— impacts the Bullwhip Effect has been largely discussed in the literature; e.g. see Chen et al. (2000) and Disney et al. (2016). The structure of the supply chain is another contextual factor that contributes to the generation of the Bullwhip Effect, e.g. see Dominguez et al. (2015), as well as the production and distribution capacities, e.g. see

[3] Although the literature includes a number of works in the search of these causes—see Bhattacharya and Bandyopadhyay (2010) for a review—, we strongly concur with this view: the Bullwhip Effect can be (simply) understood as the consequence of the distortion on information transferred in the form of orders along the supply chain potentiated by the contextual factors mentioned below.

[4] This refers to the practice of decision makers adjusting the parameters of the replenishment rule, e.g. the demand forecasts and the target stocks levels. Forrester (1958) encountered this amplification in real-world supply chains and explored it via simulation.

[5] It involves the common practice of delivering purchase orders (or placing manufacturing orders) in batches with the aim of gaining economies of scale. Its negative impact on the supply chain was investigated by Burbidge (1994).

[6] Houlihan (1985) highlighted that, when shortages occur in supply chains, customers tend to over-load their orders. He noticed that it significantly contributed to the amplification of the variability of orders throughout the system.

Buchmeister et al. (2014). We could also cite several other factors that do not directly create this effect but can amplify or alleviate it, but we will focus on the most important one according to the huge amount of works involved around it. It is the attitude of the various supply chain nodes towards the management of the supply chain, defining two opposite approaches: local versus global optimization[7].

In the first scenario, the different members of the system aim to optimize their local performance, so the supply chain emerges as the result of the interaction of these individual strategies. On the contrary, the second context entails the search of the optimal strategy for the overall supply chain. In other words, the different supply chain nodes involve in a collaborative process aimed at designing an efficient, flexible, and robust system, from which companies can improve their individual performance (Simatupang and Sridharan, 2002). The nodes will then behave accordingly to the overall strategy, which come first. This distinction defines two different lenses from which analyzing the supply chain problem.

Given that supply chains are increasingly built on interdependences, its management is undeniably a field where thinking globally makes a big difference. The improvement generated by collaboration has been widely demonstrated both in theoretical studies and in practice, which emphasizes the relevance of this approach built on looking the supply chain in its entirety. Sterman (1989) showed that the interaction of self-centered decisions within production and distribution systems acts as a major source of inefficiencies derived from augmenting the information distortion the systems and the lack of coordination between processes, which undesirably increases the Bullwhip Effect. In addition, collaborative supply chain practices like the Vendor Managed Inventory (VMI) (Waller et al., 1999) and the Collaboration Planning, Forecasting, and Replenishment (CPFR) (Fliedner, 2003) have proven to generate breakthrough improvements in dealing with traditional supply chain issues, like the Bullwhip Effect; see Sari (2008).

However, this collaborative approach to supply chain management is still far from being widespread in practice. Meaningful barriers emerge, such as mistrust between partners,

[7] In this work, we will recall them as holism and reductionism from the two diametrically opposed strategies for problem solving, see chapter 6.

deficiencies in shared vision, and misalignments in organizational compatibility (Mentzer et al., 2001), which may lead the collaborative process to failure (Fawcett et al., 2015). This situation —together with the intrinsic importance of analyzing systems by carefully studying the different parts that form them— makes the investigation of the role of the single echelon within the supply chain also essential in order to comprehend the dynamics of supply chains; which also is a fruitful research area in the literature.

Objectives and Structure of this Research Project

This Doctoral Thesis explores the supply chain management field by means of artificial intelligence (AI)-based techniques. *In broad terms, we first aim to investigate how these modern approaches may contribute to the mitigation of the Bullwhip Effect phenomenon. From this point on, we use these techniques to design, develop and assess an integrative framework for managing the supply chain from a systemic perspective.* Hence, two research lines (RL) can be clearly identified within this research project, which are also related to the aforementioned lenses for the analysis of the supply chain.

On the one hand, we focus on the single echelon, understanding demand forecasting as a powerful mechanism for Bullwhip alleviation. Thus, *our objective in RL1 has been to evaluate the potential of AI-based tools in terms of supply chain management.*

We first have developed an AI forecasting system based on artificial neural networks (ANNs) with the aim of demonstrating its strength in the reduction of Bullwhip Effect (*chapter 1*). Then, the forecasting system has been enhanced by adding other forecasting techniques under a global structure in the form of a multi-agent system. The additional intelligence added to the system allows it to select at each moment the best forecast, which increases its performance (*chapter 2*). Last, we have explored how these agent-based techniques can be used to build an intelligent decision support system for achieving an efficient, flexible, and robust management of the supply chain (*chapter 3*).

It should be noted that these works have been carried out for a specific supply chain (i.e., water distribution). This results as a consequence of my engagement during the development of the present dissertation in the project "Distributed Artificial Intelligence for Managing Water Demand in the Municipality of Gijon", which was carried out in 2013 and 2014 together with the Municipal Water Company of Gijón (Empresa Municipal de Aguas) thanks to the funding that the Instituto Universitario de Tecnología

Industrial de Asturias (IUTA) awarded us. Nonetheless, the main conclusions derived from RL1 can be extrapolated to the generic field of supply chain management.

On the other hand, we consider the implementation of collaborative solutions throughout production and distribution systems, understanding its deployment as a major challenge for 21st-century supply chains (Schweitzer et al., 2009). Thus, *our objective in RL2 has been to propose an overall framework*[8] *for the implementation of collaborative supply chains.*

In this research stream, we first have devised a conceptual framework for supply chain collaboration and discuss its application in real supply chains (*chapter 4*). This is based on the Simatupang and Sridharan's (2005) integrative scheme and is built on the integration between the Viable System Model (VSM) (Beer, 1885) and the Theory of Constraints (TOC) (Goldratt, 1990). While the VSM orchestrates the collaboration (defines the systemic structure of the supply chain), TOC governs the material flow and defines the global performance metrics (implements and guides the systemic behavior).

From this point on, we have deepened the implementation of the systemic behavior of the supply chain —which encompasses process integration and decision synchronization— through the Drum-Buffer-Rope (DBR) methodology (Schragenheim and Ronen, 1990) (*chapter 5*). We have employed agent-based modeling and simulation techniques to explore its implementation in a serial four-echelon supply chain, and we have shown that it leads the system to a dramatic operational improvement, expressed in terms of Bullwhip Effect, in comparison with traditional (non-collaborative) alternatives.

Finally, we have completed this study by translating the operational study into a financial analysis (*chapter 6*). We have also augmented the noise conditions[9] of the supply chain in order to derive managerial implications in a wide range of scenario. In this sense, we have provided evidence of the superiority of the collaborative approach in financial terms.

---

[8] With the phrase "overall framework", we intend to highlight the importance of understanding "supply chain collaboration" in all its essential dimensions, which are detailed in chapter 4.

[9] We understand noise in experimental terms. Noise factors are parameters causing variation in the performance of the supply chain that are uncontrollable during the real operation of the system but can be controlled during the simulations.

This work has also shown how IA-based multi-agent techniques (Gilbert, 2008) allow practitioners to develop awareness of complex organizational problems, so these prototypes must be interpreted as powerful laboratories for business transformation.

Finally, it should be acknowledged that this Doctoral Thesis is presented as a compendium of publications, since it comprises six articles (i.e., chapters 1 to 6) published in journals indexed within the Journal Citation Report (JCR). The reference to each one of the publications as well as the required mention to the impact factor of each journal have been included as a footnote in the first page of each chapter. Please note that this different journal publication styles justify some slight differences in the format among the six chapters.

References

1. Beer, S. (1984). The viable system model: Its provenance, development, methodology and pathology. *Journal of the Operational Research Society*, 35(1), 7-25.
2. Bhattacharya, R., & Bandyopadhyay, S. (2011). A review of the causes of bullwhip effect in a supply chain. *International Journal of Advanced Manufacturing Technology*, 54(9-12), 1245-1261.
3. Buchmeister, B., Friscic, D., & Palcic, I. (2014). Bullwhip effect study in a constrained supply chain. *Procedia Engineering*, 69, 63-71.
4. Burbidge, J. L. (1994). The use of period batch control (PBC) in the implosive industries. *Production Planning & Control*, 5(1), 97-102.
5. 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.
6. Croson, R., & Donohue, K. (2006). Behavioral causes of the bullwhip effect and the observed value of inventory information. *Management Science*, 52(3), 323-336.
7. Disney, S. M. & Lambrecht, M.R. (2008). On replenishment rules, forecasting and the bullwhip effect in supply chains. *Foundations and Trends in Technology, Information and Operations Management*, 2(1), 1–80.
8. Disney, S. M., Maltz, A., Wang, X., & Warburton, R. D. (2016). Inventory management for stochastic lead times with order crossovers. *European Journal of Operational Research*, 248(2), 473-486.

9. Dominguez, R., Cannella, S., & Framinan, J. M. (2015). The impact of the supply chain structure on bullwhip effect. *Applied Mathematical Modelling*, 39(23), 7309-7325.

10. Fawcett, S. E., McCarter, M. W., Fawcett, A. M., Webb, G. S., & Magnan, G. M. (2015). Why supply chain collaboration fails: the socio-structural view of resistance to relational strategies. *Supply Chain Management: An International Journal*, 20(6), 648-663.

11. Fliedner, G. (2003). CPFR: an emerging supply chain tool. I*ndustrial Management & data systems*, 103(1), 14-21.

12. Forrester, J. W. (1958). Industrial dynamics: a major breakthrough for decision makers. *Harvard Business Review*, 36(4), 37-66.

13. Gilbert, G. N. (2008). *Agent-based models*. London: Sage.

14. Goldratt, E. M. (1990). *Theory of Constraints*. Croton-on-Hudson: North River.

15. Houlihan, J. B. (1985). International supply chain management. *International Journal of Physical Distribution & Materials Management*, 15(1), 22-38.

16. Isaksson, O. H., & Seifert, R. W. (2016). Quantifying the bullwhip effect using two-echelon data: A cross-industry empirical investigation. *International Journal of Production Economics*, 171, 311-320.

17. Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information distortion in a supply chain: The bullwhip effect. *Management Science*, 43(4), 546-558.

18. Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., & Zacharia, Z. G. (2001). Defining supply chain management. *Journal of Business Logistics*, *22*(2), 1-25.

19. Metters, R. (1997). Quantifying the bullwhip effect in supply chains. *Journal of Operations Management*, 15(2), 89-100.

20. Sari, K. (2008). On the benefits of CPFR and VMI: A comparative simulation study. *International Journal of Production Economics*, 113(2), 575-586.

21. Schisgall, O. (1981). *Eyes on tomorrow: the evolution of Procter & Gamble*. New York: J.G. Ferguson.

22. Schragenheim, E., & Ronen, B. (1990). Drum-buffer-rope shop floor control. *Production and Inventory Management Journal*, 31(3), 18-22.

23. Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A., & White, D. R. (2009). Economic Networks: The New Challenges. *Science*, 325 (5939), 422-425.

24. Simatupang, T. M., & Sridharan, R. (2002). The collaborative supply chain. *International Journal of Logistics Management*, 13(1), 15-30.

25. Simatupang, T. M., & Sridharan, R. (2005). An integrative framework for supply chain collaboration. *International Journal of Logistics Management*, 16(2), 257-274.

26. Sterman, J. D. (1989). Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Science*,35(3), 321-339.

27. Towill, D. R., Zhou, L., & Disney, S. M. (2007). Reducing the bullwhip effect: Looking through the appropriate lens. *International Journal of Production Economics*, 108(1), 444-453.

28. Waller, M., Johnson, M. E., & Davis, T. (1999). Vendor-managed inventory in the retail supply chain. *Journal of Business Logistics*, 20(1), 183.

29. Wang, X., & Disney, S. M. (2016). The bullwhip effect: Progress, trends and directions. *European Journal of Operational Research*, 250(3), 691-701.

CHAPTER 1

THE BULLWHIP EFFECT IN WATER DEMAND MANAGEMENT: TAMING IT THROUGH AN ARTIFICIAL NEURAL NETWORKS-BASED SYSTEM[♦]

[♦] This article was published in *Journal of Water Supply: Research and Technology—AQUA* (IWA Publishing), vol. 64, no. 3, pp. 290-301 in May 2015. DOI: 10.2166/aqua.2015.087. Authors (in order): Borja Ponte, Laura Ruano, Raúl Pino, David de la Fuente.

Journal of Water Supply: Research and Technology—AQUA, indexed in the ISI, is a Q3-journal in the category 'Engineering: Civil' (Science Edition). Impact Factor 2015: 0.807.

Abstract

The Bullwhip effect (BE) refers to the amplification of the variance of orders and inventories along the supply chain as they move away from the customer. This is considered as the main cause of inefficiencies in the management of a traditional supply chain. However, the BE is not relevant in the classic system of water distribution, based on long-term supply management. Nevertheless, current circumstances have drawn a new context, which has introduced the concept of water demand management, in which efficiency and sustainability are of great importance. Then, the time horizon of management has decreased enormously and the supply time takes on an important role. Therefore, the BE must be considered, as it significantly raises the costs of management. On the one hand, this paper brings evidence that the BE appears in a system of real-time management of water demand. On the other hand, it proposes the application of artificial intelligence techniques for its reduction. More specifically, an advanced forecasting system based on artificial neural networks has been used. The BE is heavily damped.

Keywords

Artificial Neural Networks; Bullwhip Effect; Water Demand Management.

1. Introduction

The concept of Bullwhip Effect emerged in the early 90s in some large companies, when the new competitive context conceded strategic importance to Supply Chain Management (SCM). Some businesses began to understand SCM as a source of competitive advantages and studied it in detail, trying to optimize its performance. At that time, Procter & Gamble realized that the purchase orders received in one of its flagship products, Pampers diapers, fluctuated significantly, while the product demand in the retailer was almost constant. They also found out that the variability in orders transmitted to their suppliers were much higher. It was called the Bullwhip Effect (Lee et al., 1997).

The growing importance of logistics in the doubtful environment currently faced by businesses has prompted the development of this concept, which is considered to be the main cause of inefficiencies in SCM (Disney et al., 2005). For this reason, lots of various supply chains have focused on reducing the Bullwhip Effect, with the aim of minimizing the derivatives overruns. By contrast, in some particular supply chains, this phenomenon has not been relevant and it has not been widely studied. The water supply system is one of them.

Nevertheless, the perspective of municipal policies about water management has changed significantly over the last two decades, mainly due to the pressures generated by the population growth and the industrialization. Hence the concept of Water Demand Management (WDM) has developed significantly. Brooks (2006) proposed a current definition of WDM with five components: (1) reducing the quantity or quality of water required to accomplish a specific task; (2) adjusting the nature of the task so it can be accomplished with less water or lower quality water; (3) reducing losses in movement from source through use to disposal; (4) shifting time of use to off-peak periods; and (5) increasing the ability of the system to operate during droug*hts*

This burgeoning concern over efficiency and sustainability around WDM (Charlesworth and Adeyeye, 2013) has led to a reduction in the time horizon. Some years ago, long term forecasting was enough for the design of the system and the development of plans (among others, Willsie and Prat, 1974). However, nowadays, short term forecasts are required for attaining high efficiency in operation and management (among others, Gato et al., 2007). Herrera et al. (2010) defend that the ready availability of hourly predictions of water demand is crucial due to three main reasons: (1) it allows to determine the optimal regulation and pumping systems to meet the predicted demand, which promotes energy efficiency (operative point of view); (2) it allows to combine water sources in the most appropriate way to achieve a preset standard in the supply water (quality point of view); and (3) it allows to detect failures and network losses through the comparison of the actual and expected flow (vulnerability point of view). It *can be called real-time WDM.*

In a long term WDM system, the Bullwhip Effect does not arise. If the time horizon is very long, the supply time becomes trivial and does not determine the performance of the replenishment policy. However, reducing this time horizon introduces in the study the need to consider the supply time, and therefore the menace of Bullwhip Effect surges. It must be taken into account in order to avoid the negative consequences that it can have on the supply system. Thereby, one of the main objectives of this paper is to bring evidence via simulation of the appearance of the Bull*whip Effect in a real-time WDM.*

Furthermore, this work proposes a solution to the identified problem, based on the application of Artificial Intelligence techniques in forecasting the hourly water demand. More specifically, an advanced forecasting system, whose core are Artificial Neural Network (ANNs), has been developed. This methodology has been widely used in the forecasting of series of a similar nature, as the short-term electricity load (see Hippert et al., 2001, for a review). Herrera et al. (2010) showed that predictive models, among which ANNs are included, provide great performance in forecasting the hourly water consumption. This research has tried to reduce the error even further by developing a double-loop system that chooses at all times the optimal network structure (both input variables and hidden neurons). Therefore, the second goal of this paper is to demonstrate that these smart tools can cause a large decrease in the Bullwhip Effect generated in the water distribution system and, consequently, it can lead to improve the management.

2. Background: The Bullwhip Effect in Supply Chains

Although research on the Bullwhip Effect was strengthened two decades ago when large companies experimented the problem, Forrester (1961) long before noted the amplification of demand variability along a generic supply chain through a simulation model. Thereby, many authors express mathematically the Bullwhip Effect generated at level *n* of a linear supply chain (BE^n) as the quotient of the variance of the orders issued to the upper level supply chain (${\sigma^2_{POE}}^n$) and the orders received from the lower level of the same (${\sigma^2_{POR}}^n$). As this metric only evaluates the output variance compared with the input variance, it should be supplemented by another one that provides the variation in the level of inventories (i.e., the structure that causes the above variation). Therefore, some authors (e.g., Disney and Towill, 2003) propose an alternative metric of the quotient of the variance of the stock (${\sigma^2_{STOCK}}^n$) and the variance of the demand (${\sigma^2_{POR}}^n$). It can be named Alternative Bullwhip Effect (ABE^n) and is expressed by (2).

$$BE^n = \frac{{\sigma^2_{POE}}^n / {\mu_{POE}}^n}{{\sigma^2_{POR}}^n / {\mu_{POR}}^n} = \frac{{\sigma^2_{POE}}^n}{{\sigma^2_{POR}}^n} \tag{1}$$

$$ABE^n = \frac{{\sigma^2_{STOCK}}^n}{{\sigma^2_{POR}}^n} \tag{2}$$

The Bullwhip Effect involves large economic losses in the supply chain, by increasing missing sales, obsolescence, and labor, transportation and storage costs, so it can be considered a major cause of inefficiencies within SCM (Disney et al., 2005).

Lee et al. (1997) showed that there are five main causes that lead to this phenomenon: (1) errors in demand forecasting; (2) non-zero lead times; (3) order batching; (4) price fluctuations; and (5) supply shortages. The famous 'Beer Game' proposed by the MIT and analyzed by Sterman (1989) brings evidence that the Bullwhip Effect is generated along the supply chain even if the last three causes are not considered. Obviously, if lead time was null, the supply from the factory would instantly respond to customer requirements and the Bullwhip Effect would not appear. And if there were no errors in the forecasting, each level would know exactly what it needs, so the Bullwhip Effect would not surge either.

SCM is a very complex problem, which is conditioned by the interaction of multiple agents, each one of which has to weight a large number of variables. Thus, modern Artificial Intelligence tools have been widely used in order to optimize the management and to buffer the Bullwhip Effect. Next, a brief literature review on this subject is shown. In the beginning, the Metamorph tool, based on multi-agent methodology and developed by Maturana et al. (1999), can be highlighted. In 2010, Hong et al. designed an ANNs based controller and using RFID technology. Jaipuria and Mahapatra (2014) developed an advanced forecasting system (ANNs and Wavelet Discrete Transform) to reduce the

Bullwhip Effect in a generic supply chain. Also, the recent and relevant works carried out by Bahroun et al. (2010), Saberi et al. (2012) and Zarandi et al. (2013) should be mentioned.

3. The Bullwhip Effect in Real-Time Water Demand Management

The main hypothesis of this work is that the Bullwhip Effect appears in real-time WDM systems, and therefore it must be controlled due to the consequences that it could bring to the system.

Under these conditions, the Bullwhip Effect in a water supply network is the increasing variability of the demand transmitted along the same as it moves away from the final points of consumption. This phenomenon directly causes the increase of the variations in the water flow conveyed along the distribution network and also in the increase of the variations in the water stored in the supply tanks. Therefore, it tends to oversize the system (distribution network, supply tanks and treatment equipments), although the infrastructure oversize is more influenced by other reasons –reliability and security against unforeseen, but possible, events. Moreover, the Bullwhip Effect also generates cost overruns in the works of water pumping, collection and purification, as the contracted power is greater when the variability of the system requirements over time is large. Hence, taming the Bullwhip Effect leads to improvements of the management.

3.1. Simulation model

In order to demonstrate the generation of the Bullwhip Effect along a real-time WDM system, this research has considered a simple structure of a water supply network, which consists of three main levels interconnected by the distribution piping: (1) natural sources (catchment points), where water is collected; (2) points-of-use (POU), representing the distributed water demand; and (3) supply tanks (storage reservoirs), which receive water from the natural sources and send it to the POU. Then, a discrete simulation model has been developed in MATLAB R2014a of a supply system managed hourly, focused on the supply tanks.

Other assumptions adopted to model the supply system are the following: (1) stochastic POU demand (see section of results, as the same time series has been used); (2) fixed supply time: 1 hour (on the one hand, from natural sources to supply tanks and, on the other, from supply tanks to POU); (3) unconstrained catchment, storage and transportation system; (4) water is pumped to the supply tanks in order for them to store at the beginning of each hour—order-up-to point—the forecast plus a security level, with the aim of protecting against shortage; and (5) non negative condition of the order quantity (water cannot be returned to the previous level). Obviously, it is a simplified model of the reality, but it considers the main causes that surge the Bullwhip Effect in real-time WDM systems.

Next, the mathematical formulation of the model is described. Water pumped at the end of each hour from natural sources to supply tanks (WP_t) can be expressed as the difference between the demand forecast for the next period ($\widehat{D_{t+1}}$) and the water stored in the tanks at the end of that period (WT_t), also considering security level which must be kept in the tank (SL), by (3). Along the same line, the water stored in the tanks at the end of each period (WT_t) is the water stored in the tanks at the end of the previous period (WT_{t-1}), adjusted by the water pumped in the previous period from natural sources (WP_{t-1})—as the lead time is 1 hour—and by the demand (D_t), unless this difference is less than 0, according to (4). In that case, it is not possible to meet all the demand, and a deficit of unmet demand (UMD_t) is generated, by (5). Furthermore, logically, the water sent from the supply tanks to the POU (WS_t) is the demand (D_t), unless the water stored at the end of the previous period (WT_{t-1}) was lower, according to (6).

$$WP_t = \max\{\widehat{D_{t+1}} - WT_t + SL, 0\} \tag{3}$$

$$WT_t = \max\{WT_{t-1} - D_t + WP_{t-1}, 0\} \tag{4}$$

$$UMD_t = \max\{-(WT_{t-1} - D_t + WP_{t-1}), 0\} \tag{5}$$

$$WS_t = \min\{D_t, W_{t-1}\} \tag{6}$$

The operational logic of the simulation system is illustrated in figure 1. As above mentioned, it is based on the supply tanks—and the Bullwhip Effect can be observed when comparing the demand transmitted from POU to supply tanks and from supply tanks to natural sources. The system is controlled by the user through an interface, and it is connected to a database with the aim of storing and analyzing the results. It should be noted that there are two main flows: the water flow, from natural sources to POU and constrained by the lead time (supply time), and the order flow, in the opposite direction. The flow chart of the operations in the supply tanks corresponds to the previous equations.

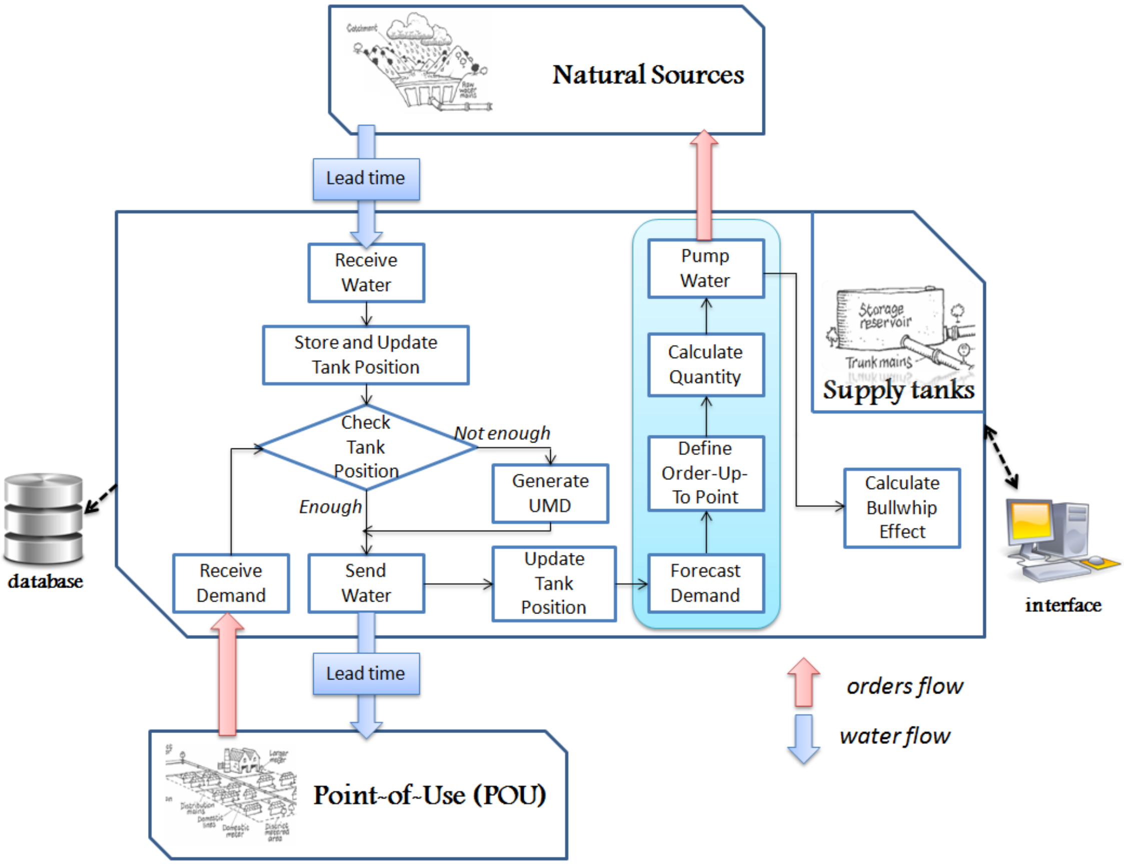


Figure 1. Outline of the simulation system.

3.2. Simulation results

In order to calculate the forecast for the next time period ($\widehat{D_{t+1}}$), moving averages (Holt, 2004) of 3 and 6 periods and simple exponential smoothing (Gardner, 2006) with coefficients 0.5 and 0.9 have been used. Additionally, three different tests with each forecasting method have been carried out, as the value of the security level with which the tanks works has also been modified. Table 1 shows the results of the twelve simulations using in all cases the same week (randomly chosen) of the time series.

Table 1 demonstrates the generation of Bullwhip Effect in the twelve tests (since the ratio is greater than 1 in all cases), in which different forecasting methods and security levels have been used. In the best situation (test 4), the amplification of the variance of the demand is 9%. Although not included for the sake of simplicity, tests carried out with changes in the supply time or the pumping policy also evidence the existence of this phenomenon. Thereby, in this real-time WDM system, there is amplification in the variability of the demand.

Table 1. Results of the simulation.

Test	*FM*	*SL*	*BE*	*ABE*	*UMD*
1	MA3	200	1.19	0.18	5,730
2	MA3	400	1.26	0.27	744
3	MA3	600	1.27	0.29	0
4	MA6	200	1.09	0.29	13,336
5	MA6	400	1.19	0.45	4,656
6	MA6	600	1.28	0.58	515
7	ES0.5	200	1.17	0.16	4,439
8	ES0.5	400	1.23	0.23	269
9	ES0.5	600	1.23	0.24	0
10	ES0.9	200	1.19	0.10	1,229
11	ES0.9	400	1.20	0.11	0
12	ES0.9	600	1.19	0.11	0

Note: The columns refer to the number of the test (Test), the forecasting method (FM), the security level of the tanks in cubic meters (SL), the Bullwhip Effect (BE), the Alternative Bullwhip Effect (ABE), and the unmet demand in cubic meters (UMD).

The results presented in table 1 show a straightforward (and easy to understand) relationship: the higher the security level, the lower the unmet demand. Furthermore, it brings evidence that the higher the security level, the higher the variations along the system, which typically results in an increase of the Bullwhip Effect.

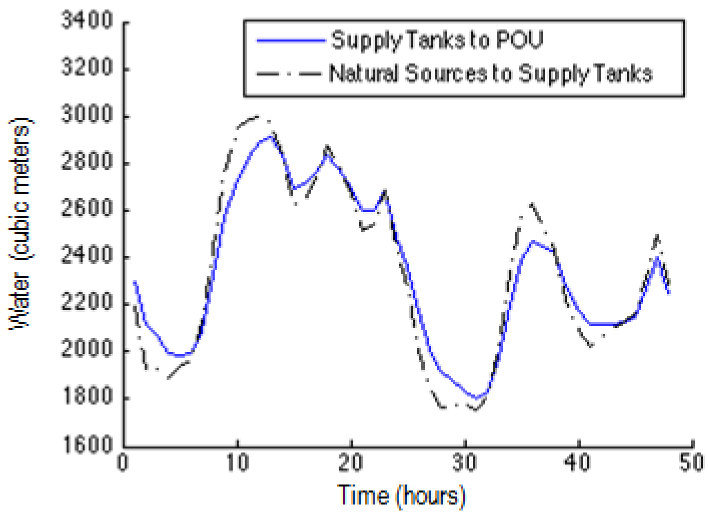


Figure 2. Water received and sent by the supply tanks for 48 hours (corresponding to Friday and Saturday) in test 3.

The Bullwhip Effect generation on the water supply network, by way of example, can be seen graphically in figure 2, which represents the water conveyed between supply tanks and POU and between natural resources and supply tanks for two days of test 3. In it, the amplification of the variance is 27%. Figure 3 displays, for the same time period, the volume of water in the supply tanks in tests 1, 2 and 3. These variations produce the magnification of the Alternative Bullwhip Effect when the security level increases, although the unmet demand obviously decreases. Thereby, the consequences of the Bullwhip Effect in the WDM system are evidenced.

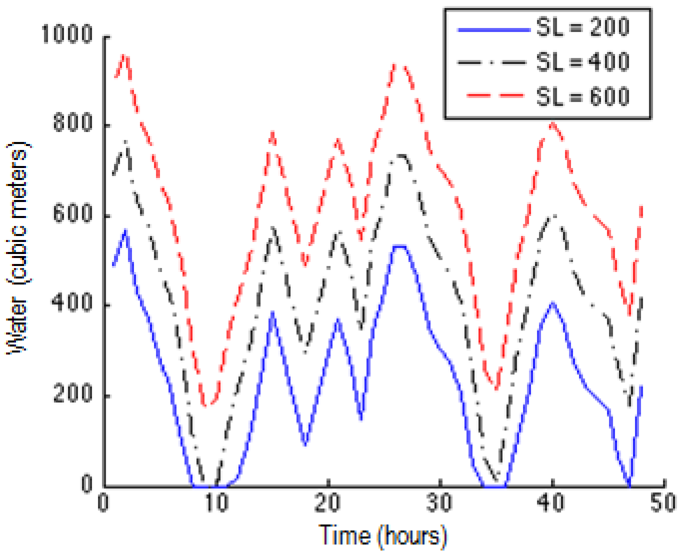


Figure 3. Volume of water in the supply tanks for 48 hours (corresponding to Friday and Saturday) in tests 1, 2 y 3.

4. Description of the Forecasting System

The forecasting errors are the main cause of the Bullwhip Effect. Hence a system based on an Artificial Neural Networks (ANNs) structure has been developed to forecast the hourly demand with the aim of minimizing the errors. The results will be evaluated by comparing them with the ones provided by statistical methods, which will be detailed afterwards.

4.1. ANNs Forecasting System

ANNs are computational models inspired by an animal's central nervous system, which are capable of machine learning, as well as pattern recognition. They are systems of interconnected neurons, distributed in different layers, which can compute values from inputs. Two characteristics of ANNs that make them particularly useful for forecasting time series are the ability to approximate practically any function (even non-linear ones) and the opportunity for “piece-wise” approximations of the functions. For a more detailed description of ANNs as a forecasting method and its contrast with other traditional tools, see Pino et al. (2008).

In particular, the model used for this study is the nonlinear autoregressive network with exogenous inputs (NARX), where the next value of the dependent output signal is forecast ($\hat{y}(t) = \widehat{D_t}$) as a regression on previous values of the output signal ($y(t) = D_t$) and previous values of an independent (exogenous) input signal ($x_t = x(t)$). The NARX model is developed, among others, in the work of Piroddi and Spinelly (2003). The software that has been used is MATLAB R2014a.

Figure 4 shows the architecture of the forecasting system—it is called Multi-Layer Perceptron (MLP). From a set of inputs, the system is capable of building a response. In particular, the program takes not only the previous demands, but also the hour (ranged from 00h to 23h), the week day (from 1, corresponding to Mondays, until 7, corresponding to Sundays) and an extra variable, related to the main feature of the day, which differences working days (1), Saturdays (2) and Sundays and holidays (3)—due to the nature of this time series: from Monday to Friday, consumption of water remains pretty similar, while it decreases on Saturday, and keeps falling on Sundays (holidays can be approximated to Sundays).

MLP are networks that have more than one layer of adaptive weights (Bishop, 1995). It has three layers of units taking values in the range 0-1, and each layer is nourished with the previous ones. Any number of weighted connections can be used, but MLPs with two weighted connections are very much capable of approximation just about any functional mapping. The MLP can be mathematically represented by (7), where $\mathrm{y_t}$ represents the output (forecast), $\mathrm{f_{outer}}$ represents de output layer, $\mathrm{f_{inner}}$ represents the input layer transfer function, w_{xy} represents the weights and biases ($i \in [1, (3m+3)]$ refers to the input neurons and $j \in [1, n]$ refers to the hidden neurons) and *(z)* represents the z-th layer.

$$\widehat{D}_t = y_t = f_{outer}\left[\sum_{j=1}^{n} w_{1j}{}^{(2)} \cdot f_{inner}\left(\sum_{i=1}^{3m+3} w_{ji}{}^{(1)} \cdot x_i + w_{j0}{}^{(1)}\right) + w_{10}{}^{(2)}\right] \tag{7}$$

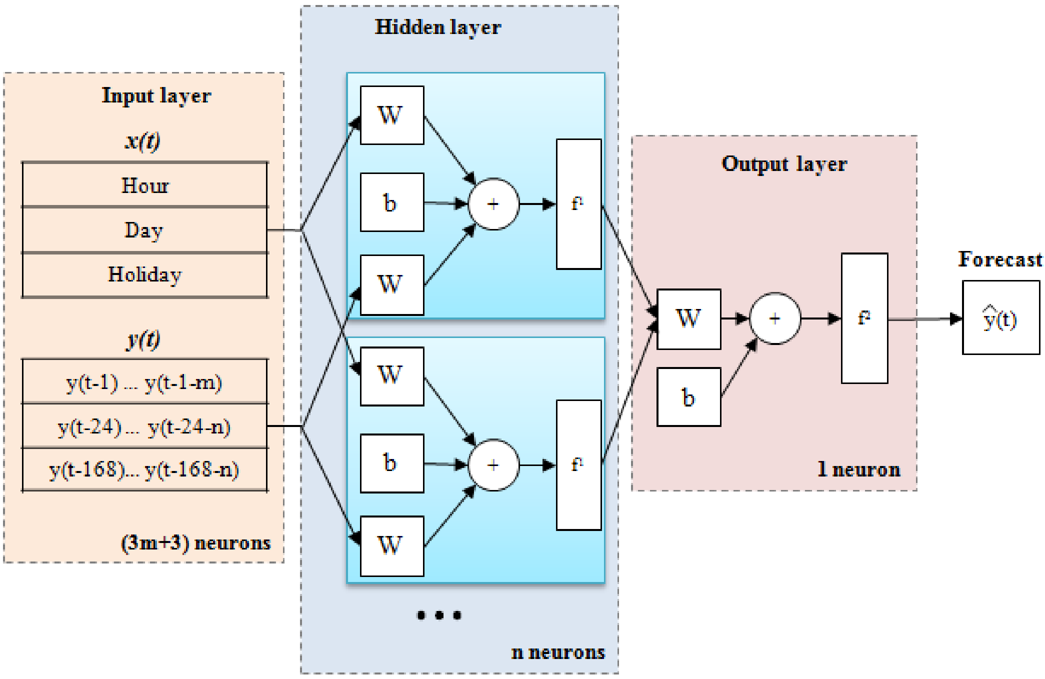


Figure 4. ANNs architecture of the forecasting system.

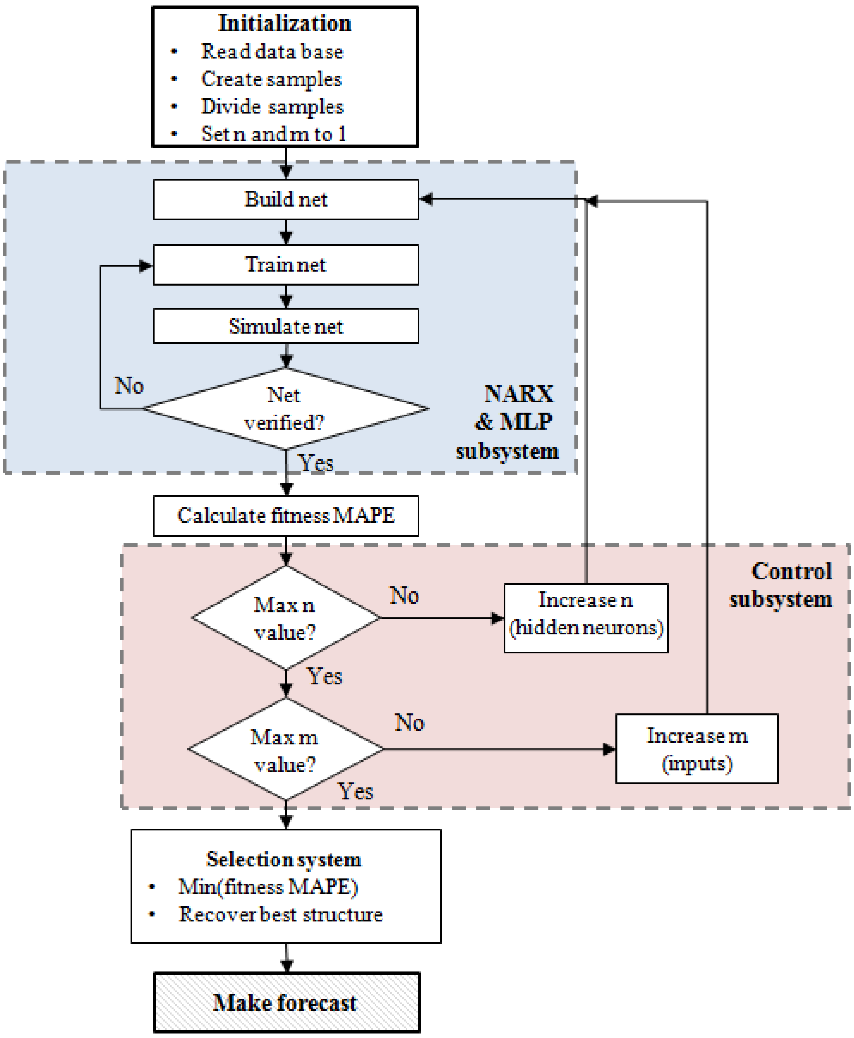


Figure 5. Flow chart followed by the Forecasting System to make the prediction.

Figure 5 points out a brief explanation of the structure and operation of the ANNs forecasting system. It makes an hourly forecast, when it receives the last demand from the measurement equipment and the information is stored in the database. Then, it reads the database and selects the last 1,008 samples, which correspond to an entire period of 6 weeks (the hourly demands of 42 days). Samples are randomly divided (except last 12) into three different groups: 70% of them are classified as training data, for adjusting the network according to its error; 15% as validation data, used to measure network performance and to halt training when it stops improving; and the remaining 15% as testing data, which provides an independent measure of network performance during and after training.

The used training function updates weight and bias values according to Levenberg-Marquardt optimization, which uses this approximation to the Hessian matrix in a Newton-like update (see Moré, 1978). In order to verify the training of the ANNs and to avoid overfitting, the early-stopping method (Sarle, 1995) has been used, as the number of training examples is sufficiently large. It presents interesting advantages in terms of speed and ease of application in comparison with cross-validation (Kohavi, 1995), which is much more suitable when the number of examples is low. Training stops when any of these conditions occurs: the maximum number of repetitions (100) is reached; the maximum amount of time is exceeded (10 minutes); the performance gradient falls below the value defined (10^{-10}); validation performance has increased more than the times defined (6) since the last time it decreased; or the scalar value exceeds its maximum value (10^{10}).

In the search of the structure that fits best the time series, two things are varied by the control subsystem: the number of neurons in the hidden layer and the number of delays (hence the number of variables that are considered to forecast). Therefore, the system chooses at each time the optimal structure of the network, seeking for a better performance of the tool than if the same structure was always imposed.

About the first loop, it should be kept in mind that the more number of hidden neurons (n) are chosen, the more complexity the structure will have, and it requires a higher time of carrying out. However, this does not always translate into a better outcome of the system because overfitting problems are more prone to occur if there are too many neurons in the hidden layer. The system builds a network for each one of the different values of the variable from $n = 2$ to $n = 20$, with jumps of 2 units.

In addition, another loop is made in order to look for the combination of delays that best fits the input data. As the hourly consumption time series shows trend and double periodicity, the best way of defining a new value for the curve of water consumption is by choosing the demands of the previous hours (from $y(t-1)$ to $y(t-1-m)$), the demands of the previous day at the same hour and the previous ones (from $y(t-24)$ to $y(t-24-m)$), and the demands of the previous week at the same day hour and the previous ones (from $y(t-168)$ to $y(t-168-m)$. The system evaluates alternatives from $m = 1$ to $m = 8$.

To evaluate the performance of the forecasting system, the criterion of the mean absolute percentage error (MAPE), introduced by Makridakis (1993), is used. It can be expressed by (8), where p is the time horizon:

$$MAPE = \frac{1}{p}\sum_{t=1}^{p}\left|\frac{D_t - \widehat{D}_t}{D_t}\right| \tag{8}$$

The last twelve demands are saved as testing samples, in order to orientate the double-loop to determine the optimal ANNs architecture at every moment. Therefore, after each iteration, the system calculates the MAPE of the last twelve demands as an indicator of network performance in recent hours (fitness MAPE). Once the double-loop process of building networks has ended, the system chooses the structure that has generated a minimum fitness MAPE and new predictions are made with this architecture.

4.2. Statistical models

Traditional methods are used to compare their results with the developed system, and to show the improvements on the Bullwhip Effect reduction. Three statistical techniques have been used. The system chooses the best of the three at any time using the same criterion (fitness MAPE minimization).

First, an autoregressive model (AR) is used (Akaike, 1969). The algorithm for computing the least squares AR model is the forward-backward approach, which minimizes the sum of a least squares criterion for a time-reversed model. The second model corresponds to

IVAR, which estimates the AR model using the instrumental variable method (Arellano and Bover, 1995). Both algorithms treat noise differently. AR assumes white noise, while the IVAR is not sensitive to noise color. The third one corresponds to the ARMA model (Jones, 1980). It includes a moving average component to consider the relation of the series with past values of the errors.

5. Results and Discussion

In order to evaluate the effectiveness of the forecasting system in the Bullwhip Effect reduction, a simulated time series with the hourly water demand in 2009 and 2010 in Gijón (a municipality of 300,000 inhabitants in the north of Spain) has been used. Validated by the municipal water company—real data are not available as this company still do not carry out an hourly management—, this series was created through the monthly water demand of the city, a distribution model of hourly water demand for a city in south-eastern Spain (Herrera et al., 2010), and random parameters. The information obtained from the literature was used to create a consumption modulation curve describing the behavior of the hourly water demand along the different days of the week. To adjust properly the vertical scale (in cubic meters) —and hence including the long-term trend of the series—, each month's water demand (known for 2009 and 2010) has been applied. This simulation was run for the above mentioned time horizon, adding random parameters with the aim of slightly modifying the curve at every moment and creating short-term trends in the series. Holidays have also been considered.

This way, the time series replicates a real hourly water demand series, which is a complex series with double seasonality and trend. On the one hand, it has a daily periodicity, as every 24 hours the series shows a similar structure. On the other hand, the consumption significantly varies on Saturdays and Sundays (and on holidays if there are), hence there is a weekly periodicity (168 hours). Moreover, the time series does not remain in a constant range, but it exhibits the above mentioned trends both in mean and variance.

In this study, different days and hours have been selected randomly with the aim of evaluating the performance of the system in different situations. Table 2 presents the different periods that have been chosen. In its last column, it differentiates between working days (1), Saturdays (2) and Sundays and holidays (3), according to the classification above mentioned.

Table 2. Training period (6 weeks) and testing periods (24 hours) of the eight tests that have been performed.

Test	*Training period*		*Testing period*		*Testing day*
	From	*To*	*From*	*To*	*Kind*
1	24/01/09	06/03/09	0h	23h	Saturday
2	28/07/09	08/09/09	5h	4h	Holiday
3	17/12/09	28/01/10	17h	16h	Working day
4	30/12/09	10/02/10	12h	11h	Working day
5	10/01/10	21/02/10	4h	3h	Holiday
6	22/01/10	05/03/10	21h	20h	Saturday
7	12/03/10	23/04/10	5h	4h	Working day
8	28/07/10	08/09/10	14h	13h	Holiday

The discrete simulation model described in the third section has been used to calculate de Bullwhip Effect and the Alternative Bullwhip Effect with the ANNs forecasting system in the eight tests. The chosen security level in the supply tanks is 500 cubic meters—this value was selected because there would not be unmet demand in none of the eight cases.

Table 3 depicts the final results obtained in this research. They point out, broadly speaking, the huge efficiency of the ANNs forecasting system versus the statistical methods in the reduction of the Bullwhip Effect. As expected, an improvement in the forecasting MAPE usually implies an improvement in both indicators of the Bullwhip Effect.

The ANNs forecasting system leads to the achievement of minor errors. By selecting at each time the best architecture of the network, forecasting errors around 1% are obtained in the tests performed, below those achieved by the traditional statistical methods. Thus, the Bullwhip Effect—that is evident and a major threat to the WDM system with the statistical models (the amplification varies between the 11% in test 2 and the 53% in test 5)—experiences a great reduction when using the ANNs system. In other words, this forecasting system makes the amplification of the variability of the demand along the supply network non-significant. Similarly, variations in the water volume at the supply tanks are largely reduced. This leads to conclude that the negative consequences of the Bullwhip Effect in the hourly-managed water distribution system are remarkably attenuated with the system that has been implemented.

Table 3. Results of the simulation.

Test	*Artificial Neural Networks*			
	Structure	*MAPE*	*BE*	*100ABE*
1	12-2-1	0.70%	0.98	0.59
2	12-2-1	0.98%	1.03	0.86
3	9-8-1	1.65%	1.00	6.71
4	9-10-1	0.58%	1.02	0.57
5	9-12-1	0.90%	1.00	0.99
6	9-12-1	0.76%	0.95	0.92
7	15-4-1	1.02%	1.00	0.76
8	12-2-1	1.03%	1.00	0.83

Test	*Statistical Methods*			*Reduction*		
	MAPE	*BE*	*100ABE*	*MAPE*	*BE*	*ABE*
1	2.64%	1.33	10.94	73.48%	26.32%	18.54
2	1.58%	1.11	2.38	37.97%	7.21%	2.77
3	2.09%	1.28	7.89	21.05%	21.88%	1.18
4	1.86%	1.29	7.73	68.82%	20.93%	13.56
5	3.38%	1.53	16.37	73.37%	34.64%	16.54
6	2.65%	1.34	12.65	71.32%	29.10%	13.75
7	2.73%	1.12	5.21	62.64%	10.71%	6.86
8	2.64%	1.33	10.94	60.98%	24.81%	13.18

Note: The columns contain: the MAPE of the forecasting (MAPE), the Bullwhip Effect generated in the distribution system (BE) and the Alternative Bullwhip Effect multiplied by 100 (ABE), both when the ANNs forecasting system is used and when the best statistical model is used to forecast. In addition, the comparison between both methodologies is displayed through the percentage reduction of MAPE and Bullwhip Effect and through the quotient between the Alternative Bullwhip Effect obtained in both cases. It also includes the ANNs structure used by the system to forecast the hourly demand in each test (Struct.).

Regarding the system's architecture, Table 3 brings evidence that there is not a direct relationship between the complexity of the network and the accuracy of their forecasts. For working days, in most cases, the system finds that the best architecture corresponds to the selection of the minimum value of m, so that the number of inputs is usually smaller (tests 3, 4 and 6) than in weekends and holidays. However, if the number of hidden neurons in each test is analyzed, it can be noted that weekends and holidays generally need fewer neurons in the hidden layer (tests 1, 2 and 8).

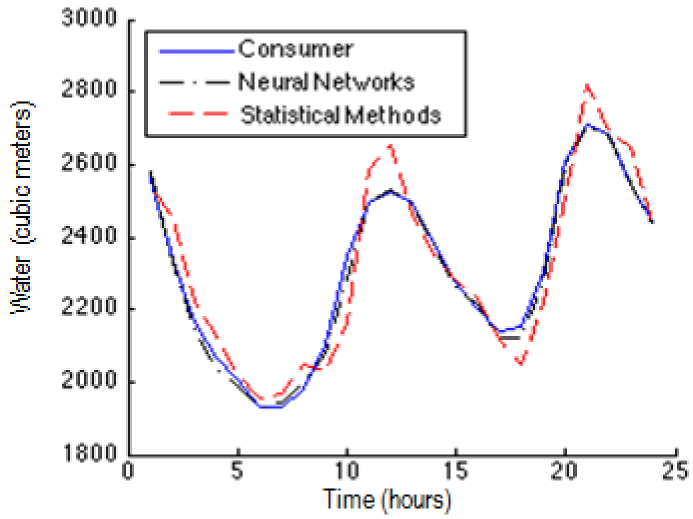


Figure 6. Differences between the demand and the two forecasts for test 1.

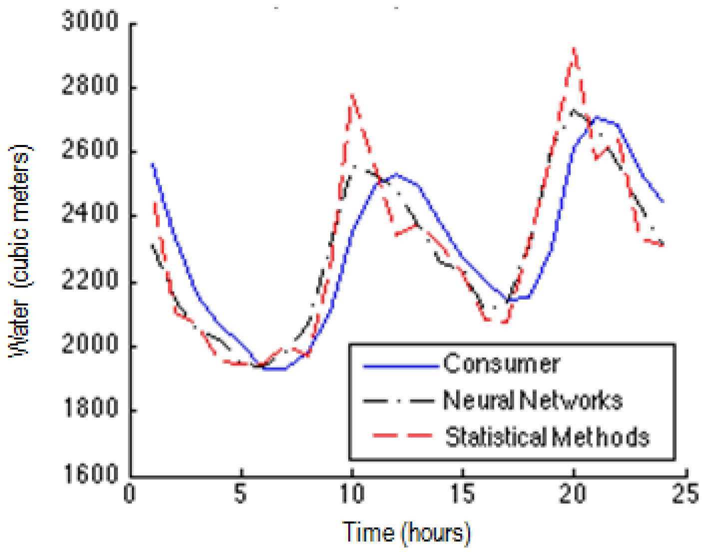


Figure 7. Variations of the real and the transmitted demand for test 1.

By way of example, test 1 is a clear example in which the results of the ANNs forecasting system significantly decreases the MAPE obtained with the statistical methods—and as a result the Bullwhip Effect is minimized. Figure 6 shows the real consumption and the two forecasts. The ANNs system (0.70% MAPE) offers better performance (2.64% MAPE of the best statistical model). The graph shows that it captures very accurately the periodicity and the trend of the consumption. Meanwhile, figure 7 displays the difference between the POU's consumption (from supply tanks to POU) and the transmitted demands (from natural sources to supply tanks)—one when ANNs are used and the other with statistical methods. It shows that the distortion introduced to the WDM system is much smaller with the ANNs, so that the tank requirements vary much less. The figure shows that the accuracy of the forecasting system causes that the water conveyed between natural resources and supply tanks approximates closely to the POU's consumption, but displaced—the supply time is the time difference between them. Thus, the Bullwhip Effect is greatly reduced.

6. Conclusions and Future Research Lines

In this paper, the Bullwhip Effect is studied for the first time in the context of water supply networks. Even though it was not a relevant concept in a traditional long term WDM system, the Bullwhip Effect is emphasized nowadays with new approaches based on hourly management, that look for efficiency optimization. Under these circumstances, demand forecasting is an essential practice and supply time must be taken into account. As a consequence of both, Bullwhip Effect comes out. Through a discrete simulation model, its generation has been showed, as well as the consequences it has on a real-time system: system's oversize, risks of shortage, and energy expenditure increase. Therefore, the Bullwhip Effect should be considered as a head cause of inefficiencies in WDM.

One way to reduce the Bullwhip Effect and to mitigate its damage is the use of advanced forecasting tools. Hence this research has developed a double-loop forecasting system which chooses at each time the most appropriate architecture of the network (both the inputs to be considered and the neurons in the hidden layer). With this ANNs-based system, very low errors in forecasting the hourly demands are achieved in comparison with traditional statistical methods. The tests performed at random moments of time point out that the mean absolute percentage error reduction leads to a large decrease of the Bullwhip Effect, Thereby, the use of the intelligent forecasting system reduces the distortion induced in the water supply network, so that the inefficiencies in WDM are significantly mollified.

There are two main lines of future works that this research group is planning as next steps on this topic. The first one of them is to extend this model to a larger noise conditions scenario, as well as to use a more complex supply structure. Considering these new factors can provide insights to other relevant insights on this issue. More specifically, it is planned to study the Bullwhip Effect in WDM from a supply approach, as many real systems are greatly influenced by hydrological uncertainty (could a reverse Bullwhip Effect exist?). The second line is to integrate this forecasting system within a larger system aimed at optimizing the management.

References

1. Akaike, H. 1969 Fitting autoregressive models for prediction. Annals of the institute of statistical mathematics. 21(1), 243-247.
2. Arellano, M. & Bover, O. 1995 Another look at the instrumental variable estimation of error-components models. Journal of Econometrics. 68(1), 29-51.
3. Bahroun, Z., Moalla, M., Baazaoui, G. & Campagna J. P. 2010 Multi-agent modelling for replenishment policies simulation in supply chains. European Journal of Industrial Engineering. 4(4), 450-470.
4. Bishop, C. M. 1995 Neural Networks for Pattern Recognition. Oxford University Press.
5. Brooks, D. B. 2006 An Operational Definition of Water Demand Management. International Journal of Water Resources Development, 22(4), 521-528.
6. Charlesworth, S., & Adeyeye, K. 2013 Editorial: Water Efficiency and Management. Journal of Water Supply Research and Technology - Aqua, 62(8), 493-495.
7. Disney, S. M., Farasyn, I., Lambrecht, M., Towill, D. R., & Van de Velde, W. 2005 Taming the Bullwhip Effect whilst watching customer service in a single supply chain echelon. European Journal of Operational Research. 173(1), 151-172.
8. Disney, S. M. & Towill, D. R. 2003 The effect of Vendor Managed Inventory (VMI) dynamics on the Bullwhip Effect in supply chains. International Journal of Production Economics. 85, 199-215.
9. Forrester, J. W. 1961 Industrial dynamics, MIT Press. Cambridge, MA.
10. Gato, S., Jayasuriya, N. & Roberts, P. 2007 Temperature and rainfall thresholds for base use urban water demand modeling. Journal of Hydrology. 337(1), 364–376.
11. Gardner, E. S. 2006 Exponential smoothing: The state of the art – Part II. International Journal of Forecasting. 22(4), 637-666.
12. Herrera, M., Torgo, L., Izquierdo, J. & Pérez-García, R. 2010 Predictive models for forecasting hourly urban water demand. Journal of Hidrology. 387(1), 101–140.
13. Hippert, H. S., Pedreira, C. E. & Sourza, R. C. 2001 Neural networks for short-term load forecasting: a review and evaluation. IEEE Transactions on Power Systems. 16 (1), 44-55.
14. Hong, S. R., Kim, S. T. & Kim, C. O. 2010 Neural network controller with on-line inventory feedback data in RFID-enabled supply chain. International Journal of Production Research. 48(9), 2613-2632.

15. Holt, C. C. 2004 Forecasting seasonals and trends by exponentially weighted moving averages. International Journal of Forecasting. 20(1), 1-5.

16. Jaipuria, S. & Mahapatra, S. S. 2014 An improved demand forecasting method to reduce Bullwhip Effect in supply chains. Expert Systems with Applications. 41(5), 2395-2408.

17. Jones, R. H. 1980 Maximum Likelihood fitting of ARMA models to time series with missing observations. Technometrics. 22(3), 389-395.

18. Kohavi, R. 1995 A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedins of the 14th International Joint Conference on Artificial Intelligence, 2, 1137-1143.

19. Lee, H. L., Padmanabhan, V. & Whang, S. 1997 The Bullwhip Effect in supply chains. Sloan Management Review. 38(3), 93-102.

20. Makridakis, S. 1993 Accuracy measures: theoretical and practical concerns. International Journal of Forecasting. 9(22), 527-529.

21. Maturana, F., Shen, W. & Norrie D. H. 1999 Metamorph: An adaptive agent-based architecture for intelligent manufacturing. International Journal of Production Research. 37(10), 2159-2174.

22. Moré, J. J. 1978 The Levenberg-Marquardt algorithm: Implementation and theory. Lecture Notes in Mathematics. 630(1), 105-116.

23. Pino, P., Parreño, J., Gómez, A. & Priore, P. 2008 Forecasting next-day price of electricity in the Spanish energy market using artificial neural networks. Engineering Applications of Artificial Intelligence. 21(1), 53-62.

24. Piroddi, L. & Spinelli, W. 2003 An identification algorithm for polynomial NARX models based on simulation error minimization. International Journal of Control. 77(16), 1767-1781.

25. Saberi, S., Nookabadi, A. S. & Hejazi, S. R. 2012 Applying agent-based system and negotiation mechanism in improvement of inventory management and customer order fulfillment in multi echelon supply chain. Arabian Journal for Science and Engineering. 37(3), 851-861.

26. Sarle, W. S. 1995 Stopped Training and other remedies for overfitting. Proceedings of the 27th Symposium on the Interface of Computing Science and Statistics, 352-360.

27. Sterman, J. 1989 Modelling managerial behaviour: Misperceptions of feedback in a dynamic decision making experiment. Management Science, 35(3), 321-339.

28. Willsie, R. H., & Pratt, L. H. 1974 Water use relationships and projection corresponding with regional growth Seattle region. Water Resources Bulletin. 10(2), 360-371.

29. Zarandi, M. H. F. & Gamasaee R. 2013 A type-2 fuzzy system model for reducing Bullwhip Effect in supply chains and its applications in steel manufacturing. Scientia Iranica. 20(3), 879-899.

CHAPTER 2

REAL-TIME WATER DEMAND FORECASTING SYSTEM THROUGH AN AGENT-BASED ARCHITECTURE♦

♦ This article was published in *International Journal of Bio-Inspired Computation* (Inderscience), vol. 7, no. 3, pp. 147-156 in June 2015. DOI: 10.1504/IJBIC.2015.069559. Authors (in order): Borja Ponte, David de la Fuente, Raúl Pino, Rafael Rosillo.

International Journal of Bio-Inspired Computation, indexed in the ISI, is a Q2-journal in the category 'Computer Science: Theory and Methods' (Science Edition). Impact Factor 2015: 1.390.

Abstract

Water policies have evolved enormously since the Rio Earth Summit (1992). These changes have led to the strategic importance of water demand management. The aim is to provide water where and when it is required using the fewest resources. A key variable in this process is the demand forecasting. It is not sufficient to have long term forecasts, as the current context requires the continuous availability of reliable hourly predictions. This paper incorporates artificial intelligence to the subject, through an agent-based system, whose basis are complex forecasting methods (Box-Jenkins, Holt-Winters, multi-layer perceptron networks and radial basis function networks). The prediction system also includes data mining, oriented to the pre and post processing of data and to the knowledge discovery, and other agents. Thereby, the system is capable of choosing at every moment the most appropriate forecast, reaching very low errors. It significantly improves the results of the different methods separately.

Keywords

Agent-based architecture; Artificial Neural Networks; Box-Jenkins; Data Mining; Demand forecasting; Holt-Winters; Hourly forecasting; Multi-Layer Perceptron; Radial Basis Functions; Water Demand Management.

1. Introduction

Water is a basic resource for human life and for the economic growth of any region. The traditional water management is based on extracting new water resources and making them serve human purposes. This way, large amounts of public money have been invested to finance water projects in order to stimulate the economic development. However, this approach, which is based on the supply increase, has barely taken into account that water is a finite and fragile resource, whose availability depends on the functioning of the hydrological cycle. For this reason, the concept of Water Demand Management has significantly evolved over the last years. Especially since the Earth Summit held in Rio de Janeiro (1992), due to the pressures generated by the population growth, the urbanization and the industrialization, the strategic importance of WDM is understood, as well as its relevance in the efficiency of municipal management (Mohamed and Savenije, 2000). Brooks (2006) proposes an operational definition of WDM with five components: (1) reducing the quantity or quality of water required to accomplish a specific task; (2) adjusting the nature of the task so it can be accomplished with less water or lower quality water; (3) reducing losses in movement from source through use to disposal; (4) shifting time of use to off-peak periods; and (5) increasing the ability of the system to operate during droughts.

A key aspect in any water management plan is demand forecasting. An accurate forecast can minimize the water used to meet demand, but besides it also results in a reduction of the energy used in the process of catchment, purification and distribution of water and it also produces a saving in the resources spent on sizing the storage and distribution system. The traditional approach to water management required only long term forecasts expressed in annual demands or even decades (Willsie and Pratt, 1974). They were enough for the design of the system (capacity of the tanks, dimension of the pipes and connections between the various nodes) and for the development of plans for meeting the demand. Nevertheless, with the passing of time, this horizon has become shorter. In fact, for attaining high efficiency in the WDM, reliable short-term forecasts are required (Gato et al., 2007). Daily forecasts involve the implementation of supply plans, by setting the system to that effect. The next step is hourly water forecasting. According to Herrera et al. (2010), the ready availability of hourly predictions of water demand is crucial due to three main reasons: it allows to determine the optimal regulation and pumping systems to meet the predicted demand, which promotes energy efficiency (operative point of view);

it allows to combine water sources in the most appropriate way to achieve a preset standard in the supply water (quality point of view); and it allows to detect failures and network losses through the comparison of the actual and expected flow (vulnerability point of view).

The literature on the subject contains several works of short term demand forecasting. The first one was written by Maidment et al. (1985), who used statistical models (in particular, ARIMA methodology) to express the daily water demand as a function of ambient temperature and volume of rain. In a later work, the same authors (Maidment and Miaou, 1986) proved its efficiency for nine American cities. Other authors (e.g., Shvarster et al., 1993; An et al., 1995) followed this line, using statistical methods and climatic factors in the prediction. Lertpalangsunti et al. (1999) were pioneers in the introduction of artificial intelligence (AI) in the study. They developed a complex forecasting system, which integrated fuzzy logic, artificial neural networks (ANNs) and case-based reasoning, which was tested with high efficiency to forecast the daily water demand in the city of Regina (Canada). Msiza et al. (2007) introduced support vector machines (SVM) in the subject, in order to compare its performance with ANNs, using two different structures: Multi - Layer Perceptron (MLP) and Radial Basis Function (RBF). They conducted the study on the daily demand of the province of Gauteng (South Africa) and the ANNs outperformed the SVM. Herrera et al. (2010) further reduced the time horizon and they evaluated six predictive models (ANNs, projection pursuit regression, multivariate adaptive regression splines, SVM, random forests and a weighted pattern-based models) in forecasting the hourly demand of the city of Valencia (Spain). The authors justify that in this modern environment the ready availability of hourly water demand predictions is crucial. Bio-inspired algorithms have also been used in other aspects around WDM –e.g. Liu and Lv (2009) used the particle swarm optimization algorithm to forecast the residual life of underground pipelines.

On the one hand, one of the main conclusions of the literature review is that these advanced methodologies are proven to give a great performance in the forecasting of short term water demand, both daily and hourly. There are not big differences between their results, as the choice of the optimal one depends on the characteristics of the study period and its recent past. On the other hand, most of the authors use climatic factors in the predictions, as they lead to improve the results. However, the just-in-time availability of these climatic factors in order to perform the hourly forecasting could be a hurdle difficult to overcome by a real-time WDM system. Therefore, considering those factors could be a constraint for the implementation. This way, Nasseri et al. (2011) developed a model based on AI techniques (genetic algorithms and Kalman filter) with excellent results, taking only in consideration data from previous demand.

Under these circumstances, this paper shows the development of a system for the real-time water demand forecasting based on AI techniques. More specifically, we use an agent-based architecture to construct the system, whose core are the advanced forecasting agents but it is also formed by other agents which carry out other important functions, which will be described next. The system continuously receives values from water hourly demand and it is capable of choosing the most reliable forecasting technique at each moment. This way, it could be implemented in different scenarios, as it has the ability of adapting to them. So, after the literature analysis, the idea of this article is to combine different tools in order to obtain a forecasting system with greater accuracy, even without the availability of real-time information about the climatic factors. The great advantage of using the agent-based architecture is that this forecasting system can be integrated into a larger management system built under the same principles.

Our investigation line has been the following: (1) Problem world and problem statement; (2) Development of the conceptual model; (3) Implementation of the forecasting methods; (4) Construction of the real-time water demand forecasting system; (5) Experimentation and obtaining results; and (6) Problem analysis and deriving conclusions. Such work structure is spread across this paper, which is divided into four main sections, including this introduction. Section 2 describes the forecasting system that we have created, with the different agents that form it and their purpose, the structure that encompasses all and the relationships between them. Section 3 contains the numerical results after testing the system with hourly water demand time series and the discussion thereof. Finally, section 4 presents the main conclusions that we have obtained based on the stated objectives, as well as the future investigation lines.

2. Description of the Real-Time Water Demand Forecasting System

Figure 1 shows schematically the forecasting system that we have devised and implemented. It consists of nine different agents: the Interface Agent, the Storage Agent, the Data Mining Agent, the Fitness Agent and the five Forecasting Agents (Naïve Agent, Box-Jenkins Agent, Holt-Winters Agent, MLP-NN Agent and RBF-NN Agent). It should be noted that we are using forecasting techniques of different nature. The system receives hourly data about the water demand from the measurement equipment and shows the real-time forecast to the user, in order to the decision-making process. The agent-based approach also allows its connection to a larger management system. Below, we detail the functionality of each agent, and the relationships among them.

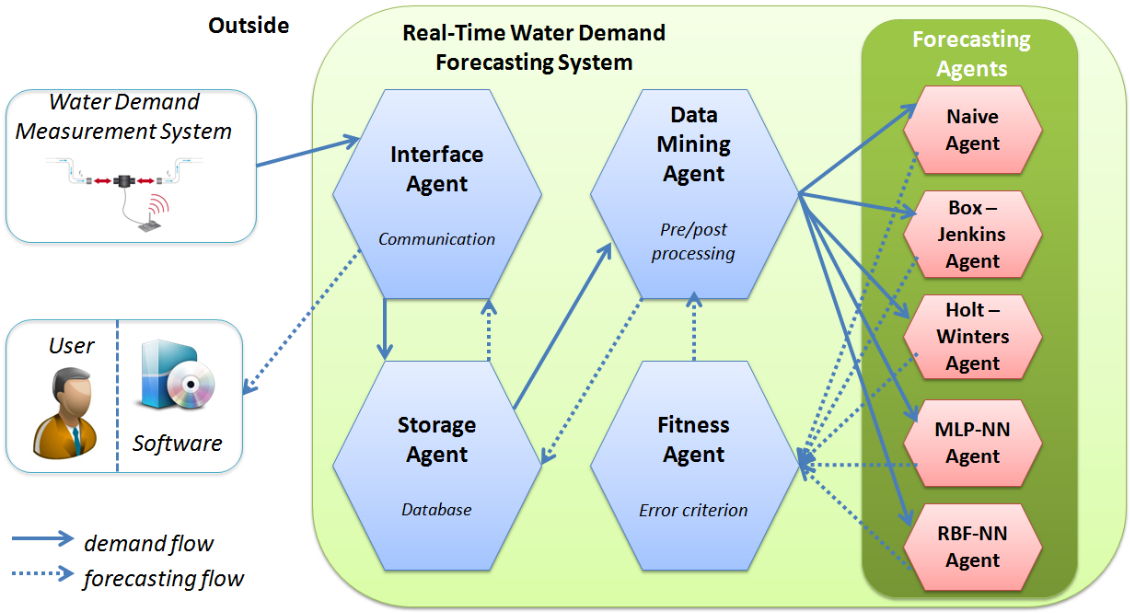


Figure 1. General outline of the real-time water demand forecasting system, with the various agents that form it and the relationships among them (two main flows) and with the outside.

2.1. Transmission Agents

The *Interface Agent* connects the forecasting system with its environment. That is to say, it acts as the intermediary between the rest of the agents and the outside with the aim of reaching the homogeneity in the agent-based system. Thus, it works in a double way: (1) it transmits the demands received hourly from the measurement equipment to the data base; and (2) it transfer the best forecast at each hour to the outside.

The *Storage Agent* manages a database attached to the system that saves hourly the values of both actual demands and the forecasts performed by the five agents. Besides, it also saves the best forecast performed at every hour. It is necessary to store all this information (not only the best forecast) because past forecasts will influence in the selection of the best forecast in future. Therefore, the *Storage Agent* is in permanent contact with the *Interface Agent* and the *Data Mining Agent* to store and move information from the outside to the other forecasting agents (demand flow), and in the opposite direction (forecasting flow).

The *Data Mining Agent* carries out the pre-processing of the information stored in the database and the post processing of the predictions. On the one hand, this involves extracting the last 1020 hourly demands (6 weeks + 12 hours, see section 2.9) from the database. It has proven to be a suitable time period, in terms of identifying the seasonality and trend of the series. On the other hand, with the aim of performing the neural networks forecasting, it involves the creation of thirteen time series with the demands displaced (displacement from 1 to 4, from 24 to 28 and from 168 to 172 hours, given the double periodicity of the series, and because the other values have not proved to be significant) to find inference rules and try to explain each demand based on past data. In addition, the *Data Mining Agent* is connected with the Forecasting Agent, to perform transformations on the variables (e.g., logarithmic, differentiation or quantification of non-numerical variables) when it is needed.

2.2. Forecasting Agents

These agents are the real core of the real-time water demand forecasting system. We are using naïve models, classical statistical methods and AI-based techniques, in order to try to combine the advantages of each alternative.

The *Naïve Agent* performs the demand forecast using a naive method, which estimates the hourly demand ($\widehat{\mathrm{D}}_{\mathrm{t}}$) as the demand in the previous hour ($\mathrm{D}_{\mathrm{t-1}}$), adjusted by the increase (or decrease) in the demand in the same time interval of the previous week ($\mathrm{D}_{\mathrm{t-168}} - \mathrm{D}_{\mathrm{t-169}}$), by (1). This is a very simplified model –and hence it requires a insignificant calculation time– but it offers good performance in regular series, like the one we have.

$$\widehat{D}_t = y_t = D_{t-1} + (D_{t-168} - D_{t-169}) \quad (1)$$

The *Box–Jenkins Agent* performs the forecast using the ARIMA methodology (Box and Jenkins, 1970). These models can be expressed by $(\mathrm{p,d,q})(\mathrm{P,D,Q})_{\mathrm{n}}$, where the parameters are the orders of autoregression (p, P), differentiation (d, D) and moving average (q, Q). Lowercase variables are not seasonal components, while the uppercase ones are seasonal, with periodicity n. In our case, n=168. These models consider that the future value of the differentiated variable ($\widehat{\Delta^d D_t}$) can be expressed as a function of past observations ($D_{t-i}, i \in [1, n]$) and a random error ($\varepsilon_{t-j}, i \in [1, q]$). It is expressed in (2), where Δ is the differentiation operator, γ is the constant model, φ_i are the parameters associated with autoregression, and θ_j are the parameters associated with the moving average.

$$\widehat{\Delta^d D_t} = y_t = \gamma + \sum_{i=1}^{p} \varphi_i \, \Delta^d D_{t-i} + \sum_{k=1}^{P} \varphi_{kn+p} \, \Delta^d D_{t-(kn+p)} - \sum_{j=1}^{q} \theta_j \, \varepsilon_{t-j} - \sum_{m=1}^{Q} \theta_{mn+q} \, \varepsilon_{t-(mn+q)} \tag{2}$$

The method of obtaining the statistical model $(p, d, q)(P, D, Q)_n$ associated with each time series is based on the sequential process of: (1) identifying the possible model; (2) parameter estimation; and (3) validation. It is repeated until the model is validated through their autocorrelation functions and until its forecasts are validated by a given error criterion. In our case, the *Box–Jenkins Agent* seeks the model that best fits the input time series, using the following statistics for the comparison of the different proposed models: goodness-of -fit according to the criteria of MAPE; residual simple autocorrelation function (ACF); and residual partial autocorrelation function (PACF). The method of obtaining the model and calculating the coefficients is described in more detail in Box and Jenkins (1970).

The *Holt–Winters Agent* uses the Holt–Winters exponential smoothing method to forecast. Its base is a simple exponential smoothing, which express the demand as a weighted average between the demand and the forecast of the previous period. Holt (1957) modified this model so that it can be applied in trended series and Winters (1960) adapted it for series with seasonality. There are two main Holt-Winters models, depending of the type of seasonality: (1) Multiplicative; and (2) Additive. These models can be mathematically expressed by (3) and (4), in the previous order, where y_t represents the forecast, $\overline{R_{t-1}}$ is the estimate of the deseasonalized level or overall smoothing in the previous period, $\overline{G_{t-1}}$ is the estimate of the trend or smoothing of the trend factor in the previous period, and $\overline{S_{t-L}}$ is the estimate of the seasonal component or smoothing of L (the seasonal index) periods ago. In our case, L=168.

$$\widehat{D}_t = y_t = (\overline{R_{t-1}} + \overline{G_{t-1}}) \cdot \overline{S_{t-L}} \tag{3}$$

$$\widehat{D}_t = y_t = \overline{R_{t-1}} + \overline{G_{t-1}} + \overline{S_{t-L}} \tag{4}$$

It should be noted that each one of the previous parameters depends on a different smoothing constant. The procedure for the estimates of model parameters is detailed, among others, in Kalekar (2004). In our case, the *Holt–Winters Agent* looks for the model that best fits the input time series using the same statistics for the comparison of three alternatives (the multiplicative model, the additive model, and the simple seasonal model, where there is no trend) as the one used in the *Box-Jenkins Agent.*

The *MLP–NN Agent* and the *RBF–NN Agent* estimate the hourly demand through an Artificial Neural Network (ANN) with three levels: an input layer (predictor variables, which are obtained by means of the *Data Mining Agent*), a hidden layer (composed by nodes that, during optimization process, attempt to functionally map the model inputs to the model outputs) and an output neuron (variable to predict). Figure 2 shows schematically the general structure of the ANN that we have used.

In both cases, the data available for each forecast (1008 hourly water demands) are randomly separated into two groups. 70% is oriented to the batch training of the network, by means of the back-propagation algorithm. The remaining 30% has been directed for verifying the network. We use different stopping criteria (maximum number of steps without reducing error: 1000; maximum workout time: 1 minute; minimal relative change in training error: 0.0001; minimal relative change in error rate training: 0.001). The steps for developing the ANNs are similar to those detailed in Pino et al. (2008).

There are various ANN architectures. On the one hand, the *MLP–NN Agent* focus on the Multi-Layer Perceptron (MLP). These are networks that have more than one layer of adaptive weights. A MLP has three layers of units taking values in the range 0-1, and each layers is nourished with the previous ones. Any number of weighted connections can be used, but MLPs with two weighted connections are very much capable of approximation just about any functional mapping (Bishop, 1995). The MLP can be mathematically represented by (5), where $\mathrm{y_t}$ represents the output (forecast), $\mathrm{f_{outer}}$ represents de output layer, $\mathrm{f_{inner}}$ represents the input layer transfer function, $\mathrm{w_{xy}}$ represents the weights and biases ($\mathrm{i} \in [1{,}17]$ refers to the input neurons and $\mathrm{j} \in [1, \mathrm{n}]$ refers to the hidden neurons) and $^{(z)}$ represents the z-th layer.

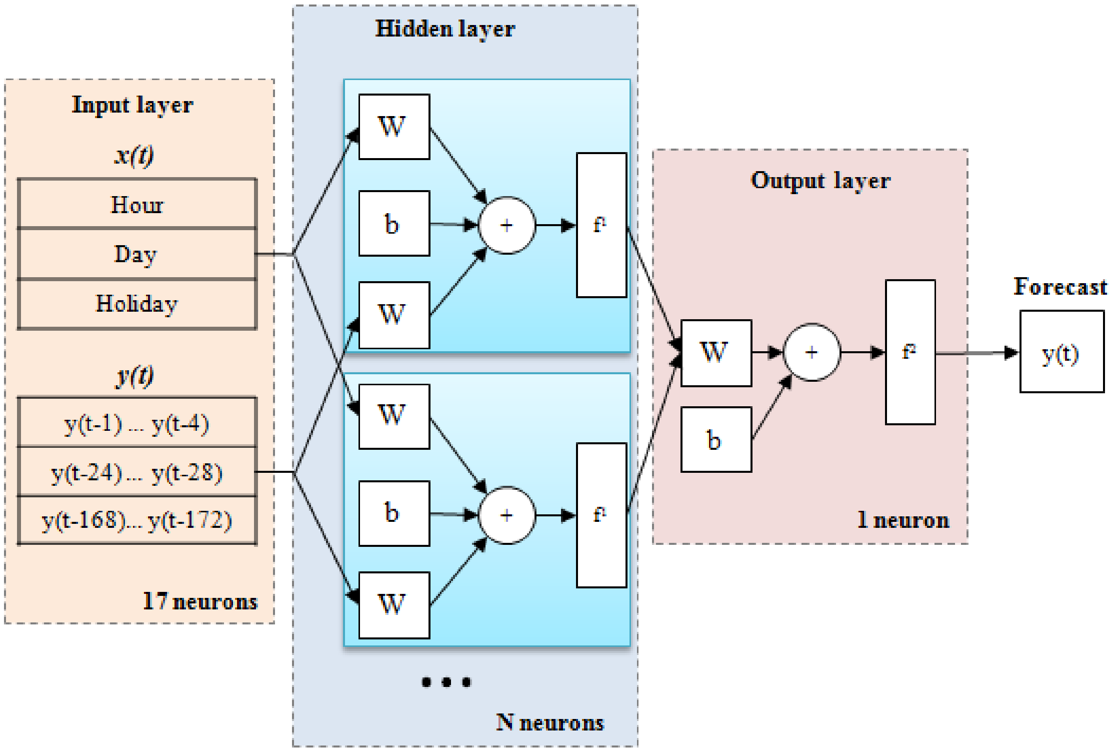


Figure 2. General structure of the ANN with its three layers (input layer, hidden layer and output layer).

Note: The 17 predictor variables are: the day and hour of the demand to forecast, a binary variable that differentiates holidays and working days, and 14 past values related to the seasonality of the time series (from y(t-1) to y(t-4), from y(t-24) to y(t-28), from y(t-168) to y(t-172)). The number of neurons in the hidden layer depends on the time se-ries. The only output neuron is related to the variable to predict, so that it performs the forecast.

$$\widehat{D}_t = y_t = f_{outer}\left[\sum_{j=1}^{n} w_{1j}{}^{(2)} \cdot f_{inner}\left(\sum_{i=1}^{17} w_{ji}{}^{(1)} \cdot x_i + w_{j0}{}^{(1)}\right) + w_{10}{}^{(2)}\right] \qquad (5)$$

On the other hand, the *RBF-Agent* performs the forecast according to the Radial Basis Function (RBF) Architecture. In the RBF, the activation of the hidden unit is determined by the distance between the input vector and the prototype vector, leading to a two stage procedure (Bishop, 1995): (1) Determination of the centre of the network using unsupervised methods; and (2) Determination of the final-layer weights. Hence, the RBF networks provide an interpolation function –called basis functions–, which passes through each and every data point. It can be mathematically represented by (6), where y_t represents the output (forecast), w_{xy} represents the weights and biases ($j \in [1, n]$ refers to the hidden neurons) and φ_j represents the activation function of the output layer.

$$\widehat{D}_t = y_t = \sum_{j=1}^{n} w_{1j} \cdot \varphi_j \tag{6}$$

2.3. Fitness Agent

The *Fitness Agent* selects the best forecast at each moment through the comparison of the last demands and the forecasts performed by the five Forecasting Agents. It uses the criterion of the minimum MAPE (mean absolute percentage error), introduced by Makridakis (1993). After evaluating different options, we have obtained the best results when the MAPE is calculated for the last 12 hours, so this agent uses this number for the selection. Figure 3 synthesizes the time horizon of the forecasting process, and the role of the *Fitness Agent* within the whole system.

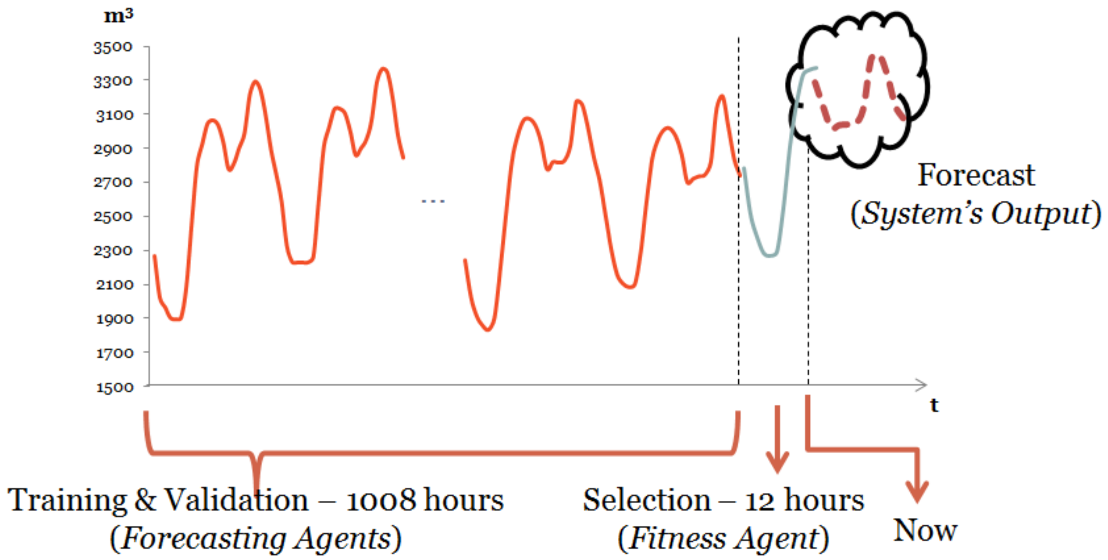


Figure 3. Time horizon of the forecasting process.

Note: The last 12 hours (both the demands and the forecasts of the five ag*ents) are used to determine the* best forecasting method in each moment, while the previous 1008 ho*urs (only the demands) are used* for the training and validation of the different forecasting methods, in order to choose the most appropriate model in each case (except the case of the Naïve Agent, whose functioning is much more simple).

3. Numerical Application and Discussion of the Results

In order to test the forecasting system, we have used a simulated time series with over 15,000 data points, which represents the hourly water demand in the city of Gijón (a municipality of 300,000 inhabitants in the north of Spain) during 21 months (years 2009 and 2010). To obtain it, we have based on the monthly demand of the city, a distribution model of hourly water demand for a city in south-eastern Spain (Herrera et al., 2010), and random parameters. It should be noted that in this city, 71% of invoiced water is oriented to domestic use, 23% of this water has an industrial use, and the remaining 6% is managed by the city council.

The time series of the hourly water demand is a complex series with a double seasonality. On the one hand, it has a daily periodicity, namely every 24 hours the series has a similar structure. There is a sharp decrease from 19h until 02h, when demand stabilizes around a daily minimum, until 06h. Then, it grows until 11am, where it sets a first local maximum. From there, demand undergoes a slight decline to local minimum at 14h, at which time it surges to a second local maximum at 19h. The mentioned times are approximate and vary according to the season of the year. On the other hand, there is a weekly periodicity (168 hours), as the structure is repeated every week, with a significantly lower consumption on Saturdays and even more on Sundays. Moreover, the time series does not remain in a constant range, but it exhibits different trends in both mean and variance, throughout the year. To illustrate the explanation, figure 4 represents two parts of the time series.

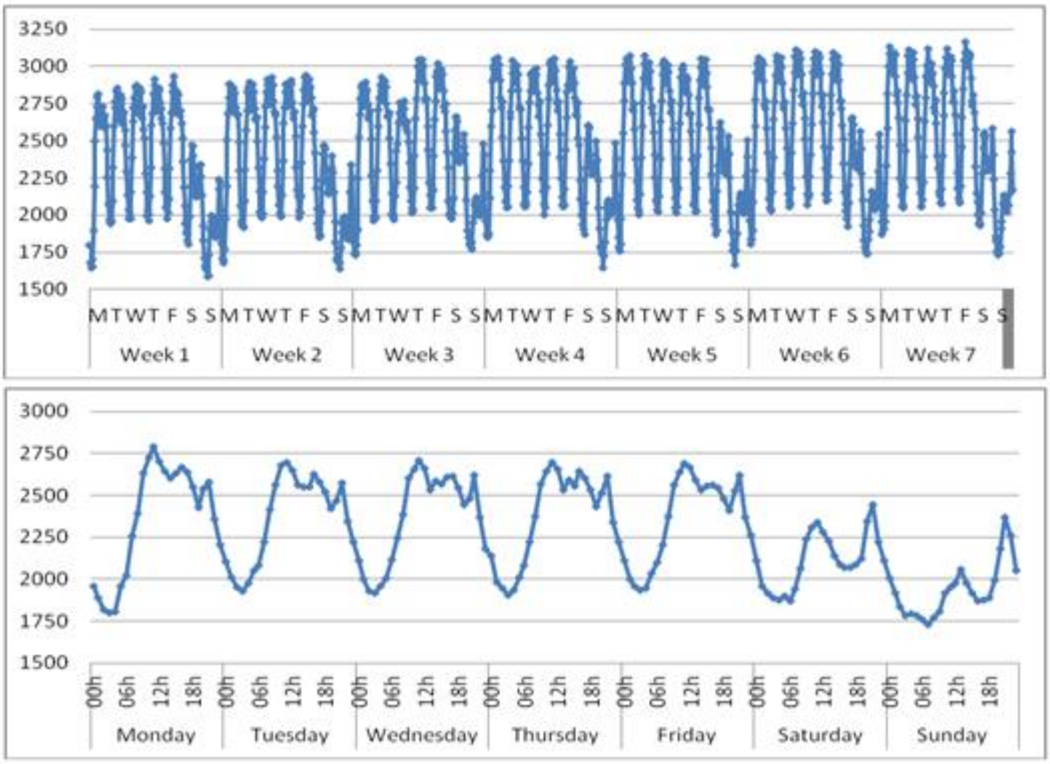


Figure 4. Two extracts from the time series (values in cubic meters / hour).

Note: The top graph (time horizon: seven weeks) brings evidence of the weekly periodicity and its trend, and the graph below (time horizon: one week) shows the daily periodicity of the time series.

Table 1. Results of the numerical test.

Forecas. Period	*Forecas. Agent*	*Features*	*Fitness MAPE*	*Forecas. MAPE*
Test I Thursday May 14, 2009 04h (*working day*)	Naive	-	1.94%	1.60%
	Holt–Winters	Simple seasonal	2.56%	1.17%
	Box–Jenkins	$(0,1,6)(0,1,1)_{168}$	2.26%	1.25%
	MLP–NN	17-8-1	1.27%	1.14%
	RBF–NN	17-10-1	1.48%	1.39%
Test II Wednesday Sept. 8, 2010 16h (*working day*)	Naive	-	1.74%	1.98%
	Holt–Winters	Simple seasonal	1.76%	1.45%
	Box–Jenkins	$(0,1,3)(1,1,0)_{168}$	1.87%	2.21%
	MLP–NN	17-9-1	1.40%	1.52%
	RBF–NN	17-10-1	1.12%	1.53%
Test III Sunday June 7, 2009 12h (*weekend*)	Naive	-	3.59%	3.93%
	Holt–Winters	Additive	4.34%	3.03%
	Box–Jenkins	$(0,1,3)(0,1,1)_{168}$	3.30%	3.22%
	MLP–NN	17-6-1	2.83%	2.91%
	RBF–NN	17-8-1	4.29%	3.85%
Test IV Friday Feb. 5, 2010 23h (*weekend*)	Naive	-	3.51%	2.50%
	Holt–Winters	Multiplicative	3.69%	2.62%
	Box–Jenkins	$(1,1,5)(0,1,1)_{168}$	2.63%	8.04%
	MLP–NN	17-9-1	2.39%	2.48%
	RBF–NN	17-11-1	4.28%	1.82%
Test V Tuesday Dec. 8, 2009 18h (*holiday*)	Naive	-	5.38%	3.19%
	Holt–Winters	Simple Seasonal	23.55%	6.09%
	Box–Jenkins	$(1,1,1)(1,1,0)_{168}$	24.44%	7.46%
	MLP–NN	17-4-1	3.74%	2.19%
	RBF–NN	17-8-1	6.20%	1.78%
Test VI Wednesday Oct. 13, 2010 04h (*after holiday*)	Naive	-	4.37%	3.22%
	Holt–Winters	Additive	5.86%	8.65%
	Box–Jenkins	$(2,1,12)(0,1,1)_{168}$	8.04%	11.48%
	MLP–NN	17-11-1	2.98%	2.03%
	RBF–NN	17-7-1	3.79%	2.00%

Note: This table contains the following five columns: (1) the beginning of the time period to predict (previous 1020 data are used by the system to forecast); (2) the forecasting method, by means of the Agent which performed the fore-cast; (3) its main feature chosen by the agent (that is to say, the Holt-Winters model chosen, the ARIMA Model, and the structure of the ANN); (4) the MAPE calculated by the Fitness Agent (12 previous demands) and which determines its selection; and (5) the MAPE obtained in the prediction made by each agent. In order to calculate the Forecasting MAPE, we use the following 12 forecasts, with the aim of looking for consistency in our results.

Data of the hourly water demand time series can be divided into three groups: (1) working days; (2) weekend; and (3) holidays (and days around them, whose forecast could be crucially affected by holidays). After several tests, Table 1 presents the numerical results for two standard cases of each group. In every test, we stand out the MAPE of the forecast performed by the system in the last column (Forecasting MAPE), which is chosen between the various methods and corresponds to the Forecasting Agent with minimizes the MAPE (Fitness MAPE). This Forecasting Agent is stood out in the second column.

In the forecasting of working days, all methods achieve low forecast errors (between 1.12% and 2.56% for the two tests shown). Therefore, all of them are capable of understanding the running of the series quite accurately. Even the Naïve Agent, which adopts an oversimplification, provides good results given the regular nature of the series. The statistical models of Box-Jenkins (ARIMA) and Holt-Winters (exponential smoothing) generally improve the results. Nevertheless, as expected, the introduction of artificial intelligence in the model, through ANNs, causes a greater decrease in the MAPE. The results of the RBF and MLP structures have a similar goodness-of-fit –there is no significant difference in its performance. By way of example, Figure 5, which represents the forecasting time period for test I, shows what we have explained.

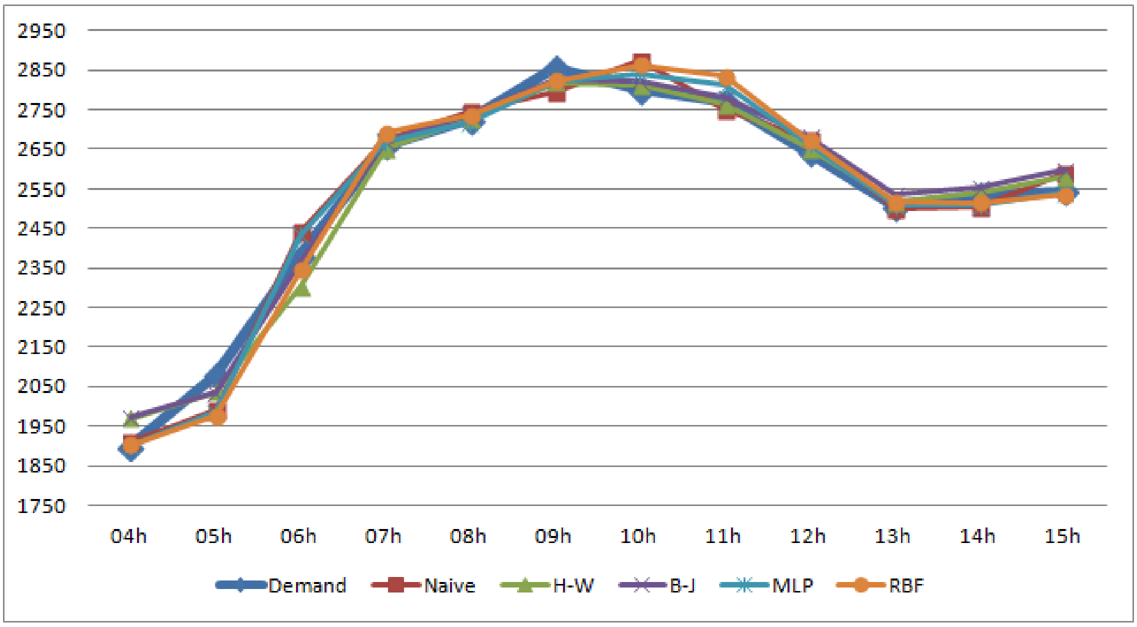


Figure 5. Actual demand and various predictions for the forecasting period in test I (values in cubic meters / hour).

Note: The MLP–NN forecasts is the one provided by the system (Forecasting MAPE 1,14%), but the different among the various methodologies are much smaller than in the other cases studied.

On weekend forecasting, all methods increase substantially the error generated. This is easily understandable, since the influence of the working days on the model is much higher. Statistical models in this case are less robust, as they show high variability in the goodness of their results. In some tests, they achieve low forecast errors but in others they are not able to accurately grasp the series. The RBF structure in ANNs shows a similar effectiveness. However, the *MLP–NN Agent* offers the best performance, reaching a MAPE less than 3% in all cases analyzed. By way of illustration, Figure 6 shows the demand and the forecasts performed by the various agents in *test III*.

The problems of statistical models are more evident on holidays and days around them. On the one hand, the system is not capable of adapting its structure in atypical days, while ANNs can manage it (see *test V*). On the other, the presence of a holiday in the days before the forecasting period introduces a distortion in the series model that deviates slightly the forecast (see *test VI*). Therefore, in this last group, the differences between the different methodologies are amplified and AI allows improving strongly the forecast. This can be shown in Figure 7, which displays the demand and the forecasts performed by the various agents in *test VI*. In holidays, again, the forecasting of ANN with MLP structure is more robust than the ANN with RBF structure.

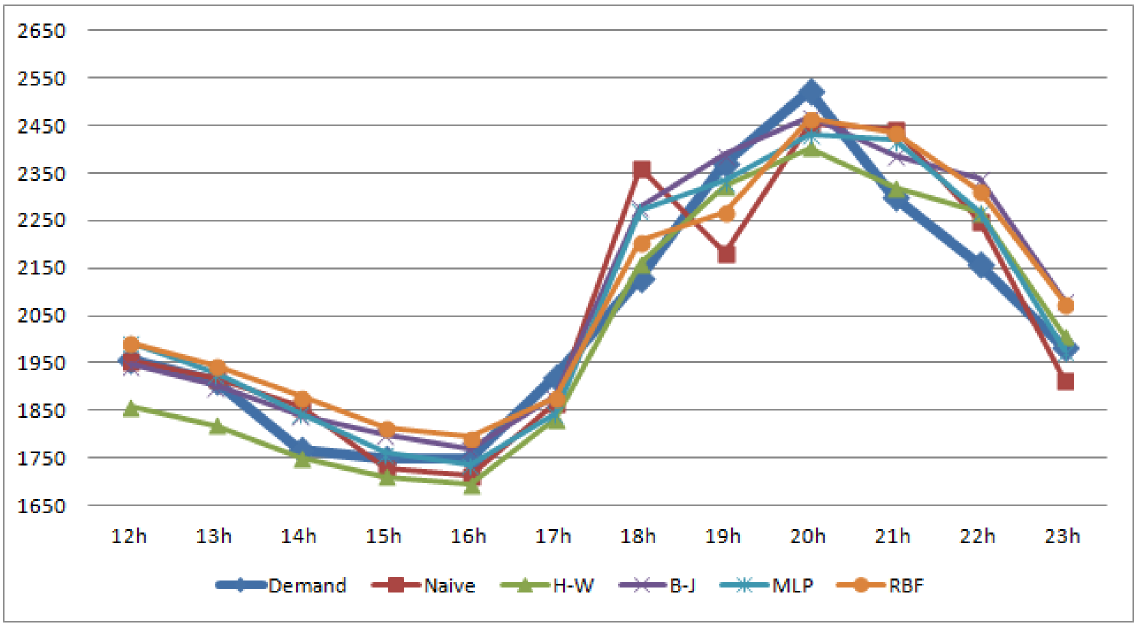


Figure 6. Actual demand and various predictions for the forecasting period in test III (values in cubic meters / hour).

Note: The MLP–NN forecasts is the one provided by the system (Forecasting MAPE 2,95%).

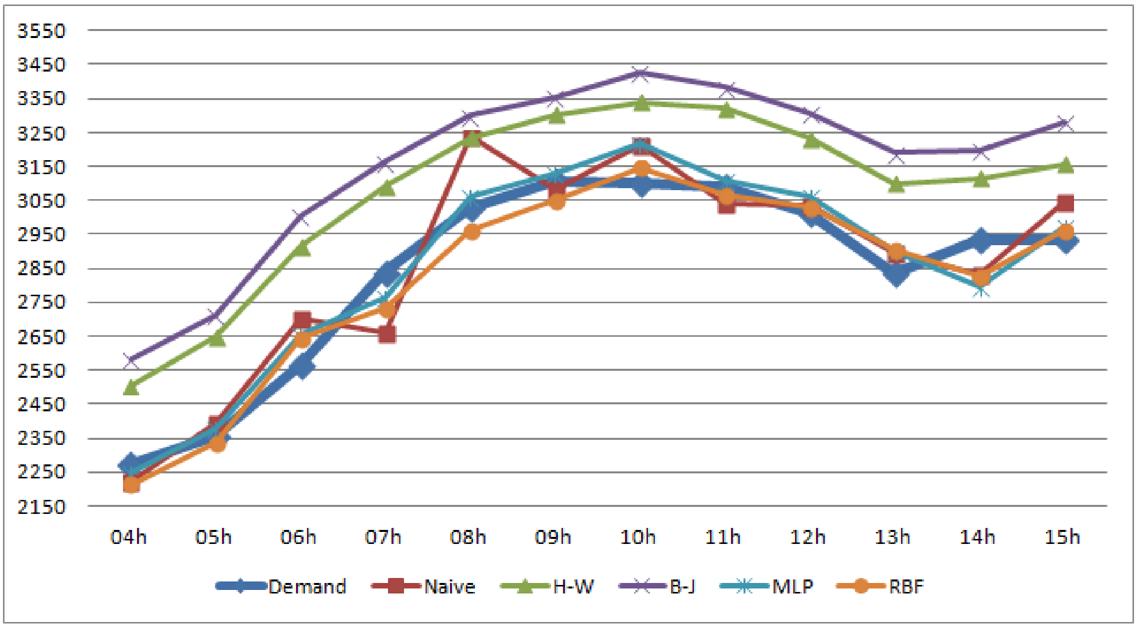


Figure 7. Actual demand and various predictions for the forecasting period in test VI (values in cubic meters / hour).

Note: The statistical methods have big difficulties to forecast accurately. The MLP–NN forecast is the one provided by the system (Forecasting MAPE 2,98%), although the RBF-NN forecast is slightly better.

4. Conclusions and Next Steps

This paper presents an application of agent-based architecture in hourly demand forecasting, a key aspect in Water Demand Management (WDM). The cores of the system are advanced statistical models (ARIMA and Holt-Winters exponential smoothing) and artificial intelligence (AI) techniques, such as Multi-Layer Perceptron (MLP) and Radial Basis Functions (RBF) Artificial Neural Networks (ANNs). Tests that have been carried out demonstrate the effectiveness of the real-time forecasting system, which selects at each moment the best forecast. Obviously, there is no way to ensure that the system always selects the prediction that will generate the lower error in future, but tests show that if the forecasting method selected is not optimal, it is closer to the optimum. The goodness-of-fit of each technique depends on the characteristics of the forecasting period, although MLP is the most robust method.

The multi-agent environment draws a very appropriate approach to tackle the problem, as the system provides at all times the forecast which it understands as the best. Under these circumstances, it allows the addition of new intelligent forecasting tools by means of new Forecasting Agents, without varying the rest of the system. In addition, this approach has enormous potential in increasing its functionality, because it allows to complete the study by adding new agents. This way, this real-time water demand forecasting system will be integrated in a larger system aimed at optimizing the management.

References

1. An, A., Shan, C. C., Cercone, N. and Ziarko, W. (1995) 'Discovering rules from data for water demand prediction', Proceedings in the Workshop on Machine Learning and Expert System, pp. 187-202.
2. Bishop, C. M. (1995) 'Neural Networks for Pattern Recognition', Oxford University Press.
3. Box, G. E. P. and Jenkins, G. M. (1970) 'Time Series Analysis: Forecasting and Control', Holden Day, San Francisco.
4. Brooks, D. B. (2006) 'An Operational Definition of Water Demand Management', International Journal of Water Resources Development, Vol. 22(4), pp. 521-528.
5. Gato, S., Jayasuriya, N. and Roberts, P. (2007) 'Temperature and rainfall thresholds for base use urban water demand modeling', Journal of Hydrology. Vol. 337, pp. 364–376.
6. Herrera, M., Torgo, L., Izquierdo, J. and Pérez-García, R. (2010) 'Predictive models for forecasting hourly urban water demand', Journal of Hidrology, Vol. 387, pp. 101–140.
7. Holt, C. C. (1957) 'Forecasting seasonals and trends by exponentially weighted moving averages' ONR Memorandum, Vol. 52, Pittsburgh, PA, Carnegie Institute of Technology.
8. Kalekar, P. S. (2004) 'Time series Forecasting using Holt-Winters Exponential Smoothing', Kanwal Rekhi School of Information Technology, Tech. Rep.
9. Lertpalangsunti, N., Chan, C., Mason, R. and Tontiwachwuthikul, P. (1999) 'A tool set for construction of hybrid intelligent forecasting systems: application for water demand prediction', Artificial Intelligence in Engineering, Vol. 13, pp. 21-42.
10. Liu, Q. M., and Lv, W. (2009) 'The forecasting residual life of underground pipeline based on particle swarm optimisation algorithm', International Journal of Bio-Inspired Computation, Vol. 1 (4), pp. 270-275.
11. Maidment, D. and Miaou, S. (1986) 'Daily water use in nine cities', Water Resources Research, Vol. 22 (6), pp. 845-851.
12. Maidment, D., Miaou, S. and Crawford, M. (1985) 'Transfer function models of daily urban water use', Water Resources Research, Vol. 21 (4), pp. 425-432.

13. Makridakis, S. (1993) 'Accuracy measures: theoretical and practical concerns', International Journal of Forecasting, Vol. 9 (22), pp. 527-529.

14. Mohamed, A. S. and Savenije, H. H. G. (2000) 'Water Demand Management: Positive Incentives, Negative Incentives or Quota Regulation?', Physics and Chemistry of the Eart Part B – Hydrology Oceans and Atmosphere, Vol. 25(3), pp. 251-258.

15. Msiza, I., Nelwamondo, F. and Marwala, T. (2007) 'Artificial neural networks and support vector machines for water demand time series forecasting', IEEE International Conference on Systems, Man and Cybernetics, pp. 638-643.

16. Nasseri, M., Moeini, A., and Tabesh, M. (2011) 'Forecasting monthly urban water demand using Extended Kalman Filter and Genetic Algorithms', Expert Systems with Applications, Vol. 38, pp. 7387-7395.

17. Pino, R., Parreño, J., Gómez, A. and Priore, P. (2008) 'Forecasting next-day price of electricity in the Spanish energy market using Artificial Neural Networks', Engineering Applications of Artificial Intelligence, Vol. 21 (1), pp. 53-62.

18. Shvarster, L., Shamir, U. and Feldman, M. (1993) 'Forecasting Hourly Water Demand by Pattern – Recognition Approach', Journal of Water Resources Planning and Management, Vol. 119 (6), pp. 611-627.

19. Willsie, R. H. and Pratt, L. H. (1974) 'Water use relationships and projection corresponding with regional growth Seattle region', Water Resources Bulletin, Vol. 10(2), pp. 360–371.

20. Winters, P. (1960) 'Forecasting sale by exponentially weighted moving averages', Management Science, Vol. 6, pp. 324-342.

CHAPTER 3

INTELLIGENT DECISION SUPPORT SYSTEM FOR REAL-TIME WATER DEMAND MANAGEMENT[♦]

♦ This article was published in *International Journal of Computational Intelligence Systems* (Taylor & Francis), vol. 9, no. 1, pp. 168-183 in February 2016. DOI: 1080/18756891.2016.1146533. Authors (in order): Borja Ponte, David de la Fuente, José Parreño, Raúl Pino.

International Journal of Computational Intelligence Systems, indexed in the ISI, is a Q4-journal in the category 'Computer Science: Artificial Intelligence' (Science Edition). Impact Factor 2015: 0.391.

Abstract

Environmental and demographic pressures have led to the current importance of Water Demand Management (WDM), where the concepts of efficiency and sustainability now play a key role. Water must be conveyed to where it is needed, in the right quantity, at the required pressure, and at the right time using the fewest resources. This paper shows how modern Artificial Intelligence (AI) techniques can be applied on this issue from a holistic perspective. More specifically, the multi-agent methodology has been used in order to design an Intelligent Decision Support System (IDSS) for real-time WDM. It determines the optimal pumping quantity from the storage reservoirs to the points-of-consumption in an hourly basis. This application integrates advanced forecasting techniques, such as Artificial Neural Networks (ANNs), and other components within the overall aim of minimizing WDM costs. In the tests we have performed, the system achieves a large reduction in these costs. Moreover, the multi-agent environment has demonstrated to propose an appropriate framework to tackle this issue.

Keywords

Water Demand Management; Decision Support System; Multi-agent Systems; Neural Networks.

1. Introduction

Water is considered to be the most important natural resource. This basic resource is essential for all kinds of social and economic activities, as well as for the life and the health of the mankind. From this perspective, it is easy to justify the great importance of water management. This fact has been accentuated in recent years, mainly since the Earth Summit held in Rio in 1992, due to pressures caused by population growth, urbanization and industrialization. That is, the scarcity of resources and the respect for the environment have drawn a new context that threatens both the quality and the availability of this resource.

For the aforementioned reasons, policies regarding water have undergone major changes over the last two decades. Accordingly, the concept of Water Demand Management (WDM) has significantly evolved[1]. Its current definition encompasses five main goals[2]: (1) reducing the quantity and quality of water required to accomplish a specific task; (2) adjusting the nature of the task so it can be accomplished with less or lower quality water; (3) reducing losses in movement from source through use to disposal; (4) shifting time of use to off-peak periods; and (5) increasing the system ability to operate during droughts.

In this context, the effectiveness of a WDM system heavily depends on demand forecasting. This is not only about minimizing the water used to meet demand, but accurate forecasts have associated other benefits, such as the reduction of energy consumption in water catchment, purification and distribution processes. This fact highlights the importance of water demand forecasting, which can be divided into:

- Very long-term forecasting (decades)[3], which crucially determines the design of the water supply system, e.g. tanks capacities and pipes dimensions.
- Long-term forecasting (years)[4], which allows managers to develop plans for managing water demand, as well as adjustments in the distribution system.
- Mid-term forecasting (months)[5], used to adjust previous planning, through comparing actual and planned data, as well as to determine water price.
- Short-term forecasting (days)[6], which involves the implementation of supply plans, by setting the necessary systems to that effect.
- Very short-term forecasting (hours)[7], which results in water conveyance from tanks to points-of-consumption when requited, in the right quantity and pressure.

Therefore, with the passing of time, smaller time horizon forecasts are demanded in order to meet the requirements of this new context. Long-term forecasts are not enough to reach a suitable water management, but in the current circumstances reliable short-term forecasts are essential[8]. Thus, to meet efficiently the demand, forecasts must be continuously available with the aim of[9]:

- In terms of operation and energy efficiency, determining optimal regulation and pumping schemes to supply the predicted demand.
- In terms of quality, combining water sources in the most appropriate way to achieve a given standard in the supplied water.
- In terms of vulnerability, detecting network losses and failures through comparing actual and expected flows.

From this perspective, this article proposes the application of modern Artificial Intelligence (AI) techniques to WDM. With the aim of drawing up a holistic approach to this issue, we have developed an Intelligent Decision Support System (IDDS). That is, we seek the system overall solution rather than tackling the different sub-problems separately. This is the main contribution of the paper in comparison to the existing literature, which includes several works focused on different aspects of the WDM system –mainly water demand forecasting. Since we aim to consider the overall problem in its entirety, the WDM system has been designed as a set of different kind of agents with complex interrelations among them. This Multi-Agent System (MAS), whose core is based on advanced forecasting techniques such as Artificial Neural Networks (ANNs), determines the optimal adjustment of pumping stations.

This research has emerged from the interest of the Water Company of Gijón[10], a municipality of 300,000 inhabitants in the north west of Spain, in these emerging management techniques. It should be highlighted that it is not a real application but it aims to show how AI techniques can be combined within a multi-agent framework in order to develop an IDDS for WDM.

This article is structured in five sections, including this introduction. Section 2 reviews the most relevant and recent literature on AI applications to WDM. Section 3 describes the MAS, with the different agents that form it and their purpose, and the structure and relationships between them. Section 4 presents and discusses the numerical results obtained after testing the system with water demand time series. Finally, section 5 concludes according to the stated objectives and defines future work lines.

2. Artificial Intelligence Applications for Water Demand Management

One of the great challenges in WDM is, undoubtedly, water demand forecasting. During the last decades of the 20th century, some statistical models were applied in this field. For example, Maidment and Miaou[11] used ARIMA methodology[12] to express daily water demand as a function of rainfall and air temperature in nine US cities. Some years later, An et al.[13] included other climatic factors within ARIMA models to forecast water demand. Shvartser et al.[14] integrated this methodology in a pattern recognition approach, explaining in detail the development of the model and evaluating it for an water supply system in Israel. Other studies[15] used these statistical methods on similar issues, such as power demand forecast.

From the beginning of this century, AI has been incorporated in WDM, since it is a highly complex problem conditioned by the interaction of multiple variables and developed in uncertain environments. AI can be defined as the discipline that builds processes that, when run on a physical structure, produce results that respond to the perceived inputs based on the stored knowledge. These techniques were mainly used to forecast water demand. Lertpalangsunti et al.[16] were pioneers and developed a forecasting system that used various classical AI tools. The study was conducted for the city of Regina (Canada) and achieved a great reduction in the forecasting error in comparison with statistical alternatives.

2.1. Machine Learning applied to water demand forecasting

In AI, Machine Learning is a subfield geared towards the creation of programs that generalize behaviors from unstructured information supplied as examples. The literature contains many Machine Learning applications to water demand forecasting.

Jain and Ormsbee[17] used ANNs to estimate the maximum weekly demand through information on some climatic factors (frequency and volume of rainfall and water temperature) and recent demand. Liu et al.[18] used the same technique on Weinan City (China). Nasseri et al.[19] developed a hybrid model of Genetic Algorithm and Kalman Filter for monthly water demand forecasting, with excellent results exclusively from previous demand data. Genetic Algorithms were also applied to other phenomena with similar characteristics[20]. Solomatine and Shrestha[21] incorporated Fuzzy Logic to the forecasting, while Tabesh and Dini[22] relied on it to predict water consumption in Tehran (Iran). The application of Support Vector Machines to the matter was studied by Msiza et al.[23], who compared its results with ANNs in different scenarios. Recently, Ponte et al.[24] developed an ANN-based system that significantly outperformed statistical techniques in the hourly forecasting and water demand. This system was used to tackle a classical problem of distribution networks, the Bullwhip Effect[25], which is not relevant in long-term management but was shown to act as a source of inefficiencies in real-time WDM. Other Machine Learning techniques that have been applied to water demand forecasting are Random Forest[26] and Multivariate Adaptive Regression Splines[27].

2.2. Multi-agent Systems in Water Demand Management

Distributed AI, another branch of AI, is oriented to the study of the necessary techniques for knowledge distribution and coordination, as well as the interactions between system and environment. In this sense, the system is designed as set of intelligent agents. An agent can be defined as a computer system capable of carrying out flexible and autonomous actions that affect their environment according to certain design goals. This multi-agent methodology allows one to tackle WDM from a new approach by integrating different breeds of agents in a more complex system, where to combine forecasting techniques with other elements. This way, the system can be directed towards another objective of greater amplitude.

The literature contains some applications of MAS in WDM. Moss and Edmonds[28] created an agent-based model, applied to the Thames basin, to analyze the effects on some social parameters on water demand. Athanasiadis et al.[29] developed a MAS that simulated customers' behavior to evaluate different pricing policies in Thessaloniki (Greece). Galán et al.[30] integrated social, urban dynamics, geography, and consumption models under a multi-agent framework, where they simulated different water demand scenarios for Valladolid (Spain). Zechman[31] used multi-agent modeling to evaluate different management strategies in water distribution systems. Giuliani and Castelletti[32] evaluated the interest of cooperation in large water resource systems. Ni et al.[33] first carried out a dynamic study of water quality based on a Q-Learning Algorithm, and subsequently[34] developed a MAS aimed at dynamic water quality assessment. Recently, Karavas et al.[35]

designed an energy management MAS for the design and control of autonomous polygeneration microgrids.

2.3. Overview of the bibliographic analysis

After the literature review, we summarize our main assessments:

i. The extensive literature brings evidence that AI is a suitable approach to address the WDM issue given its complexity and scale.

ii. Authors have used these methodologies (e.g. ANNs) mainly for consumption forecasting since it is one of the basic pillars in WDM. Small prediction errors can be reached without generalized differences among the various methods. The consideration of additional variables (e.g. climatic factors) improves the prediction but the need for the immediate availability of these data is a major limitation in real-time forecasting.

iii. Multi-agent methodology allows managers to study WDM from a broader perspective. It is possible to combine forecasting techniques with other elements (e.g. economic and social) to focus on goals of different nature (e.g. improve water quality or optimize transportation system).

Under these circumstances, this paper exhibits how AI techniques can be applied to WDM. Looking for the forecast that minimizes some error criterion can be understood as a partial solution for a larger identity problem. From a holistic approach, a MAS can be used in order to find the best overall solution. We have developed an IDSS aimed at optimizing real-time WDM, whose core are IA forecasting methods –recent demands are the only input, given the complexity of real-time availability of other data.

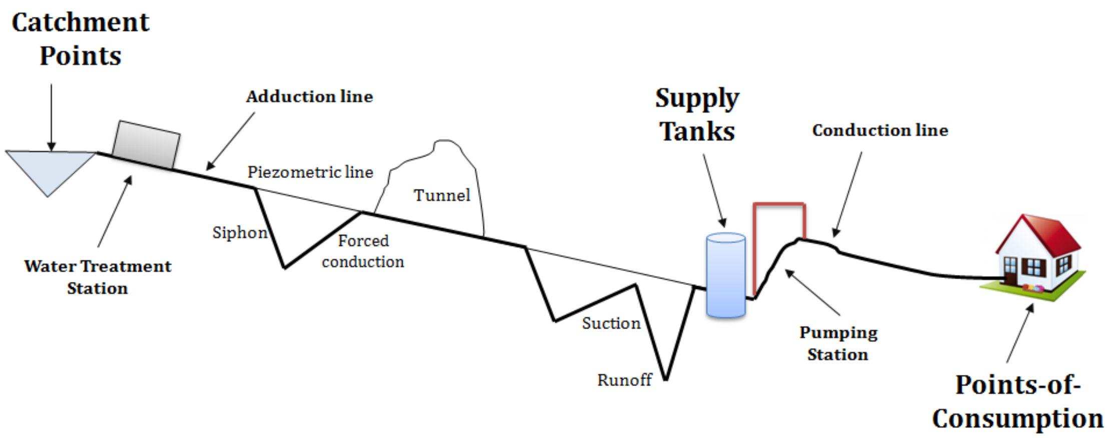


Figure 1. Overview of a water supply network.

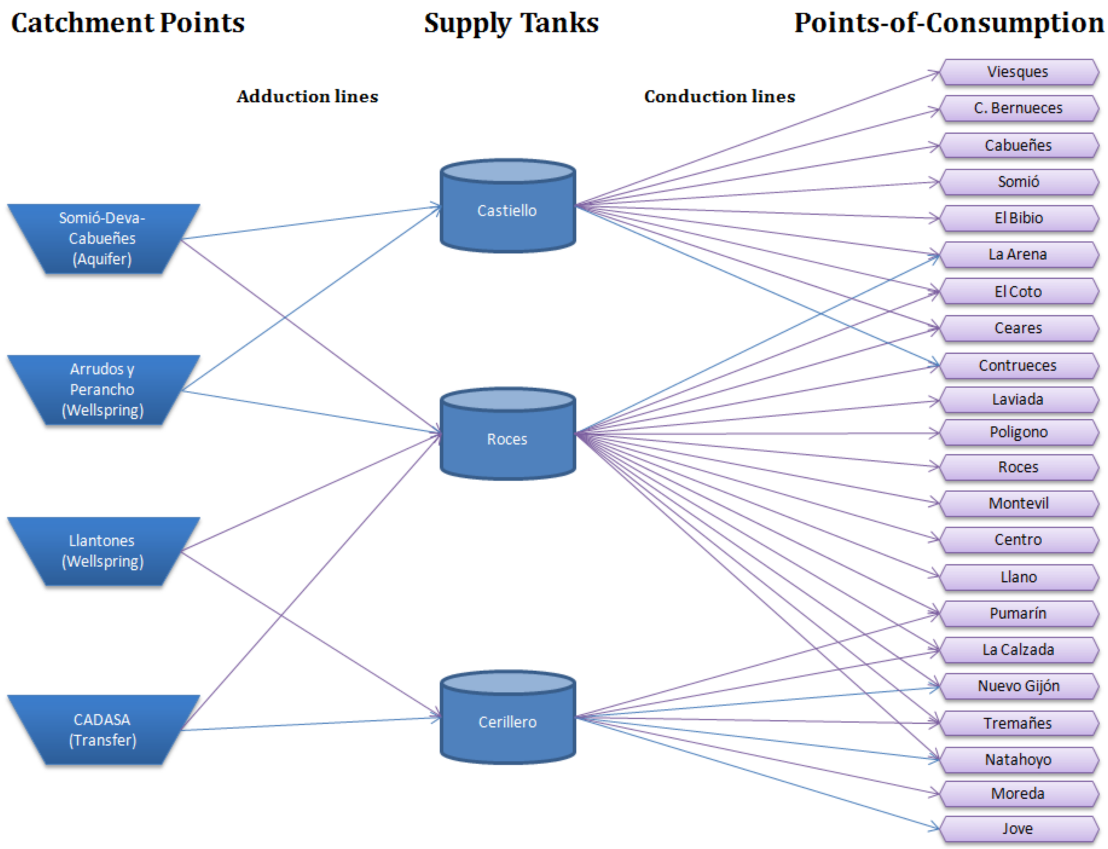


Figure 2. Basic structure of the water supply network of Gijón.

3. Description of the Intelligent Decision Support System

In order to design the MAS, we have considered a simple structure of a water supply network. In the upper level, there are various catchment points –usually natural sources, such as wellsprings and marshes fed into by rivers and groundwater. In the lower level, the points-of-consumption represent the distributed water demand. Between both levels, the supply tanks (storage reservoirs) are intermediate echelons that receive the water from the natural sources through the adduction lines and send it to the points-of-consumption through the conduction lines. The treatment station and the pumping station are also key elements. Figure 1 displays an overview of this water supply network. Figure 2 summarizes the structure of the water supply network of Gijon, which will be used as a basis in the development of the system.

Figure 3 shows an outline of the structure of the IDSS that has been designed, with its agents and the relationships between them and with the outside. Input data are the real-time water demands (from the water demand measurement system in the points-of-consumption) and the supply tanks level (from the measurement system in the supply tanks). The output is the best adjustment of pumping systems, namely the optimum quantity of water to be pumped hourly from the supply tanks.

Six breeds of agents can be identified in the MAS. The bidirectional relationships between them are represented in Figure 3. Below, we explain in detail the function of the different breeds of agents.

i. the Communication Agent,
ii. the Information Agent,
iii. the Water Demand Forecasting Agent,
iv. the Scenarios Simulator Agent,
v. the Cost Evaluation Agent,
vi. the Pumping Planning Agent.

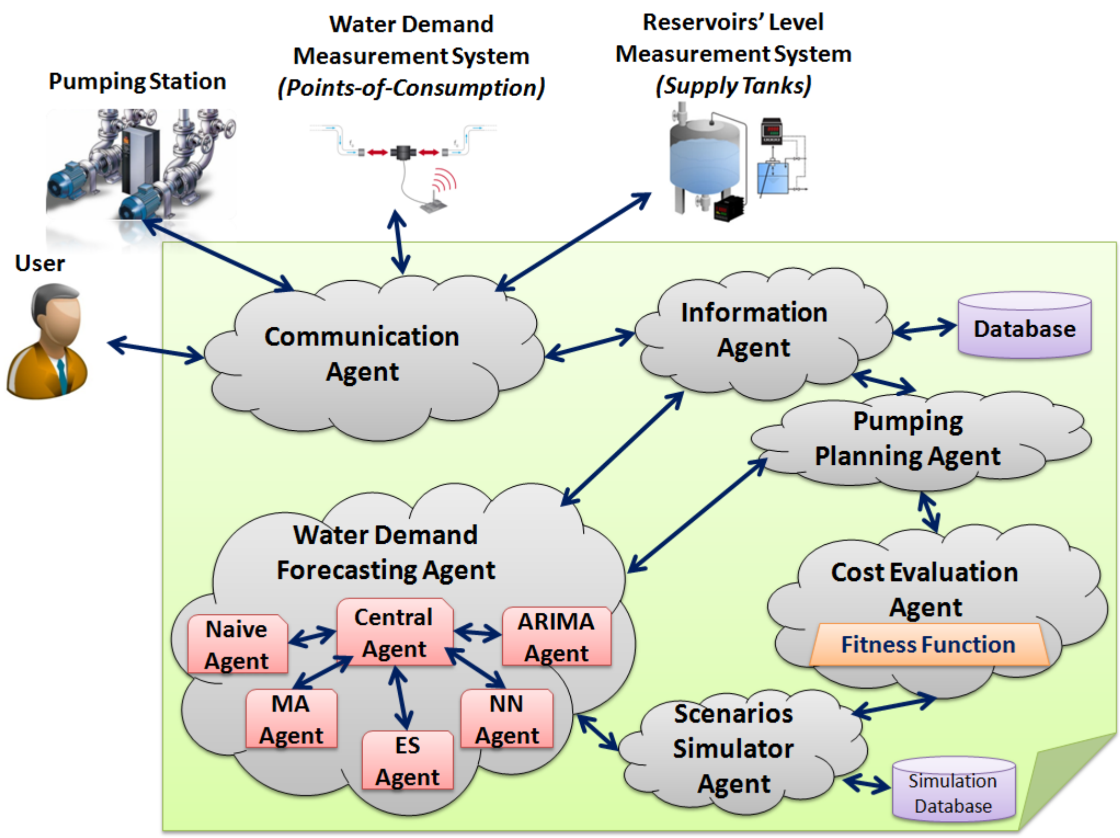


Figure 3. Outline of the IDSS that has been developed in this research work.

3.1. Communication Agent

The Communication Agent holds the interactions of the MAS with the outside. It operates in a quadruple way, as it communicates the MAS with:

- The water demand measurement system. Hence, hourly water demands are continuously stored in the database. This enables the development of reliable forecasts in real time.
- The reservoirs level measurement system. Thereby, the current level of the different supply tanks is continuously known and stored in the database, which has influence on the water to be pumped.
- The pumping stations, which operate according to the Pumping Planning Agent (it hourly determines the amount of water to be pumped).
- The user through an interface. The interface allows the user to introduce information that could alter the ordinary system operation (e.g. changes in the cost model or in the system's constraints), as well as to see the most relevant information (e.g. consumption, forecasts, and costs).

3.2. Information Agent

The database associated to the Information Agent stores hourly data related to the WDM system, so that it is available for the other agents. In particular, it saves information on: (1) hourly water demand to date; (2) the most reliable demand forecasts to date; (3) water stored in supply tanks to date; and (4) hourly water pumped to date.

Thus, the main objective of the Information Agent is mediation between the database and the other agents, both storing data and responding to information requirements. Hence, the other agents do not see a database but another agent, and the MAS reaches the essential homogeneity.

3.3. Water Demand Forecasting Agent

The five forecasting agents are the real core of the MAS. Each one of them estimates the hourly water demand according to a predetermined method. Three simple forecasting methods (naive model, moving averages, and exponential smoothing), a complex statistical method (ARIMA models), and an AI-based tool (ANNs) are used. These calculate the forecasts using historical consumption stored in the database.

Naïve Agent

The Naive Agent performs the forecast using a naive model, which estimates the hourly demand ($D^{\wedge}_t$) as the demand in the previous hour (D_{t-1}) adjusted by the increase (or decrease) in the demand in the same time interval of the previous week (D_{t-168}-D_{t-169}) by Eq. (1).

$$\widehat{D}_t = D_{t-1} + (D_{t-168} - D_{t-169}) \tag{1}$$

MA Agent

The MA Agent forecasts using a n-order moving average. It estimates the hourly demand (D^t) as the arithmetic average of the last n demands (Dt-i, i∈[1,n]). Previously, a 2nd-order differentiation (with the operator Δ) of the time series must be performed with the aim of eliminating trend and seasonality as the simple moving average method must not be applied for series with these features. The differentiation process is given by Eq. (2) and Eq. (3), while the forecast is based on Eq. (4) and requires undoing the differentiation process. The MA Agent calculates the forecast from n=1 (1 hour) to n=168 (1 week).

$$\Delta D_t = D_t - D_{t-1} \tag{2}$$

$$\Delta^2 D_t = \Delta D_t - \Delta D_{t-1} \tag{3}$$

$$\widehat{\Delta^2 D_t} = \frac{1}{n}\sum_{i=1}^{n} \Delta^2 D_{t-i} \tag{4}$$

Among all forecasts, the MA Agent selects the optimal one according to the Mean Absolute Percentage Error (MAPE) criterion36, which is expressed by Eq. (5) where m is the time horizon.

$$MAPE = \frac{1}{m}\sum_{t=1}^{m} \left|\frac{D_t - \widehat{D}_t}{D_t}\right| \tag{5}$$

ES Agent

The ES Agent forecasts according to a simple exponential smoothing with seasonality, i.e. a weighted average of the recent forecasts and the error in the same interval. Therefore, it estimates the hourly demand ($D^{\wedge}_t$) as the sum of a level function associated with the last hour (L_{t-1}) and a seasonal function associated with the demand on the same day and hour of the week before (S_{t-168}) by Eq. (6). The level function depends on the linear smoothing coefficient (α) while the seasonal function depends on the seasonality coefficient (δ), as it can be seen in Eq. (7) and (8). Hence, the ES Agent evaluates different values of coefficients of linear smoothing and seasonality, seeking to minimize the MAPE.

$$\widehat{D}_t = L_{t-1} + S_{t-168} \tag{6}$$

$$L_{t-1} = \alpha(D_{t-1} - S_{t-169}) + (1-\alpha)L_{t-2} \tag{7}$$

$$S_{t-168} = \delta(D_{t-168} - L_{t-168}) + (1-\delta)S_{t-169} \tag{8}$$

ARIMA Agent

The ARIMA Agent estimates the hourly demand using an autoregressive integrated moving average model. These models can be synthesized according to $[(p,d,q)(P,D,Q)_n]$ where p (P) is the order of the autoregression, d (D) is the order of differentiation and q (Q) is the order of the moving average. The lowercase variables are not-seasonal components, while the uppercase ones are seasonal with periodicity n. These models consider that the future value of the differentiated variable ($\Delta^d D^{\wedge}_t$) can be expressed as a function of past observations (D_{t-i}, $i\in[1,n]$) and a random error (ε_{t-j}, $i\in[1,q]$), by Eq. (9), where Δ is the differentiation operator, γ is the constant model, φ_i are the parameters associated with autoregression, and θ_j are the parameters associated with the moving average. It should be noted that it is also necessary to eliminate the differentiation.

$$\widehat{\Delta^d D_t} = \gamma + \sum_{i=1}^{p} \varphi_i \, \Delta^d D_{t-i} + \sum_{k=1}^{P} \varphi_{kn+p} \, \Delta^d D_{t-(kn+p)} - \sum_{j=1}^{q} \theta_j \, \varepsilon_{t-j} - \sum_{m=1}^{Q} \theta_{mn+q} \, \varepsilon_{t-(mn+q)} \tag{9}$$

The method of obtaining the statistical model $[(p,d,q)(P,D,Q)_n]$ associated with each time series is based on the sequential process of:

i. identifying the possible model,
ii. parameter estimation,
iii. validation.

It should be highlighted that data from the last six weeks (1,008 hourly demands) are used in the forecasting. This process[12] is repeated until both the model is verified through their autocorrelation functions and its forecasts are validated by a given error criterion. In our case, the ARIMA Agent seeks the model that best fits the input time series, using the following statistics for the comparison of the different proposed models:

- goodness-of -fit according to the MAPE criterion,
- residual simple autocorrelation function,
- residual partial autocorrelation function.

NN Agent

The NN Agent performs forecasting through ANNs with three levels: an input layer (predictor variables), a hidden layer, and one output neuron (variable to predict). The basic elements are the neurons in the hidden layer. Each one of them receives a number of inputs via interconnections and emits an output, which can be identified by three functions:

- a propagation or excitation function, which consists of the sum of each input by its interconnection weight,
- an activation function, which modifies the former,
- a transfer function, which is applied to the value returned by the above one and that limits the output.

Mathematically, the hourly demand forecast at each period ($D^{\wedge}_t$) is expressed by Eq. (10), where y_t represents the output (forecast), f_{outer} represents de output layer, f_{inner} represents the input layer transfer function, w_{xy} represents the weights and biases ($i \in [1,17]$ refers to the input neurons and $j \in [1,n]$ refers to the hidden neurons) and $^{(z)}$ represents the z-th layer.

$$\widehat{D}_t = f_{outer}\left[\sum_{j=1}^{n} w_{1j}{}^{(2)} \cdot f_{inner}\left(\sum_{i=1}^{17} w_{ji}{}^{(1)} \cdot x_i + w_{j0}{}^{(1)}\right) + w_{10}{}^{(2)}\right] \qquad (10)$$

Figure 4 outlines the designed ANN. We introduced 17 predictor variables (input neurons) in the NN Agent as the time series (see section IV) shows double periodicity: daily and weekly. The variables are: the day of the week (Day); the hour of the day (Hour); the four immediately preceding hourly demands (t-1 to t-4); the hourly demand of the day before at the same hour and the four immediately preceding demands (t-24 to t-28); the hourly demand of the previous week on the same day and the same hour and the four immediately preceding demands (t-168 to t-172); and an additional binary variable that differentiates holidays and working days (Holiday). The number of neurons in the hidden layer is a decision variable to optimize. The output neuron is related to the variable to predict: the hourly water demand. It should be clarified that we have looked for the best structure in terms of forecasting reliability, execution time (which is a significant limitation in a real-time system), and avoiding the common 'overfitting' problem (i.e. memorizing instead of learning).

The steps for developing the ANN-based system are similar to those detailed in Ref. 37, which can be considered as the preliminary step to this research work. This article forecasts hourly water demand comparing two different ANN architectures: multi-layer perceptron and radial basis functions, and concluded that the first structure tends to offer better performance. For this reason, a multi-layer perceptron has been used. The data available for each forecast, i.e. the last 6 weeks (1,008 hourly demands), were separated randomly into two groups. The 70% was oriented to the batch training of the network through the back-propagation algorithm. The remaining 30% has been used to validate the network. The following stopping criteria were defined:

- max number of steps without reducing error: 1000,
- max workout time: 1 minute.
- min relative change in training error: 0.0001,
- min relative change in error rate training: 0.001.

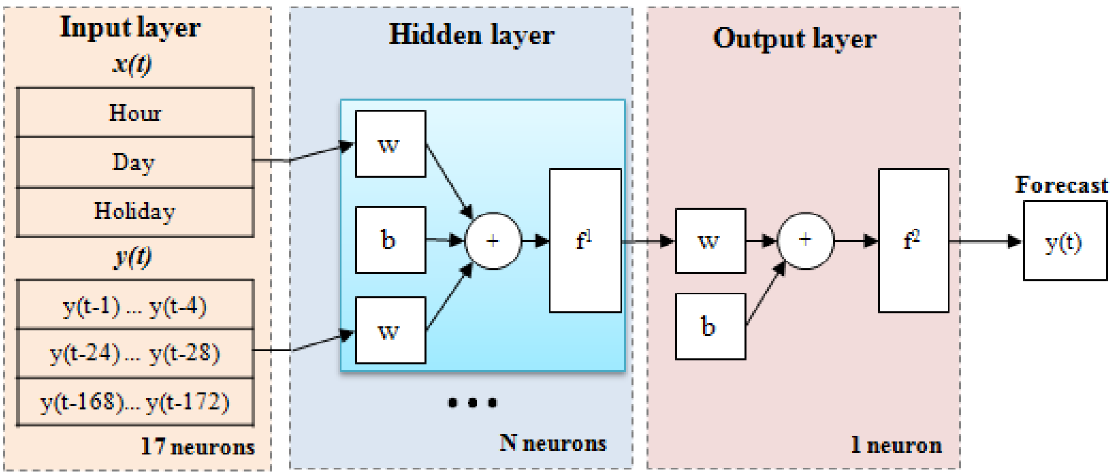


Figure 4. General outline of the ANN-based system.

Central Agent

The optimal forecast selected hourly by each one of the five agents is sent to the Central Agent. These hourly demand forecasts are stored in the database through the Information Agent. Moreover, the Central Agent sends these forecasts to the Planning Pumping Agent. Therefore, the Central Agent only acts as interconnection between the Forecasting Agents and the other agents.

3.4. Scenarios Simulator Agent

In WDM systems, water pumping is performed at specific times whose frequency varies greatly from one city to another. The new context stresses the importance of continuous pumping in order to reduce operating costs. The Scenarios Simulator Agent uses information stored in the database to perform a simulation of the last 24 hours in different scenarios defined by the forecasts transmitted by the Central Agent. These will be studied by the Cost Evaluation Agent that seeks the one that minimizes WDM costs.

The assumptions that we have made in the development of this simulator are the following:

- fixed supply time: 1 hour (both, on the one hand, from natural sources to supply tanks and, on the other hand, from these to points-of-consumption),
- unconstrained catchment, storage and transportation systems,
- water is pumped to the supply tanks in order to store at the beginning of each hour the quantity that has been forecast,
- when there is risk of shortage, the pumping is carried out urgently at a higher cost,
- water cannot be returned to the previous echelon.

In order to determine the water stored at the beginning of each hour in Supply Tanks (IW_t), this agent adds the water stored at the end of the last hour (FW_{t-1}) and the water pumped during that time (WP_{t-1}), by Eq. (11). The water stored at the end of each hour in Supply Tanks (FW_t) is expressed as the difference between water stored at the beginning of this hour (IW_t) and the hourly demand (D_t), except if this value is negative. In that case, the emergency water pumping (EWP_t) would be carried out, and the water stored would be zero. It is expressed by Eq. (12) and (13). Finally, the water pumped in each period (WP_t), a process that is supposed to be done at the end of it, is the difference between the demand forecast for the next period ($\hat{D}_{t+1}$) and the final status of tank (FW_t), if this value is greater than zero, according to Eq. (14).

$$IW_t = FW_{t-1} + WP_{t-1} \tag{11}$$

$$FW_t = \max\{IW_t - D_t, 0\} \tag{12}$$

$$EWP_t = \max\{D_t - IW_t, 0\} \tag{13}$$

$$WP_t = \max\{\widehat{D}_{t+1} - FW_t, 0\} \tag{14}$$

The operational logic of the simulation system is illustrated in Figure 5. It should be noted that there are two main flows: the downstream water flow, from natural sources to points-of-consumptions and constrained by the lead time (supply time), and the upstream demand flow, in the opposite direction.

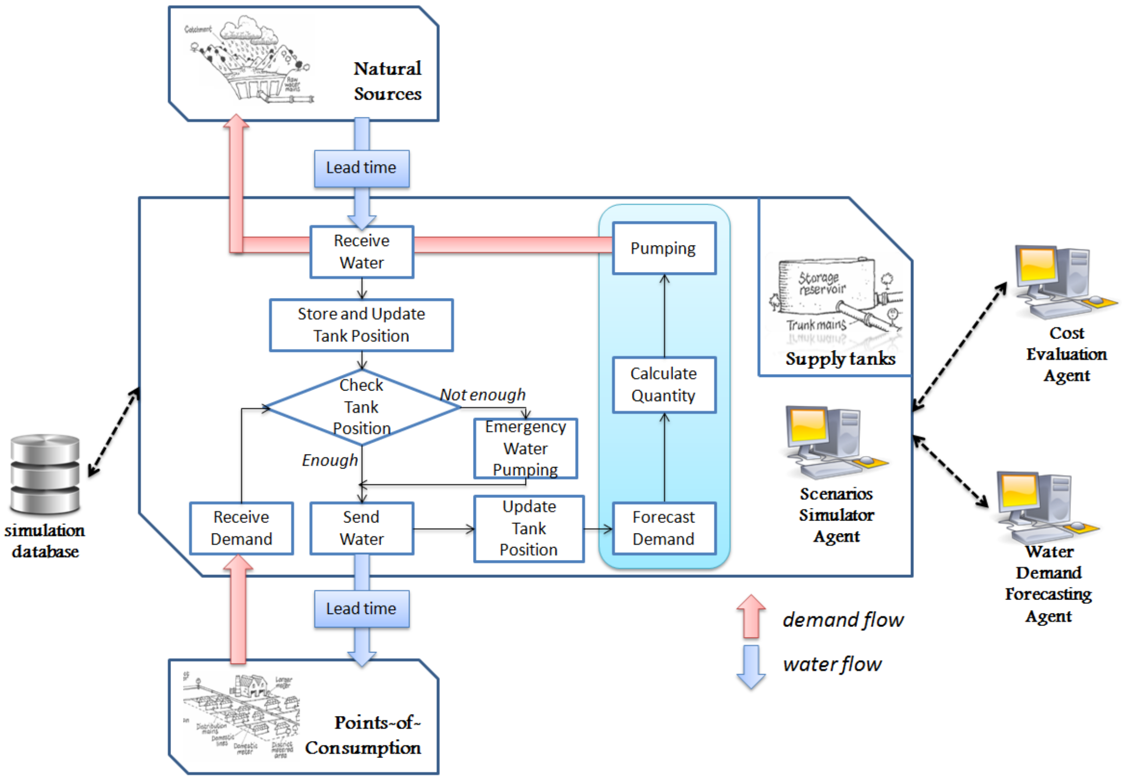


Figure 5. Overview of the simulation model, by means of a flow chart.

3.5. Cost Evaluation Agent

The Cost Evaluation Agent hourly decides which one of the five scenarios presented by the Scenarios Simulation Agent minimizes WDM costs and, therefore, it is optimal for managing the system at that time. In the simple structure previously defined and according to the stated assumptions, we consider that there is a cost associated with pumping, treatment and storage of water, which depends entirely on demand. That is to soy, water must be transported to the demand points, which involves some costs that are independent of the planning. Moreover, there are two cost overruns (both can be expressed as a ratio of a unit of currency to a unit of volume of water), which are inputs for the system:

- An additional cost associated with emergency water pumping (c_{ewp}), defined as the difference between the emergency water pumping cost and the scheduled water pumping cost.
- An additional cost related to excessive storage of water in Supply Tanks (c_{fw}), in relation to tanks capacity and supply problems resulting from excessive storage of water, i.e. an opportunity cost.

Thus, the Cost Evaluation Agent determines in each scenario the WDM cost overrun (WMO), which arises from the errors in the forecast. It can be estimated as the sum of the cost overrun associated to the emergency water pumping (WMO_{ewp}) and the cost overrun associated to the excessive storage of water (WMO_{fw}) throughout the simulation time interval (m). The first one is the product of emergency water pumped (EWP_t) and its associated cost overrun (c_{ewp}). The second one is the product of the quantity of water stored at the end of each interval (FW_t) and their additional costs associated storage (c_{fw}). Hence, the fitness function to be minimized is expressed by Eq. (15) to (17).

$$WMO_{ewp} = \sum_{t=1}^{m} EWP_t \cdot c_{ewp} \qquad (15)$$

$$WMO_{fw} = \sum_{t=1}^{m} FW_t \cdot c_{fw} \qquad (16)$$

$$WMO = WMO_{ewp} + WMO_{fw} \qquad (17)$$

3.6. Pumping Planning Agent

The Pumping Planning Agent uses the optimal scenario selected by the Cost Evaluation Agent to choose the forecast which must be used to adjust the pumping system. Hence, it determines the water to be pumped hourly by Eq. (14), but considering the position level of the supply tanks (instead of the simulation calculations) leading to real-time WDM. These data are stored in the database via the Information Agent, and they are carried to the Pumping Station through the Communication Agent.

4. Results and Discussion

In order to test the IDSS, twelve time series with 1,032 data (WD01 to WD12) have been used. Each one of them contains hourly water demands of 43 consecutive days in the city of Gijón (Spain).

These time series have been created through simulation using the monthly water demand of this city, a distribution model of water demand along each week, and random parameters to introduce different sources of uncertainty. It should be clarified that the average hourly demand in the township in 2012 and 2013 was 2,455.44 m^3/hour, while we have used the model that, according to Ref. 9, best fits the urban consumption in the city of Valencia (Spain).

Within each series, we have used 97.7 % of the data (1,008 hourly demands corresponding to 6 weeks) for training of forecasting methods and determining the optimal alternative for the pumping systems adjustment through simulation scenarios, while the remaining 2.3% (24 time demands, corresponding to 1 day) was used to test the system with the solution provided. From that point on, we have analyzed the reduction achieved in management costs.

The time series we have chosen span training and testing periods of very different nature. Table 1 contains the relevant information about the training period (first and last day) and testing period (day and hour of beginning and end). Note that in all cases the testing period begins when the training period ends. Six series have been chosen as working days (see [1]), which would be the usual case in the practical implementation of the system. Three series correspond to weekend days (see [2]). The remaining three are related to holidays (i.e. holidays or days around them, see [3]), since we aim to evaluate the effectiveness of the developed application in these special cases.

The twelve time series show a similar structure. These hourly time series displays double seasonality and trend. On the one hand, there is a daily frequency (each 24 hours). There is a night sharp decrease from 19h until 02h, when demand stabilizes around a daily minimum until 06h. At that time, demand grows strongly until 11h, when it sets a first local maximum. From there, demand undergoes a slight decline to set a local minimum at 14h, and then it begins to increase until the establishment of a second local maximum at 19h. Note that the previously given hours are approximate and vary according to the season of the year. On the other hand, there is a weekly frequency, namely the structure is similar each 7 days (168 hours). During weekends, a significant decline in consumption can be observed –first on Saturdays, and even larger on Sundays, where the morning cycle is especially small compared to the afternoon one. In addition, these time series do not remain in a constant range, but they show a significant trend throughout the year, both in average and variance, which must be considered in the forecasting process.

Figure 6, which has been included to illustrate the explanation, shows the training and testing periods for the time series WD03. Notice the decreasing trend from weeks 2 to 6, which seems to be reversed during the last week. Besides, the holidays in weeks 3 and 6, when demand drops considerably, greatly influence the forecast –especially the Friday of week 6, given the weekly frequency of the series, makes complicated the forecast.

Table 1. Training and testing period for the time series.

Test	*Training period*		*Testing period*		*Testing day*
	From	*To*	*From*	*To*	*Kind*
WD01	28/06/12	09/08/12	Thur 14 h	Frid 13 h	[1]
WD02	08/02/13	19/03/13	Tues 03 h	Wedn 02 h	[1]
WD03	28/09/13	09/11/13	Frid 11 h	Satu 10 h	[3] (*)
WD04	18/02/212	30/03/12	Satu 21 h	Sund 20 h	[2]
WD05	15/08/12	26/09/12	Wedn 16 h	Thur 15 h	[1]
WD06	12/09/12	31/10/12	Mond 09 h	Tues 08 h	[1]
WD07	09/11/12	22/12/12	Sund 00 h	Sund 23 h	[2]
WD08	04/01/12	15/02/12	Wedn 04 h	Thur 03 h	[1]
WD09	23/02/13	06/04/13	Satu 13 h	Sund 12 h	[2]
WD10	02/04/12	14/05/12	Mond 22 h	Tues 21 h	[1]
WD11	15/05/12	26/06/12	Tuesd 16 h	Wedn 15 h	[3] (**)
WD12	24/10/12	06/12/12	Tues 06 h	Wedn 05 h	[3]

Notes: (*) It corresponds to a week after a holiday on Friday; (**) It corresponds to a week after a holiday on Tuesday.

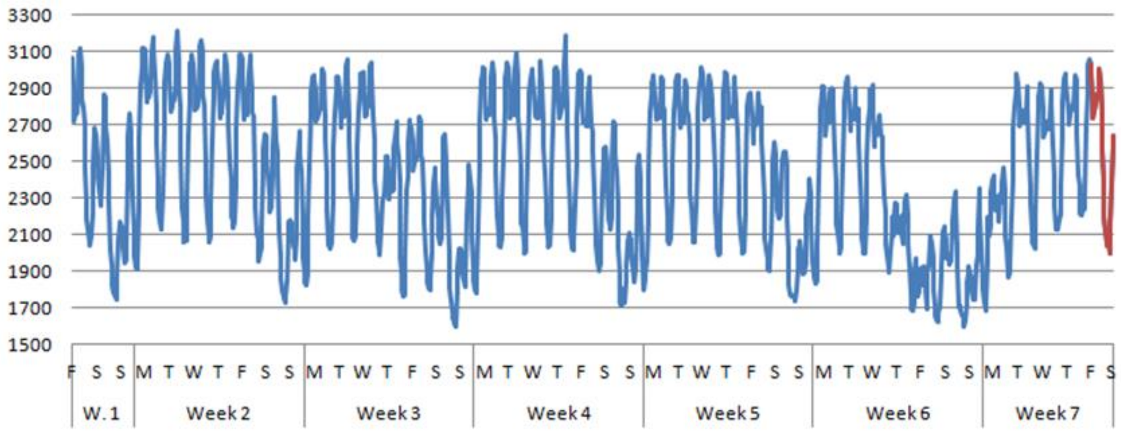


Figure 6. Training period and testing period for the time series WD03 (values in cubic meters).

4.1. Time series forecasting

The reduction in WDM costs is based on the accuracy of the system forecasts. Table 2 shows for each series the best result of the five forecasting agent. It contains the following information: the name of the time series (Time series); the agent that forecasts (Forecasting); the optimal features of the method that minimize the MAPE during the training period, i.e. the time horizon if it is a moving average, the linear smoothing and seasonality coefficient in the case of exponential smoothing, the ARIMA model that best fits the input data for the Box Jenkins methodology, and the optimal structure of the ANN through the neurons in each layer (Features); and the MAPE of the forecast calculated on the 24 testing data (MAPE). We have stood out in bold the agent which achieves a minimum error for each series.

Table 2 compares the results of the five agents with a base method, which distinguishes three kinds of days: regular working days (Monday to Friday except holidays and eve of holidays), holidays (including Sundays), and eve of holidays (including Saturdays). Thus, this base method estimates the hourly demand on any day as the demand in the previous day of the same kind. The results show that in most cases this method achieves small errors due to the regular nature of the studied time series.

Table 2 highlights that the Naive Agent achieves in all cases forecast errors lower than 5%. It is especially efficient in forecasting working days with an error lower than 2.5% in series of this type –and sometimes even below 1%. This model is fairly easy since it greatly simplifies the series operation; nonetheless its performance is positive in view of the results. Notice that in nine of the twelve cases, the Naive Agent provides better results than the other two simple methods of forecasting (exponential smoothing and moving averages), although the theoretical foundation of these others is more complex.

Table 2. Results of the time series forecasting.

Time series	*Forecasting. Agent*	*Features*	*Fitness MAPE*
WD01	Base Method	-	1.11%
	Naive Agent	-	1.25%
	MA Agent	5	2.81%
	ES Agent	α=0,8 ; δ=2,8·10-5	3.72%
	ARIMA Agent	(0,1,13)(1,1,0)168	1.53%
	NN Agent	17-6-1	0.92%
WD2	Base Method	-	2.73%
	Naive Agent	-	2.39%
	MA Agent	5	3.81%
	ES Agent	α=0,7 ; δ=2,1·10-5	4.48%
	ARIMA Agent	(0,1,14)(0,1,1)168	2.26%
	NN Agent	17-8-1	1.54%
WD03	Base Method	-	10.08%
	Naive Agent	-	3.00%
	MA Agent	5	3.07%
	ES Agent	α=1 ; δ=3,3·10-5	4.92%
	ARIMA Agent	(0,1,5)(0,1,1)168	5.34%
	NN Agent	17-7-1	2.44%
WD04	Base Method	-	3.77%
	Naive Agent	-	2.14%
	MA Agent	4	3.25%
	ES Agent	α=0,6 ; δ=4,4·10-5	2.83%
	ARIMA Agent	(0,1,14)(0,1,1)168	2.14%
	NN Agent	17-7-1	2.19%
WD05	Base Method	-	1.69%
	Naive Agent	-	1.77%
	MA Agent	5	3.63%
	ES Agent	α=0,5 ; δ=6,0·10-6	1.87%
	ARIMA Agent	(0,1,6)(0,1,1)168	1.77%
	NN Agent	17-9-1	1.21%
WD06	Base Method	-	9.43%
	Naive Agent	-	1.59%
	MA Agent	5	3.63%
	ES Agent	α=0,8 ; δ=5,8·10-6	1.60%

	ARIMA Agent	(0,1,5)(0,1,0)168	2.00%
	NN Agent	16-9-1	1.91%
WD07	Base Method	-	3.52%
	Naive Agent	-	3.76%
	MA Agent	3	3.90%
	ES Agent	α=0,6 ; δ=2,4·10-5	3.89%
	ARIMA Agent	(0,1, 3)(0,1,1)168	3.26%
	NN Agent	17-6-1	2.87%
WD08	Base Method	-	1.87%
	Naive Agent	-	1.86%
	MA Agent	5	3.37%
	ES Agent	α=0,6 ; δ=3,2·10-5	1.60%
	ARIMA Agent	(0,1, 3)(1,1,0)168	2.04%
	NN Agent	17-6-1	1.43%
WD09	Base Method	-	3.05%
	Naive Agent	-	1.78%
	MA Agent	2	2.78%
	ES Agent	α=0,8 ; δ=1,2·10-5	2.09%
	ARIMA Agent	(1,1,0)(0,1,1)168	2.32%
	NN Agent	17-9-1	2.32%
WD10	Base Method	-	3.43%
	Naive Agent	-	0.82%
	MA Agent	5	3.18%
	ES Agent	α=0,8 ; δ=3,7·10-6	1.87%
	ARIMA Agent	(1,1,0)(0,1,1)168	1.96%
	NN Agent	17-5-1	0.72%
WD11	Base Method	-	2.63%
	Naive Agent	-	3.79%
	MA Agent	5	3.48%
	ES Agent	α=0,8 ; δ=4,3·10-5	5.70%
	ARIMA Agent	(2,1,12)(0,1,1)168	9.76%
	NN Agent	17-7-1	2.51%
WD12	Base Method	-	8.31%
	Naive Agent	-	4.28%
	MA Agent	4	2.67%
	ES Agent	α=1 ; δ=0	14.82%
	ARIMA Agent	(0,1,0)(1,1,0)168	16.04%
	NN Agent	17-6-1	2.96%

Moving averages generate errors between 2.5% and 4% in all series, showing a more robust performance as they are less sensitive to the type of day than other methods, which becomes an advantage when the forecast is complex. For this reason, it offers interesting solutions on holidays or days around them. For example, in WD12, the moving average achieves the lowest error, even better than ANNs. However, results obtained by this agent are greatly improved by other methods in working and weekend days.

The ES Agent provides similar results to the Naive Agent, although slightly worse in most cases. It is capable of achieving good performance in forecasting working days (e.g. WD06, WD08 and WD10 with a MAPE lower than 2%), but it is not reliable in forecasting holidays (e.g. the MAPE is about 15% in WD12). In addition, it does not offer a good performance in days around holidays as these days significantly modify the model of the series, and hence decreasing the accuracy of forecasts.

The ARIMA models have the same deficit as the exponential smoothing on holidays and days around them, which can be justified from the same perspective. However, the ARIMA Agent usually improves the results of the ES Agent on working days and weekends, being able to understand very precisely the trend and seasonality of the series, with errors less than 2.5%, except in the WD07 series in which its results are only enhanced by ANNs. In WD04, the ARIMA Agent achieves the lower forecast error.

The results from the four methods analyzed so far are considerably outperformed by the NN Agent, which achieves the smallest error in 8 out of the 12 cases. The network built by this agent can explain very precisely the past of the series and makes very accuracy forecasts for the future. Even when forecasting the demand is difficult and the other agents do not offer precise results, the NN Agent responds with reliable forecasts. As expected, this fact brings evidence that the incorporation of AI to the model increases the confidence in forecasts, because they make the system trained for understanding unexpected changes in trends and they deal appropriately with the seasonality.

By way of illustration, Figure 7 shows the testing period for the WD03 series, as well as the forecasts of the Naive (3.00% MAPE), ARIMA (5.44% MAPE) and NN Agents (2.44% MAPE). The ANNs generate the best approximation. Notice that the ARIMA method is not capable of accomplishing a good result, since the holiday just a week before the testing day acts as a source of error. Figure 8 shows the same information for the

WD05 series that corresponds to a working day (1.77% MAPE for the naive model and ARIMA techniques and 1.21% for the ANNs). It can be noted that in this case all forecasts are considerably more accurate. Again, the NN Agent offers the best performance.

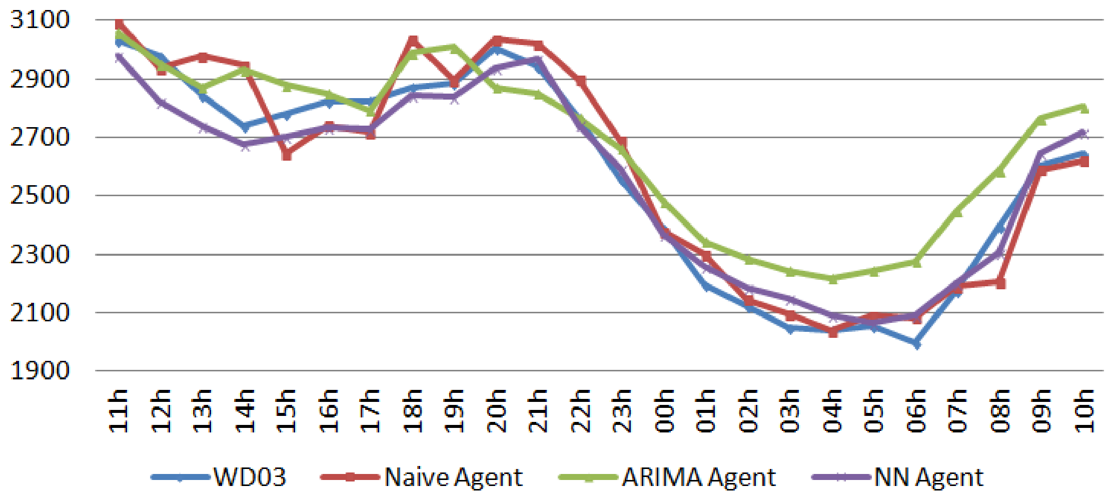


Figure 7. Testing period of the time series WD03: hourly water demand and forecasts (values in cubic meters).

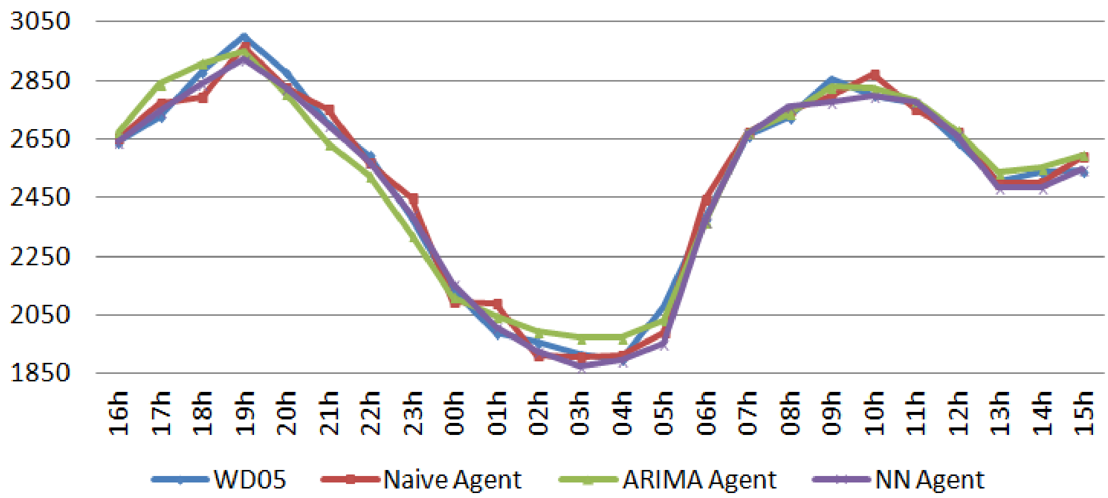


Figure 8. Testing period of the time series WD05: hourly water demand and forecasts (values in cubic meters).

4.2. Cost overrun reduction

The developed IDSS requires the introduction as an input of the unit cost overruns related to excessive storage and emergency water pumping. However, the really significant in terms of the solution provided by the system is the relationship between both according to Eq. (17). For this reason, the study of each series has been divided into three scenarios:

i. Scenario 1: when both are equal (c_{ewp}=2 uc; c_{fw}=2 uc),

ii. Scenario 2: when the excessive storage overrun cost is three times the emergency water pumping overrun cost (c_{ewp}=1 uc; c_{fw}=3 uc),

iii. Scenario 3: when the relationship is the opposite (c_{ewp}=3 uc; c_{fw}=1 uc).

Tables 3, 4 and 5 comprise the economic results of the tests in the three scenarios. That is to say, for each series, the costs of the solution provided by the MAS are shown. In order to compare the results, we have taken as a reference the solution provided by the base method. This tables contains the following data: the name of the series (Time Series); the WDM Cost Overrun in the testing period if the pumping system adjustment is made according to the base method (Cost Overrun Base Method); the forecasting agent that minimizes the overrun cost with the training data (Forecasting Agent); the WDM cost overrun in the testing period with the solution provided by the MAS (Cost Overrun MAS); and the percentage reduction achieved in comparison with the base method (Reduction Over Base Method).

Table 3. Results of the MAS oriented to minimizing costs (in units of currency) for Scenario 1.

Time series	*Cost Overrun Base Method*	*Forecasting Agent*	*Cost Overrun MAS*	*Reduction over Base Method*
WD01	1,094	NN	872	*20.29%*
WD02	2,730	NN	1,730	36.63%
WD03	10,022	NN	2,740	72.66%
WD04	3,338	Naive	1,872	43.92%
WD05	1,854	NN	1,416	23.62%
WD06	10,912	Naive	1,688	84.53%
WD07	2,906	NN	2,654	8.67%
WD08	2,276	NN	1,548	31.99%
WD09	2,844	Naive	1,718	39.59%
WD10	3,816	NN	910	76.15%
WD11	3,312	NN	3,230	2.48%
WD12	9,122	MA	2,786	69.46%

Results demonstrate the high efficiency of the system. The MAS for real-time WDM can achieve large reductions in cost overrun in comparison with the results obtained if the base method is used to adjust the pumping systems. In 32 of the 36 total tests, the system achieves a reduction in costs, and only in the remaining 4, the solution provided by the system would cause a higher cost. The average reduction is 42.50% for the first scenario, 37.53% for the second one, and 39.23% in the third scenario.

Table 4. Results of the MAS oriented to minimizing costs (in units of currency) for Scenario 2.

Time series	*Cost Overrun Base Method*	*Forecasting Agent*	*Cost Overrun MAS*	*Reduction over Base Method*
WD01	989	NN	720	27.20%
WD02	2,971	NN	1,967	33.79%
WD03	5,017	NN	2,394	52.28%
WD04	1,872	Naive	2,026	-8.23%
WD05	2,577	NN	996	61.35%
WD06	7,136	Naive	1,574	77.94%
WD07	2,353	ARIMA	3,293	-39.95%
WD08	3,036	ES	1,099	63.80%
WD09	3,358	Naive	1,659	50.60%
WD10	2,884	NN	767	73.40%
WD11	3,008	NN	3,665	-21.84%
WD12	13,386	MA	2,675	80.02%

Table 5. Results of the MAS oriented to minimizing costs (in units of currency) for Scenario 3.

Time series	*Cost Overrun Base Method*	*Forecasting Agent*	*Cost Overrun MAS*	*Reduction over Base Method*
WD01	1,199	ARIMA	984	17.93%
WD02	2,489	NN	1,493	40.02%
WD03	15,027	ES	2,826	81.19%
WD04	2,640	ARIMA	1,267	52.01%
WD05	1,131	ARIMA	1,822	-61.10%
WD06	14,688	ES	1,156	92.13%
WD07	3,459	NN	1,925	44.35%
WD08	1,516	NN	1,312	13.46%
WD09	2,330	NN	1,415	39.27%
WD10	4,748	Naive	819	82.75%
WD11	3,616	NN	2.795	22.70%
WD12	4,561	NN	2,460	46.07%

Moreover, these tables show that the choice of the forecasting method that results in the optimal alternative to adjust the pumping equipment varies according to the relationship between the cost overruns. For example, in WD03 series, ANN-based forecast generates the optimal setting in the first two cases, while in the latter the best forecast is provided by the ES Agent. Ten of the twelve series (all series except the WD02 and the WD11) clearly show this idea. Analyzing the previous tables in more detail (comparing it with Table 2), it can be observed that when the performance of a forecasting method is clearly superior to the others, the system tends to resort to it in order to minimize costs. However, when the difference is not very significant, choosing the best alternative to adjust the pumping depends on the ratio of costs. In these ten cases, an intermediate ratio could be found that differentiate the case when some forecasting method is optimal and the case when another is more appropriate.

These results confirm that limiting the study of WDM to the forecasting of this variable only implies finding a partial solution to the problem, which does not always lead to the best overall solution. It must be highlighted that WDM is a complex problem that should be understood as a whole and not as a collection of parts, and hence multi-agent techniques draw an appropriate framework to deal with it.

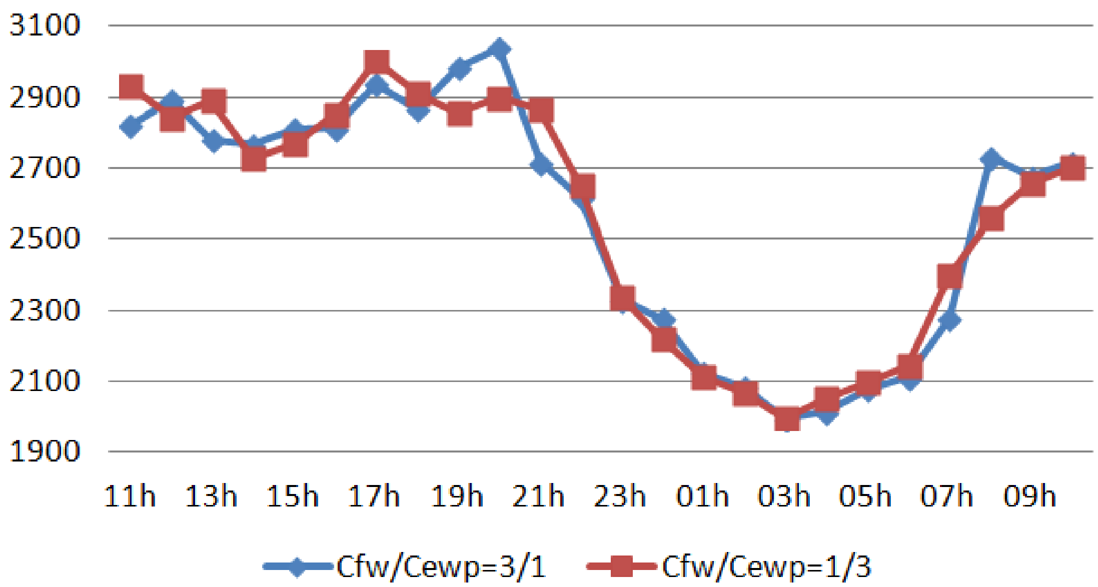


Figure 9. Testing period of the WD03 time series:
optimal adjustment of the pumping system (values in cubic meters).

4.3. Pumping systems adjustment.

As previously mentioned, the main output of the IDSS for real-time WDM is the optimal adjustment of the pumping systems. The system determines hourly the quantity of water to be pumped in order to minimize costs. As an example, Figure 9, based on the testing period of WD03, represents the solution proposed by the MAS for optimal adjustment of the pumping equipment in two opposite scenarios: when the excessive storage overrun cost is three times higher the emergency water pumping overrun cost (ANNs are used to forecast) and when the ratio is inverse (exponential smoothing is applied to forecast).

5. Conclusions and Future Work

This paper aims to show how the multi-agent methodology can be applied to WDM, a concept that has gained great importance in recent years given the requirements imposed by the new context marked by the scarcity of resources and the respect to the environment. In order to do this, an IDSS has been designed and implemented. It integrates sophisticated forecasting methods and management components under a structure that simulates a municipal water distribution system, and determines in real time the optimum adjustment of the pumping systems in order to minimize WDM costs.

Tests on time series with hourly water demand demonstrate the high efficiency of the developed system. They show that the introduction of AI techniques in the forecasting process, such as ANNs, can significantly decrease the error when compared with other traditional techniques, especially on holidays and days around them, because they have a greater capability of adapting to unexpected changes. This leads to a big reduction in WDM costs. However, the tests also show that the choice of the optimal alternative for adjusting the pumping systems depends on the input variables. Therefore, limiting the study to the search of the best forecasting method represents only a partial solution to the problem, which does not have to lead always to the best overall solution for the WDM system.

Again, it should be highlighted that this is a preliminary work or pilot system, where some simplified assumptions have been adopted (e.g. regarding the distribution system or the cost model). Translating this model into a real system would require reformulating these assumptions, as well as covering other common problems in real water distribution systems, such as leakages. In this regard, the main contribution of this work is that multi-agent methodology has proven to be not only a suitable tool to address this issue but also a necessary approach to study it, as WDM must be analyzed in its entirety from an holistic approach. In addition, this approach has enormous potential in increasing its functionality, as it allows managers to complete the study by adding new agents with the aim of increasing the scope of the system. It would also be possible to integrate this system into a MAS of greater magnitude.

References

1. A. S. Mohamed and H. H. G. Savenije, Water Demand Management: Positive Incentives, Negative Incentives or Quota Regulation?, Phys Chem Earth Pt B. 25 (2000) 251-258.
2. D. B. Brooks, An Operational Definition of Water Demand Management, Int J Water Resour D. 22 (2006) 521-528.
3. R. A. Young, Price Elasticity of Demand for Municipal Water: A Case Study of Tucson, Arizona, Water Resour Res. 9 (1973) 599-610.
4. R. H. Willsie and L. H. Pratt, Water use relationships and projection corresponding with regional growth Seattle region, Water Resour Bull. 10 (1974) 360–371.
5. D. R. Maidment and E. Parzen, Cascade model of monthly municipal water use, Water Resour Res. 20 (1984) 15–23.
6. R. B. Billings and D. E. Agthe, State-space versus multiple regression for forecasting urban water demand, Nord Hydrol. 35 (1998) 411–430.
7. S. Gato, N. Jayasuriya and P. Roberts, Temperature and rainfall thresholds for base use urban water demand modeling, J Hydrol. 337 (2007) 364–376.
8. J. Bougadis, K. Adamowski and R. Diduch, Short-term municipal water demand forecasting, Hydrol Process. 19 (2005) 137-148.
9. M. Herrera, L. Torgo, J. Izquierdo and R. Pérez-García, Predictive models for forecasting hourly urban water demand, J Hydrol. 387 (2010) 101–140.
10. Empresa Municipal de Aguas de Gijón (Spain). http://agua.gijon.es/. Last Access: October 30, 2015.
11. D. Maidment and S. Miaou, Daily water use in nine cities, Water Resour Res. 22 (1986) 845-851.
12. G. E. P. Box and G. M. Jenkins, Time Series Analysis: Forecasting and Control (Holden Day, San Francisco, CA, 1970)

13. A. An, C. C. Shan, N. Cercone and W. Ziarko, Discovering rules from data for water demand prediction, in Proceedings in the Workshop on Machine Learning and Expert System (1995), pp. 187-202.

14. L. Shvartser, U. Shamir, and M. Feldman, Forecasting Hourly Water Demands by Pattern – Recognition Approach, J Water Res Pl-ASCE. 119 (1993) 611-627.

15. G. A. Darbellay and M. Slama, Forecasting the short-term demand for electricity – Do neural networks stand a better chance?, International Journal of Forecasting. 16 (2000) 71-83.

16. N. Lertpalangsunti, C. Chan, R. Mason, and P. Tontiwachwuthikul, A tool set for construction of hybrid intelligent forecasting systems: application for water demand prediction, Artif Intell Eng. 13 (1999) 21-42.

17. A. Jain and L. E. Ormsbee, Short-term water demand forecasting modeling techniques-conventional versus AI, Journal AWWA. 94 (2002) 64-72.

18. J. Liu, H. G. Savenije, and J. Xu, Forecast of water demand in Weinan city in China using WDF-ANN model, Phys Chem Earth. 28 (2002) 219-224.

19. M. Nasseri, A. Moeini, and M. Tabesh. Forecasting monthly urban water demand using Extenden Kalman Filter and Genetic Programming, Expert Syst Appl. 38 (2011) 7387-7395.

20. L. Zhang, L. B. Jack and A. K. Nandi, Fault detection using genetic programming forecasting, Mech Syst Signal Pr. 19 (2005) 271-289.

21. D. Shrestha and D. Solomatine, Machine learning approach for estimation of prediction interval for the model output, Neural Networks. 19 (2006) 225-236.

22. M. Tabesh and M. Dini, Fuzzy and neuro-fuzzy models for short-term water demand forecasting in Tehran, Iran J Sci Technol B. 33 (2009) 61-77.

23. I. Msiza, F. Nelwamondo and T. Marwala, Artificial neural networks and support vector machines for water demand time series forecasting, in IEEE International Conference on Systems, Man and Cybernetics (2007), pp. 638-643.

24. B. Ponte, L. Ruano, R. Pino and D. De la Fuente, The Bullwhip effect in water demand management: taming it through an artificial neural networks-based system, J Water Supply Res T 64 (2015) 290-301.

25. J. Costas, B. Ponte, D. De la Fuente, R. Pino and J. Puche, Applying Goldratt's Theory of Constraints to reduce the Bullwhip Effect through agent-based modeling. Expert Syst Appl. 42 (2015) 2049-2060.

26. L. Breiman, Random Forests, Mach Learn. 45 (2001) 5-32.

27. G. Moisen and T. Frescino, Comparing five modeling techniques for predicting forest characteristics, Ecol Model. 157 (2002) 209-225.

28. S. Moss and B. Edmonds, Sociology and Simulation: Statistical and Qualitative Cross – Validation, Am J Sociol. 110 (2005) 1095-1131.

29. N. Athanasiadis, A. K. Mentes, P. A. Mitkas and Y. A. Mylopoulos, A hybrid agent-based model for estimating residential water demand, Simul-Trans Soc M S. 81 (2005) 175-187.

30. M. Galán, A. López–Paredes, A. and R. Del Olmo, An agent-based model for domestic water management in Valladolid metropolitan area, Water Resour Res. 45 (2009) w05401.

31. E. M. Zechman, Agent-Based Modeling to Simulate Contamination Events and Evaluate Threat Management Strategies in Water Distribution Systems, Risk Anal, 31 (2011) 758-772.

32. M. Giuliani and A. Castelletti. Assessing the value of cooperation and information exchange in large water resources systems by agent-based optimization, Water Resour Res. 49 (2013) 3192-3296.

33. J. Ni, L. Ren, M. H. Liu and D. Q. Zhu, A Multi-agent Dynamic Assessment Approach for Water Quality Based or Improved Q-Learning Algorithm, Math Probl Eng, (2013) ID 812032.

34. J. Ni, M. Liu, L. Ren, and S. X. Yang, A multiagent Q-learning-based optimal allocation approach for urban water resource management system, IEEE T Autom Sci Eng. 11 (2014) 204-214.

35. C. S. Karavas, G. Kyriakarakos, K. G. Arvanitis and G. Papadakis, A multi-agent decentralized energy management system based on distributed intelligence for the design and control of autonomous polygeneration microgrids. Energ Convers Manage. 103 (2015) 166-179.

36. S. Makridakis, Accuracy measures: theoretical and practical concerns, Int J Forecasting. 9 (1993) 527-529.

37. B. Ponte, D. De la Fuente, R. Pino and R. Rosillo, Real-time water demand forecasting system through an agent-based architecture, Int J Bio-Inspir Com. 7 (2015) 147-156.

CHAPTER 4

SYSTEMIC APPROACH TO SUPPLY CHAIN MANAGEMENT THROUGH THE VIABLE SYSTEM MODEL AND THE THEORY OF CONSTRAINTS♦

♦ This article was published in *Production Planning & Control: The Management of Operations* (Taylor & Francis), vol. 27, no. 5, pp. 421-430 in January 2016. DOI: 10.1080/09537287.2015.1132349. Authors (in order): Julio Puche, Borja Ponte, José Costas, Raúl Pino, David de la Fuente.

Production Planning & Control: The Management of Operations, indexed in the ISI, is a Q2-journal in the category 'Operations Research and Management Science' (Science Edition). Impact Factor 2015: 1.532.

Abstract

In today's environment, Supply Chain Management (SCM) takes a key role in business strategy. A major challenge is achieving high customer service level under a reasonable operating expense and investment. The traditional approach to SCM, based on local optimisation, is a proven cause of meaningful inefficiencies – e.g. the Bullwhip Effect – that obstruct the throughput. The systemic (holistic) approach, based on global optimisation, has been shown to perform significantly better. Nevertheless, it is not widely expanded, since the implementation of an efficient solution requires a suitable scheme. Under these circumstances, this paper proposes an integrative framework for supply chain collaboration aimed at increasing its efficiency. This is based on the combined application of the Beer's Viable System Model (VSM) and the Goldratt's Theory of Constraints (TOC). VSM defines the systemic structure of the supply chain and orchestrates the collaboration, while TOC implements the systemic behaviour – i.e. integrate processes – and define performance measures. To support this proposal, we detail its application to the widely used Beer Game scenario. In addition, we discuss its implementation in real supply chains, highlighting the key points that must be considered.

Keywords

Supply Chain Management; Systems thinking; Viable System Model; Theory of Constraints; Supply chain collaboration.

1. Introduction

The revolution of information and communications technologies, the decrease of transportation costs, the geopolitical restructuring that took place as a result of the Cold War, and the liberalization of capital markets have drawn a new competitive business context marked by its complexity and dynamism. Competition must manage efficiently convoluted worldwide networks being able to agilely react to the frequent changes in customer requirements. By way of illustration, Figure 1 shows the dramatic increase of the international trade of products in the Eurozone over the last decade (from 2003 to 2014, exports and imports grew by 97.57% and 79.74% respectively), even in a recessive economic context. In this regard, competition between firms is no longer limited to the product itself, but goes much further. For this reason, the concept of Supply Chain Management (SCM) has gained strength to the point of having a strategic importance for companies, which has encouraged researchers to deepen its study and the development of proposals to increase the yield of companies involved.

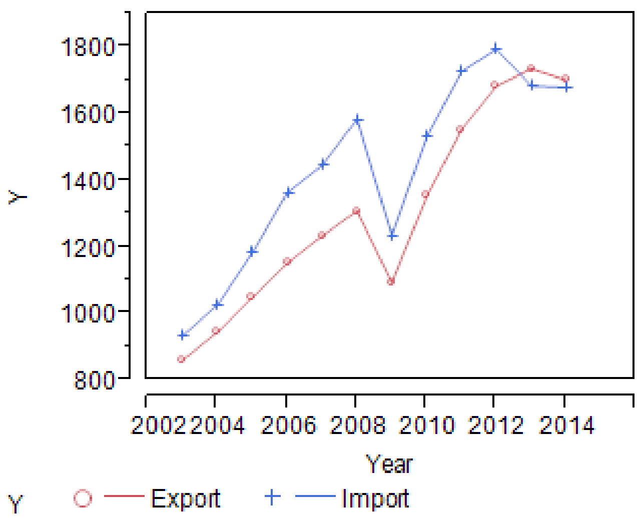


Figure 1. International trade in the Eurozone (in millions of euros). Data from Eurostat.

One of the main challenges regarding SCM is to improve the customer service level without capital outlay (CGI Group, 2013). The delivery of products on the right time and lowest cost enables a company to differentiate from its counterparts and enhances future profitability (Chopra and Meindl, 2007). Given the current complexity and dynamism, real supply chains usually exploit the throughput inappropriately, which leads to poor service levels. This issue, which can be considered as the problem statement of this research work, is especially relevant when lead times are long or the product experiences short life cycles (De Treville, Shapiro, and Hameri, 2004).

Sterman (1989) concluded that large inefficiencies occur within supply chains due to the individual adoption of local optima solutions by the various participants. This reductionist approach to SCM (the overall strategy is obtained as a sum of individual strategies), based on mass production paradigm, leads to increasing storage, shortage, labor, obsolescence, and shipping costs through the well-known Bullwhip Effect[10]. That is, this reductionist approach has shown to present several problems in terms of throughput management.

Thus, a premium has been placed on collaboration (Lehoux, D'Amours, and Langevin, 2014) as a key source of competitive advantages. This holistic or systemic approach –the overall strategy determines the individual strategies– to SCM, has been shown to outperform the traditional reductionist alternative –the overall strategy is obtained as a sum of individual strategies– (e.g. Disney and Towill, 2003; Kollberg, Dahlgaard, and Brehmer, 2006; Costas et al., 2015). However, although the improvement in operational and (consequently) financial terms is widely accepted by academics and practitioners, supply chain collaboration faces high hurdles, such as the menace of opportunistic behaviors (Simatupang, Wright, and Sridharan, 2004), which stresses the importance of defining an appropriate framework.

It should highlight that the recent economic crisis has been understood as a consequence of the fact that globalization still has not been able to develop systemic dynamic properties to deal with a growing variety of requirements (Schweitzer et al., 2009). This fact has increased the interest for new approaches to business based on holistic paradigms. Hence, supply chains must be underscored as boiling areas for innovation.

The main contribution of this research article is the proposal of a systemic approach to SCM, where to take advantage of the benefits derived from collaboration. Within the framework proposed by Simatupang and Sridharan (2005), the Theory of Constraints (TOC) (Goldratt, 1990) is the mechanism used to improve supply chain efficiency, while the Viable System Model (VSM) (Beer, 1979; 1981; 1985) orchestrates the implementation of the collaborative solution.

Our research method has followed guidelines from A3 Thinking (Sobek II et al., 2011). A3 Thinking provides researchers and practitioners with an efficient way of studying and

[10] The Bullwhip Effect refers to the amplification of the variability of orders along the supply chain.

tackling business problems. Its effectiveness has been widely demonstrated within the TPS (Toyota Production System) paradigm. This Lean's tool for problem solving allows us to provide a complete structure to implement successful moves toward organizational improvement through a deeper understanding of the issue. This structure can be observed along this article: (section 1) Problem statement, background and setting goals; (section 2) Clarifying the problem after reviewing the literature; (section 3) Developing the conceptual model, through the proposal of an integrative framework for supply chain collaboration; (section 4) Detailing the application of the proposed model on the well-known Beer Game scenario; and (section 5) Discussing its applicability in real supply chains.

2. Literature review

This section reviews the main collaborative approaches to SCM and introduces the two philosophies that are combined in this research work.

2.1. Supply chain collaboration

Supply chain collaboration can be easily defined as several companies creating competitive advantage, and hence obtaining higher profits, by working together in a production and distribution system (Simatupang and Sridharan, 2002). From that point on, collaboration has been understood in very different ways by researchers and managers. In this regard, Simatupang and Sridharan (2005) propose an outstanding framework, defined by five features: (1) Information sharing; (2) Collaborative performance system; (3) Decision synchronization; (4) Incentive alignment; and (5) Integrated processes. This integrative rather than sequential approach (the output of each feature acts as an input for the others) is supported by empirical evidence –if some of the features are ignored, intrinsic barriers could derail the collaborative process.

Supply chain integration encompasses the coordination of resources, decisions and methods among the different stakeholders and is the skeleton of the overall process. Decision synchronization covers devising joint decision-making processes (includes re-allocating decision rights) with the aim of synchronizing planning and execution levels. This includes forecasts, safety stocks, order placement, order delivery, target customer service level, and pricing. In this regard, a wide variety of solutions have been proposed in the last two decades to improve the performance of the supply chain, such as Vendor

Managed Inventory (Andel, 1996) and Collaborative Planning, Forecast and Replenishment (Ji and Yang, 2005). Moreover, some systemic philosophies like Lean Production (Womack and Jones, 1996) proposed methods to manage the production flow, e.g. Kanban and CONWIP control –see Takahashi and Nakamura (2002) for a comparison, and Jasti and Kodali (2015) for a review of existing Lean SCM frameworks.

Information sharing, the main enabler of collaboration, refers to the access to private data of all members, which covers dissemination of demand conditions, inventory and order status (and locations), cost-related data, and performance indicators. With this goal, the use of information and communications technologies has shown to improve supply chain efficiency (Gunasekaran and Ngai, 2004). Measuring this efficiency through systemic performance metrics (which must be devised to guide the participants to improve overall performance) is another key feature of supply chain collaboration. Companies require different types of metrics that span the supply chain (Kaplan and Cooper, 1997). For example, Li and O'Brien (1999) use four main criteria to measure supply chain efficiency: profit, lead time performance, delivery promptness, and waste elimination. Najmi and Makui (2012) propose a conceptual model for evaluating supply chain performance that we recommend looking at.

Lastly, incentive alignment requires to share costs, risks, and benefits among the participants. That is, aligning incentives aims to motivate them to act consistent with the overall strategy, and hence eliminating the incentives to deviate. Kaplan and Narayanan (2001) propose the use of expert systems, activity-based costing, and web-based technology to trace, calculate and display the incentive scores. To implement it, the use of linear contracts (e.g. pay-for-effort and pay-for-performance schemes) is common.

2.2. Viable System Model (VSM)

The VSM (Beer, 1979; 1981; 1985) offers the possibility to scientifically design an organization as a system with regulatory, learning and adaptive capabilities necessaries to ensure its survival (viability) when facing changes that may occur in its environment over time, even though they were not foreseen in its design. To achieve this viability, the VSM proposes an invariant systemic structure based on the definition of five functions, called Systems One to Five, that are considered necessary and sufficient conditions to deal with the environment complexity[11] in which the system operates.

System One represents the operational (autonomous) units managing the different production elements. Since conflicts between processes and responsibilities of these operational units might appear, System Two –with an essential role in coordination– is essential. Controlling the performance level of the operational units is assumed by System Three, which is also responsible for defining directives, allocating resources and corresponding accountability to each operational unit, as well as identifying potential synergies that might arise. Next to System Three, System Three* is responsible for performing audit activities to operational units. Since System Three is unable to predict the future and recognize potential risks, a structural function is required to solve this problem. This function is represented by System Four. Changes in the environment are detected and analyzed with regard to the system's main objectives, leading to possible recommendations for action. Finally, System Five formulates the principles and goals of the system, playing a key role in preserving its identity.

Supplementing these five functions, the VSM is supported by instruments for unfolding variety both horizontally and vertically. The horizontal unfolding aims to balance variety through the design of mechanisms to reduce (attenuators) and amplify (amplifiers) it. Thus, each connection (with four components: transmitter, transducer, channel, and receiver), which represents communication relationships among the functions of the model and between them and the environment, considers variety attenuation and amplification mechanisms in both senses (from the environment to operations and from operations to management). On the other hand, the vertical unfolding supports the recursion of operational units to smaller subsystems. The purpose is to reduce the variety faced by each part of the system (complexity reduction).

Although this socio-cybernetic theory has got increasingly greater recognition for their plausibility, Jackson and Flood (1988) criticized: (1) their purely theoretical design and abstract nature; (2) the questionable analogy between the human brain and other organizations; and (3) their hierarchical arrangement and lack of flexibility. Nonetheless,

---

[11] The complexity is measured by the concept of variety, i.e. the number of possible states or behaviour modes that a system can adopt (Ashby, 1956).

different authors show concrete VSM applications to diagnose or design viable organizations in a multitude of sectors (Beer, 1981; Puche Regaliza, 2014; 2015). These resulted in the discovery of pathologies and, after their treatment, the VSM led to a tangible improvement in such organizations.

Its systemic and multilevel nature, its ability to handle the dynamic complexity enclosed when managing an organization, and its interaction with the environment makes us consider the advantages offered by the VSM regarding SCM. In this subject, Chronéer and Mirijamdotter (2009) proposed its utilization for shortening a product development process by better connecting the information flows and Badillo et al. (2015) used VSM to better understand the supply chain of a telecommunications firm.

2.3. Theory of Constraints (TOC)

The TOC (Goldratt, 1990), a major innovation in the production field, is a management philosophy that views any system as being limited in reaching a higher performance level only by its bottleneck. Thus, it aims to achieve breakthrough improvements by only focusing on it[12]. The TOC encompasses three main areas: logical thinking, performance measurement, and operations management.

Its logical thinking focuses on the bottleneck through a continuous improvement philosophy (Goldratt, 1992). In order to increase the overall performance, all efforts must be concentrated on the system's constraint –that is, any improvement away from the bottleneck means a waste of resources. The cycle is split into five stages: (1) Identifying the bottleneck; (2) Deciding how to exploit it; (3) Subordinating everything else in the system to the previous decision; (4) Implementing measures to elevate the constraint; and (5) Assessing whether the bottleneck has been broken, and re-starting the cycle to avoid that inertia limits the system.

The performance measurement is based on a simple idea: the only purpose of a business is to make money now and in the future. To quantify the success in achieving this goal, TOC uses three financial indicators: net income (absolute terms), return-on-investment (relative terms), and cash flow (survival terms). This theory highlights the simultaneous consideration of the three –it is not about increasing one at the expense of the others. To

[12] Unlike Lean Production that shares the effort throughout the whole system.

determine these metrics, the Throughput Accounting (Goldratt, 1994) is proposed, which considers three operational indicators: throughput (how much money does the system generate?), investment (how much money does the system need to generate throughput?), and operating expenses (how much money is required to operate?). Unlike the traditional cost accounting (aimed at cost reduction), this accounting seeks to maximize the throughput, i.e. to optimize the efficiency of the value stream. Thus, it enables managers to analyze the link between process constraints and financial performance in decision making. Consequently, it allows them to determine the real impact of their decisions.

The logistic function applies the Drum-Buffer-Rope (DBR) method (Goldratt, 1990), which aims to manage properly the bottleneck through suitable coordination (ensuring its steady supply). It is named for its three main components. The *drum* is a system pacemaker and is placed at the node that limits system performance. The other nodes follow its beat (production rate) so the drum is protected against variability by the *buffer*, whose size plays a key role (Ye and Han, 2008; Kuo-Jung, Sheng-Hung, and Rong-Kwei, 2003) –the full capacity of the bottleneck must be exploited. The *rope* is the release mechanism, which subordinates the entire system (upstream and downstream) to the drum. That is, orders must be released according to the buffer time before they are due. The planning stage, or DBR configuration, is complemented with the monitoring stage, which implies managing the buffer along the different nodes to tune the system for peak performance.

Although TOC was initially oriented to production systems, its application to other business areas has been further studied, such as Marketing (Goldratt, 1994) and Project Management (Goldratt, 1997). In SCM (Goldratt, Schragenheim, and Ptak, 2000), the early works deal with managing the system from a single company perspective (Cox and Spencer, 1998). Later studies used TOC to promote supply chain collaboration. Simatupang, Wright, and Sridharan (2004) provide a conceptual framework for using TOC in supply chains. Wu et al. (2010) developed a DBR-based replenishment model under capacity constraints. Costas et al. (2015) showed that the DBR method induces large operational improvements in the supply chain without any collateral damage. According to their practical experience, the TOC Center reports that firms adopting TOC typically gain 25-100% of additional output without significant increase in expenditure (Mabin and Balderstone, 2003).

3. Conceptual model: Integrative framework for supply chain collaboration

The top performer paradigm for production systems is the well-known Lean Production. Thus, proposing the Lean implementation seems to be the natural step to tackle the problem that consists on closing the gap in terms of throughput in the supply chain. Lean is a brilliant systemic philosophy, and we have no doubts in recommending it in each supply chain member. Nonetheless, two key points in Lean are focusing on the flow of value to customers and, simultaneously, reducing the overall MUDA (Lean's term for waste) in a systematic and continuous way. These points remain extremely important in supply chains, as these are multi-agent systems –and this makes a difference.

In our proposal, VSM is used as a framework for system design and diagnosis. According to it, when the supply chain is considered as system-in-focus and apply recursion to the subsystems (organizations that belong to the supply chain), it becomes clear that most of the reported common issues in collaborative supply chains are linked to the alignment of System Five (values, culture, principles, rules, and the overall policy) for all nodes. Hence, by taking TOC as the general paradigm for the supply chain, we strongly attenuate the variety (simplify) such a big issue. The reason is that TOC works by putting the system bottleneck in the center of attention for everyone. By means of that, the system-in-focus has a natural representative node, which is the one where the bottleneck is placed – nonetheless TOC also manages the change of the bottleneck.

Applying TOC across the supply chain, through the DBR method, warranties to concentrate all agents to what matters for the system as a whole: the throughput, the operational expense, and the investment. These indicators act as a balance scorecard to monitor the system. However, the implementation of TOC in supply chains does not come without very important challenges. Watson, Blackstone, and Gardiner (2007) can be consulted for a review on the evolution of this management philosophy, including the problems impeding greater acceptance. Here the Simatupang and Sridharan's schema is introduced.

The center of this integrative schema underscores the information sharing. Poor information sharing is a general problem, but with TOC it becomes more harmful because of the fact of applying an inventory managed policy as a need to manage the rope –the key artefact used to manage the flow. VSM cares about potential issues of this type with System Three*, which is what we propose as an element to ensure that the adequate degree of transparency and validity of the information (in terms of availability, opportunity, and cost) is surveilled.

Like Lean, TOC requires the orchestration of all processes, namely core and enablers. Core process is anyone that delivers value to customers, while enabler processes are focused on providing services inside the system. Hence, the feature regarding integrated supply chain processes in Simatupang and Sridharan's schema provide guidelines to address this issue. For core processes, TOC is self-sufficient, but not for enablers. Such orchestration is examined at the light of VSM. System Two provides the context to keep at the lowest possible place in the organization the everyday decisions to keep the system running smoothly.

Every participant takes many decisions that have an impact across the whole supply chain. For instance, launching promotions may need most of the times coordination, approvals, and other activities that must be properly synchronized. Such decision-making with the focus on the whole system is shaped using Systems Three, Four and Five (the meta-system) in order to early detect poorly structured constructions for the decision support system.

Last but not least, the issue of redistributing the overall economic profit obtained by the system must be considered. Implementing TOC means that many decisions with strong economic impact must be done according to general rules and, consequently, this can (and does) generate conflicts for the agents if the system does not take care about it. Incentive alignment and overall performance metrics are a must.

To sum up our approach, the key component to solve the problem regarding throughput in supply chains is TOC, which implements the most that system needs for the core processes through the DBR method and the Throughput Accounting. Then, once this necessity has been created, the Simatupang and Sridharan's scheme provides an appropriate framework for collaboration based on five blocks. Finally, VSM acts as a guideline for initial and permanent regular design and diagnosis about structural weaknesses of the system, caring about the five VSM functions as well as the channels properties, and making judgements in terms of the laws of requisite variety. Figure 2 highlights and summarizes the conceptual model proposed. In section 5, we introduce some of the challenges that must be taken into account so to prevent major errors.

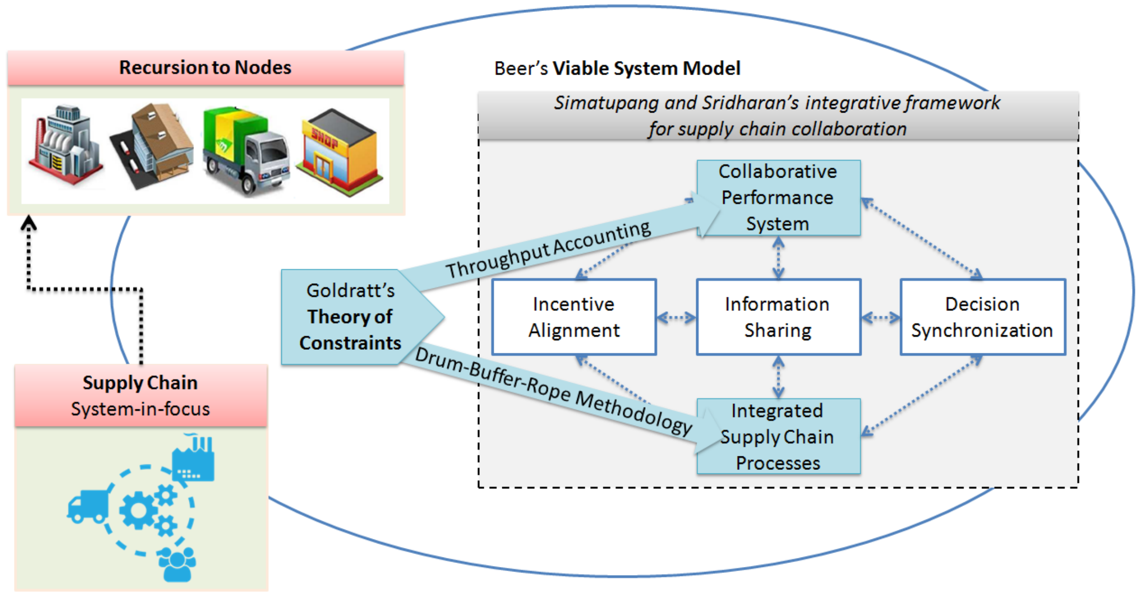


Figure 2. Collaborative model for SCM based on VSM and TOC.

4. Hypothetical case study

The Beer Game is a role-playing exercise aimed at teaching the main SCM principles that has been used in countless management courses over the last 50 years. Its experimental and counterintuitive nature has proved to be very effective in helping managers to understand the causal relationships between decision-making and supply chain behavior (Goodwin and Franklin, 1994), showing the generation of large inefficiencies (Sterman, 1989). This way, the Beer Game scenario, a single-product linear supply chain, composed of four echelons, has been widely used in literature to emphasize, investigate and analyze supply chain dynamics (Macdonald, Frommer, and Karaesmen, 2013).

Under these circumstances, the Beer Game supply chain is a suitable fit to show how our proposal can be implemented, as it: (1) covers the two main flows of the system along a significant length; (2) incorporates all commonly available sources of information within supply chains; (3) brings a rich enough sequence of events so that the belief-desire-intention of the participants can be analyzed in front of different event algebras; and (4) has widely shown in literature the problems of the reductionist approach, i.e. that the interaction of individual decisions produces a solution that is far from the optimal.

In this section, we describe the application of our systemic proposal for SCM to the Beer Game scenario. We first define the system-in-focus. Subsequently, the five VSM functions are designed pointing out where each Simatupang and Sridharan's feature must be considered.

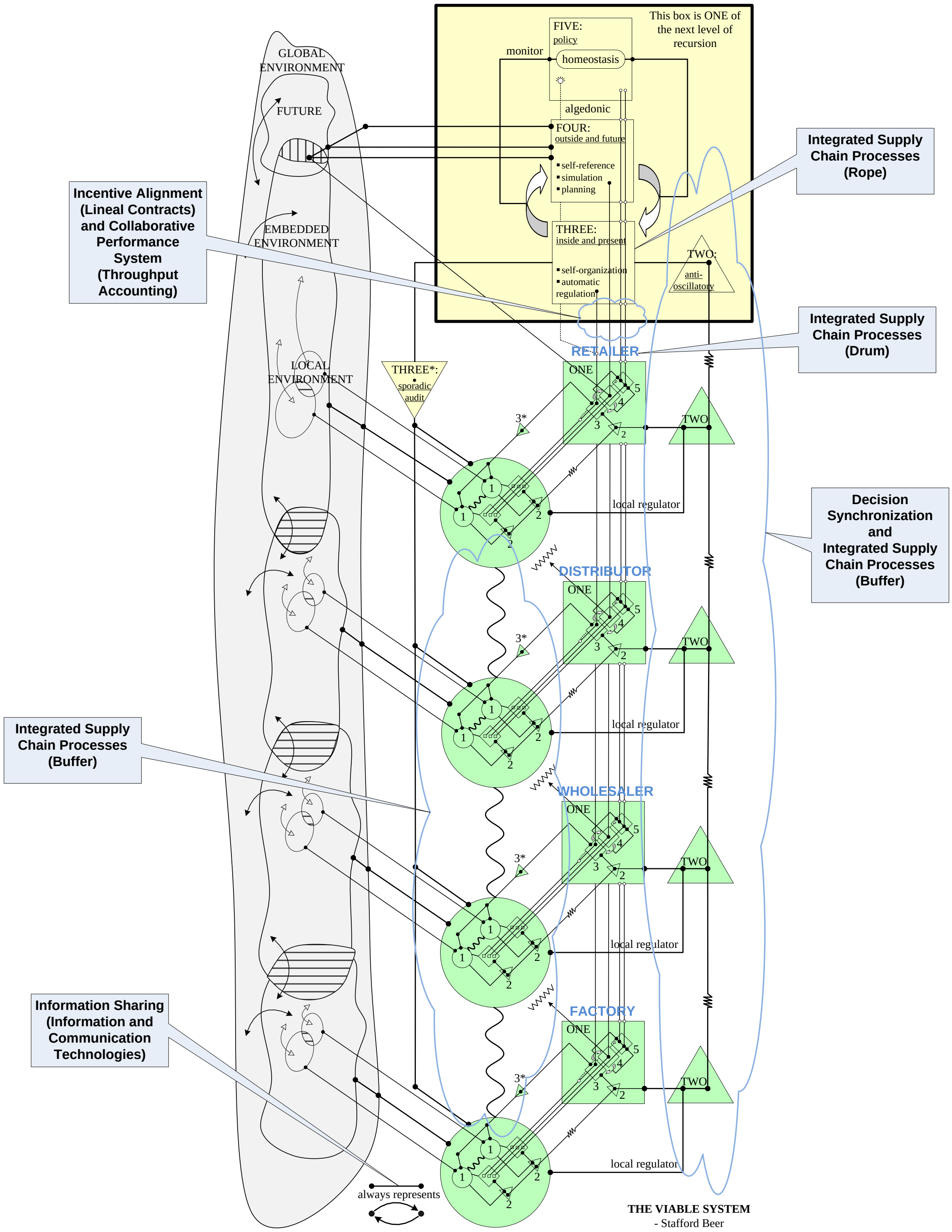


Figure 3. Systemic approach to SCM. Adapted from Beer (1985).

4.1. System-in-focus

From a TOC-based perspective, only satisfying customer requirements generates throughput for the global supply chain. Therefore, the nodes (factory, wholesaler, distributor, retailer, using the common notation in the Beer Game) must be aimed at maximizing it. Taking the supply chain itself as system-in-focus, each node represents a

System One operational unit (see Figure 3). If we enter each one of them, we find a structural replica of VSM (structural invariant), i.e. a set of nodes representing the operational units. From that point on, the next level of recursion is represented for each (system-in-focus) operational unit. Likewise, in the previous level of recursion we find all the industry. Each supply chain of it is an operational unit –we focus on one of them. We can further define previous (countries, continents, etc.) and next (departments, production lines, etc.) levels of recursion (vertical variety unfolding).

The supply chain overall function is threatened by a number of noise sources, which can be classified into four kinds[13]: (1) surrogate noise (geographical point of view, use of space, etc), (2) temporal noise (progressive deterioration, mutations in the environment, legislation, etc), (3) system noise (latencies, faults, defects, errors, etc), and (4) external noise (demand variability, raw materials, etc). The third kind is related to the System One typical functioning, while the others are related to the supply chain environment. Note that the last one refers to the upper and lower nodes of the system –retailer to customers (final product) and factory to suppliers (raw materials).

4.2. System One

Once reviewed the perspective of taking as system-in-focus the operational nodes, the systemic behaviour must be shaped according to TOC. The bottom right of Figure 3 displays the four operational elements representing the four supply chain nodes. Each one is composed of: (1) a management unit (square shape); (2) operations (circular shape) responsible for interacting with the environment to offer their products and services (left amoeba); and (3) a system of local coordination (triangular shape).

Since TOC is based on managing the supply chain through its bottleneck, the first step is to detect it. Where is the bottleneck in a supply chain? It is not a static but a continuous question. The factory could be the bottleneck if its production rate cannot cover customer demand. Intermediate echelons could be bottlenecks if their transport or storage capacities significantly limit the customer service level. Nonetheless, the supply chain bottleneck tends to be related to customer demand (Youngman, 2009). In order to maximize the flow, lost sales in the retailer must be minimized. Therefore, following the DBR method, the

[13] According to the p-diagram classification, a widely-used technique in robust engineering.

drum should be placed on the retailer, see figure 3. Nonetheless, its allocation could be displaced to other node over time.

In a supply chain, there are two main flows: the material flow (downstream product shipping) and the information flow (to monitor and control the system, and to coordinate actors for decision-making; e.g. upstream flow of orders and downstream flow of shipping notes). These flows are represented by the different connections between nodes, between them and the environment, between them and the VSM functions, between functions, and between functions and the environment. Each connection, simplified in Figure 3, represents an information or material exchange in both senses, serving also as variety amplifiers or attenuators (horizontal unfolding). The Simatupang and Sridharan's feature related to information sharing through information and communication technologies is allocated in this point. Also, it can be extended to all other connections in the VSM.

4.3. System Two

In the upper right area of Figure 3, a triangle representing the global System Two can be observed. This is responsible for the coordination of the four operational elements through its interaction with their locals Systems Two. When the demand arrives at the system, the drum makes signals to the rest of operational units or supply chain nodes. They remain subordinated to the drum through a rope, so that the customer demand estimation is linked directly to the factory. Each node calculates the length of its rope until the drum position and orders material movements on the basis of its buffer downstream until reaching the bottleneck. The drum node issues orders directly to the factory. The buffer management consists of moving the flow so that arrivals occur in time in the bottleneck[14]. The buffer represents the material release duration while the rope corresponds to the release synchronization[15].

[14] Instead of a traditional safety stock based on material quantities, TOC-based buffers depend upon the lead time.

[15] To deepen into the TOC implementation, the exceptional guide developed by Youngman (2009) is highly recommended.

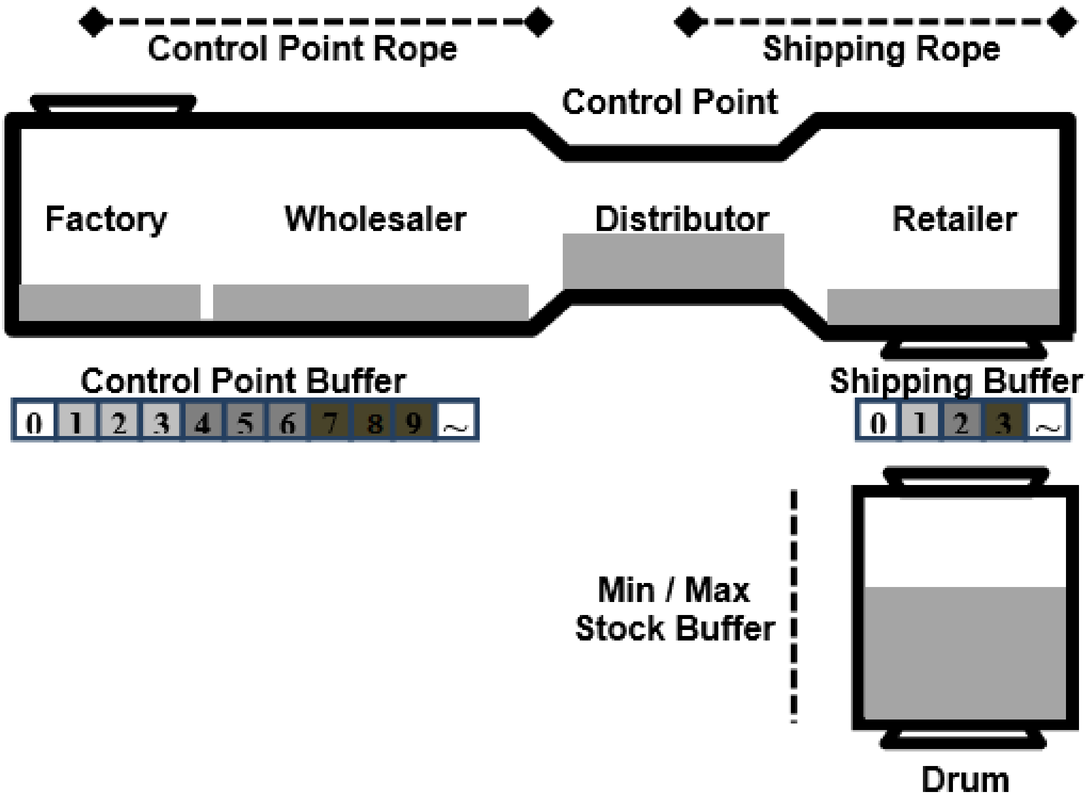


Figure 4. DBR method applied to the Beer Game scenario. Adapted from Youngman (2009).

As mentioned, the factory uses customer demand (time slot defined by rope, which is the time slot to protect) to decide the production orders that must be placed in the channel. Manufacturing time is equal to the shipping lead time in the remaining levels. Subsequently, each node except the retailer (since there are no downstream nodes) manages the buffer, which represents both time and material flow. Managing it means to compensate the downstream dissipated flow after shipment in each slot. Orders are dosed in the buffer and, consequently, are dissipative. They have not lead time, since each node decides how much to dose subordinate to the bottleneck. In addition, backorders are not generated as the new dose also obey the bottleneck. The DBR method applied to the Beer Game scenario is schematically shown in Figure 4. Although the usual case is to place the bottleneck on the retailer, we have placed it on the distributor with the aim of illustrating a more complex example, where to observe two ropes and two buffers. It should be clarified that to plot the graph, we have considered the lead time to be 3 time units, hence the control point buffer is 9 time units and the shipping buffer is 3 time units.

As previously mentioned, System One is represented by nodes that compose the supply chain, one of which is the drum. In System One, one part of the buffer can be observed, namely the squiggly lines that link supply chain nodes and that represent the material flow among them. The rest of the buffer is represented by System Two, which enables the coordination between all operational units. Regarding Simatupang and Sridharan's framework, the buffer (from the integrated supply chain processes feature) is allocated on these two points, while decision synchronization is located on the last point, see Figure 3. Synchronization mechanisms, as those mentioned in section 3, must be used to force the nodes follow the downstream material flow and upstream orders flow sequences. Thus, purchase orders, sales orders and backorders of the all different nodes are coordinated.

4.4. System Three

In Figure 3, System Three is identified in a square next to global System Two. The rope is represented by System Three. It takes care of finding synergies between nodes, assigning appropriate resources to each one, accountability of using these resources (agreement contract), and transmitting the system-in-focus rules to each node. The incentive alignment feature (from Simatupang and Sridharan's scheme), which can be implemented through linear contracts, is allocated on this point, see Figure 3. System Three* (inverted triangle) allows managers to audit the nodes performance without relying on the information they sent through System Two and central channels connecting with System Three (which forms the information flow). To deploy System Three*, audits are carried out, which enables monitoring and makes the information (shared through the overall framework) reliable.

In this point, the collaborative performance system is allocated, see Figure 3. We propose its implementation through the Throughput Accounting that is based on three main operational indicators. Firstly, the throughput expresses the rate at which money is generated, and is obtained through the difference between the revenue (sales at the retailer) and variable costs related to purchases (raw materials at the factory). Note that internal sources damage the throughput, e.g. defective products. Secondly, the operating expense refers to all that costs spent in turning inventory into throughput. It is the sum of storage, transport, labor and order costs. It should be highlighted that system thinking requires considering only overall rather than local (per echelon) indicators. Hence in the holistic approach, it would have no sense to consider (the common in the Beer Game) backlog costs among the different participants, as it is not money entering or leaving the system. Thirdly, the inventory is calculated by estimating the economic value of the products that are inventory in the system, both on-hand (net stock) and on-order (in transport), as well as all other invested money, e.g. equipment, machinery, and facilities.

Through the concept of “cost bridge”, the improvement in the aforementioned metrics leads to an increase in the financial indicators, which can be easily obtained. The net profit is the difference between throughput and operating expense. The cash flow considers, besides the above difference, the change of investment in the time horizon to analysis. The return-on-investment is the ratio of the net profit to the inventory. Figure 5 outlines the implementation of the Throughput Accounting in the Beer Game scenario. By means of this collaborative accounting, we aim to quantify the impact of the decisions, through analyzing the relationship between process constraints and financial performance in decision making. We think the Throughput Accounting proposes a suitable structure where to tackle the problem of exploiting throughput within supply chains.

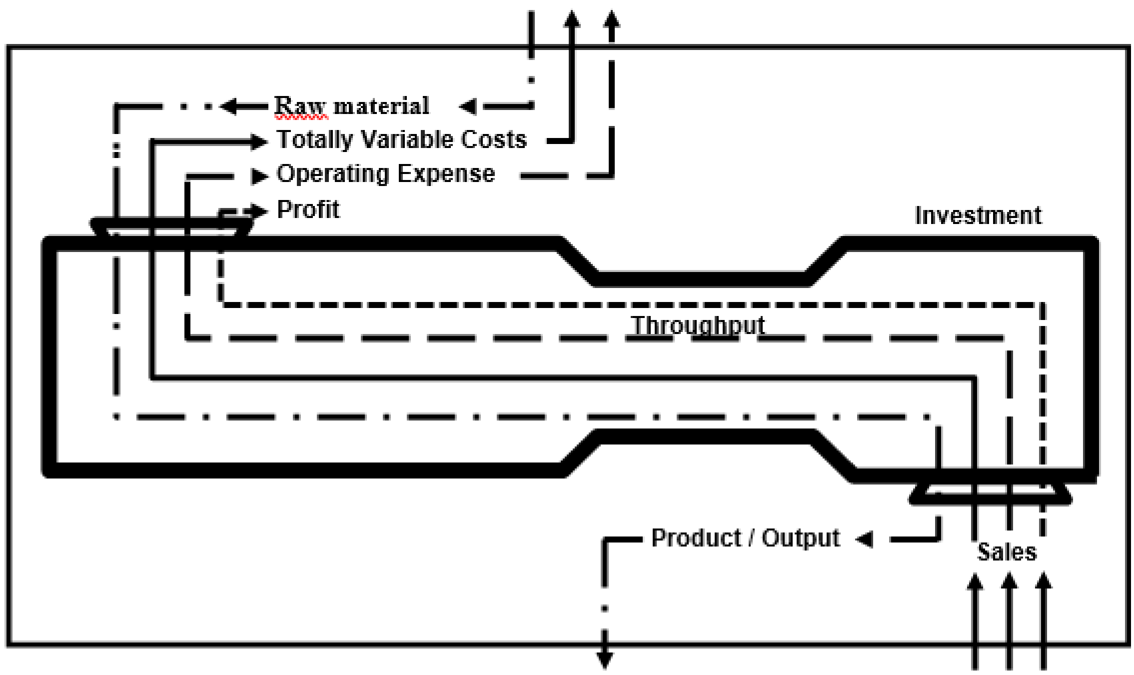


Figure 5. Throughput Accounting in the Beer Game scenario. Adapted from Youngman (2009).

4.5. System Four

In Figure 3, System Four is represented by a square shape just above System Three. In addition, it can be observed the interaction between System Four and the environment (left amoeba), allowing its inspection. The arrows between System Three and System Four enable exchanging information on what is happening in the organization and in the environment now and what will happen in the future. This last interaction is of special importance since, on the one hand, System Three adapts the organization based on the indications identified by System Four and, on the other hand, System Four inspects the environment based on what the organization is currently doing. System Four is responsible for preparing the supply chain against the possible changes that may arise in the future (prediction), providing the whole system with the necessary adaptability to maintain its viability over time. In this systemic approach, System Four is more relevant than in a reductionist approach in order to try to eliminate redundancies and minimize local Systems Four in favour of promoting the need for monitoring the environment of the supply chain as a whole. In addition, the bottleneck should acquire special importance, which can be manifested in managing demand and markets development of the supply chain and its competitors.

4.6. System Five

Finally, System Five can be identified in Figure 3 by a square shape above System Four. It is connected with the interaction between System Three and System Four interaction, which allows practitioners to solve the problems encountered when System Three and System Four do not agree on the basis of the principles defined by System Five. In addition, Figure 3 highlights the algedonic channel that connects directly (and unidirectionally) the operational elements with System Five, allowing them to alert System Five in case of serious risk. System Five defines the philosophy to be followed by the overall supply chain, which according to TOC principles is to make money now and in the future. In this case, the need for transparency of information (Simatupang and Sridharan's framework for application of TOC) and the enforcement of the Goldratt's principles as management principles for all nodes (everything is subordinated to the bottleneck) must be highlighted as key points.

5. Implementation in real supply chains: discussion and future work

One of the main challenges facing supply chains currently is to improve efficiency by increasing simultaneously the net profit, the return-on-investment and the cash flow. In other words, the throughput must be appropriately exploited along the system. Supply chain collaboration has shown to be effective to deal with this issue. Nonetheless, this systemic approach is not totally widespread within real supply chains, as some high barriers emerge.

Supply chain participants must understand that the implementation of compelling solutions based on collaboration is a complex process that requires an appropriate scheme, such as the one proposed by Simatupang and Sridharan (2005). Based on this, we propose an integrative framework to achieve an holistic SCM. The DBR method, from Goldratt's TOC, is used to integrate processes –and hence to improve supply chain efficiency, measured through the systemic Throughput Accounting. The VSM also plays a crucial role in the overall process, as it orchestrates the framework so to define the systemic structure of the supply chain. In this reciprocal approach, the interaction of different connecting features of collaboration is being addressed.

To support this proposal, we detail its application to the hypothetical and widely used Beer Game scenario, although it can be easily adapted to other supply chain topologies. This case study aims at providing managers with insight to evolve from a reductionist approach, where the global strategy is obtained as the sum of individual strategies, to a holistic approach, where the individual strategies arise from the global strategy.

Concerning the application of this framework in real supply chains, the main catalyst is the overall improvement potential for the whole group of supply chain participants induced by TOC. Mabin and Balderstone (2003) reviewed several TOC applications in practice and calculated an average increase in the throughput by 63% without a significant increase in operational expense. On the other hand, the major hurdle to overcome is the dilemma of the predatory relationships among supply chain agents, since they are not expected to be in conflict in this new paradigm, but to behave as powerful partners –i.e. their incentives to deviate must be completely removed.

According to the personal experience of the authors of this article, in order to implement the proposed framework, one relevant opportunity to capture is the central purchasing unit. If a central purchasing unit does exist in the supply chain, it can develop towards a kind of headquarter to host the VSM System Five, System Four, and System Three* –and will also play a big role in System Two.

System Five, and its recursion to every node, can be fostered by a central unit by developing task force encounters and other group techniques, where the goodwill of the supply chain is worked by groups and activities. In addition, rules derived from the TOC philosophy can be established so to reinforce the vision of the whole system and the way members are expected and required to contribute in specific manners.

System Four will take a much more effective instantiation from the supply chain as a whole rather than the intelligence generated by each node. The reason is made evident at the light of the VSM; once the system-in-focus is the whole supply chain a new SWOT analysis[16] starts. Thereby, the horizontal unfolding of variety enters to work, and generates huge value to the supply chain (applying the Forces of Porter, the Delta Hax analysis, etc). Actually, in supply chains in which there is a "lion", this function is usually taken by this node, who forces the others to follow some roadmap; if the "lion" do not convince, activity is deployed by huge effort and energy rather than smoothly.

System Three* takes place by Lean's *genchi genbutsu*, i.e. going to the place and observe critically. Cross visits, blitz events (activities lasting a few consecutive days to produce a tangible alteration, generally deploying best demonstrated practices outside) and other activities are placed by the central unit after discussion to achieve *nemawashi* (i.e. consensus obtained by applying a scientist schema) in order to fuel continuous improvement to raise common standards shared in the system, namely to protect the image that the supply chain projects to the environ (customers, public, suppliers, etc).

Regarding possible extensions of this research work, some areas require further investigation in the VSM application in supply chains, e.g. to detail how to shape all VSM components (functions, different recursion levels, etc) and variety amplifiers and attenuators. Moreover, we would like to research why the systemic approach is not widely used yet, being extensively verified that this mature theory outperforms classical approaches. We know that moving each node away from their selfish natural behaviour needs some education phases. For this reason, we also aim to focus on the transition process: from reductionism to holism in SCM. Nonetheless, and in conclusion, the good news is that a mature state for moving towards this direction is available, and as previously highlighted the literature brings evidences that expectations for success are quite high for most supply chains.

[16] SWOT (Strengths, Weaknesses, Opportunities and Threats) analysis: a tool in risk analysis and business strategy.

References

1. Andel, T. 1996. “Manage inventory, own information.” Transportation & Distribution 37 (5): 54-58.
2. Ashby, W.R. 1956. An Introduction to Cybernetics. Chapman & Hall.
3. Badillo, I., R. Tejeda, O. Morales, and A. Briones. 2015. A Systems Science/Cybernetics Perspective on Contemporary Management in Supply Chains. Applications of Contemporary Management Approaches in Supply Chains. Associate Prof. Hakan Tozan (Ed.).
4. Beer, S. 1979. The Heart of Enterprise. John Wiley & Sons.
5. Beer, S. 1981. Brain of the Firm. 2nd edition. John Wiley & Sons.
6. Beer, S. 1985. Diagnosing the System for Organization. John Wiley & Sons.
7. CGI Group. 2013. “Maximising supply chain throughput with existing infrastructure.” Accessed August 25 2015. http://www.cgi-group.co.uk/brochure/maximising-supply-chain-throughput
8. Chopra, S., and P. Meindl. 2007. Supply Chain Management: Strategy, Planning and Operations. 3rd edition. Upper Saddle River, NJ: Pearson, Prentice-Hall.
9. Chronéer, D., and A. Mirijamdotter. 2009. “Systems thinking benefits in supply change management: an illustration of the Viable Systems Model in a Supply Chain.” International Journal Intelligent Systems Technologies and Applications 6 (3/4): 227-248.
10. Costas, J., B. Ponte, D. de la Fuente, R. Pino, and J. Puche. 2015. “Applying Goldratt’s Theory of Constraints to reduce de Bullwhip Effect in a supply chain through agent-based modeling.” Expert Systems with Applications 42 (4): 2049-2060.
11. Cox, J.F., and M.S. Spencer. 1998. The Constraints Management Handbook. Lucie Press.
12. De Treville, S., R.D. Shapiro, and A.P. Hameri. 2004. “From supply chain to demand chain: the role of lead time reduction in improving demand chain performance.” Journal of Operations Management 21: 613-627.
13. Disney, S.M., and D.R. Towill. 2003. “The effect of vendor managed inventory (VMI) dynamics on the Bullwhip Effect in supply chains.” International Journal of Production Economics 85:199-215.
14. Goldratt, E.M. 1990. Theory of Constraints. North River Press.

15. Goldratt, E.M. 1992. An Introduction to Theory of Constraints: The Production Approach; Workshop Description. Avraham Y. Goldratt Institute.

16. Goldratt, E.M. 1994. It's not Luck. North River Press.

17. Goldratt, E.M. 1997. Critical Chain. North River Press.

18. Goldratt, E.M., E. Schragenheim, and C.A. Ptak. 2000. Necessary but not sufficient. North River Press.

19. Goodwin, J.S., and S.G. Franklin. 1994. "The Beer Distribution Game: Using simulation to teach Systems Thinking." Journal of Management 13: 7-15.

20. Gunasekaran, A., and E.W.T. Ngai. 2004. "Virtual supply-chain management." Production Planning & Control 15 (6): 584-595.

21. Jackson, M.C., and R.L. Flood. 1988. "Cybernetics and Organization theory: a critical review." Cybernetics and Systems 19 (1): 13-33.

22. Jasti, N.V.K., and R. Kodali. 2015. "A critical review of lean supply chain management frameworks: proposed framework." Production Planning & Control 26 (3): 1-18.

23. Ji, Y.F., and H.L. Yang. 2005. "Bullwhip Effect Elimination in Supply Chain with CPFR." Proceedings of the 2005 International Conference on Manegement Science & Engineering 1 (3): 737-740.

24. Kaplan, R.S., and R. Cooper. 1997. Cost and Effect. Harvard Business School Press.

25. Kaplan, R., and V.G. Narayanan. 2001. "Measuring and managing customer profitability." Journal of Cost Management 15 (5): 5-15.

26. Kollberg, B., J.J. Dahlgaard, and P.O. Brehmer. 2006. "Measuring lean initiatives in health care services: issues and findings." International Journal of Productivity and Performance Management 56: 7-24.

27. Kuo-Jung, Y., C. Sheng-Hung, and L. Rong-Kwei. 2003. "Enhancement of Theory of Constraints replenishment using a novel generic buffer management procedure." International Journal of Production Research 41 (4): 725-704.

28. Lehoux, N., S. D'Amours, and A. Langevin. 2014. "Inter-firm collaborations and supply chain coordination: review of key elements and case study." Production Planning & Control 25: 858-872.

29. Li, D., and C. O'Brien. 1999. "Integrated decision modelling of supply chain efficiency." International Journal of Production Economics 59 (1): 147-157.

30. Mabin, V.J., and S.J. Balderstone. 2003. “The performance of the theory of constraints methodology: analysis and discussion of successful TOC applications.” International Journal of Operations & Production Management 23 (6): 568-595.

31. Macdonald, J.R., I.D. Frommer, and I.Z. Karaesmen. 2013. “Decision making in the beer game and supply chain performance.” Operations Management Research 6 (3-4): 119-126.

32. Najmi, A., and A. Makui. 2012. “A conceptual model for measuring supply chain's performance.” Production Planning & Control 23 (9): 694-706.

33. Puche Regaliza, J.C. 2014. “Extending the Viable System Model scope on ICT-sector software projects in Castilla y León.” Kybernetes 43 (2): 192-209.

34. Puche Regaliza, J.C. 2015. “Quantitative analysis of Viable System Model on software projects in the ICT sector in Castilla y León.” Kybernetes 44 (5): 806-822.

35. Schweitzer, F., G. Fagiolo, D. Sornette, F. Vega-Redondo, A. Vespignani, and D.R. White. 2009. “Economic networks: The new challenges.” Science 325 (5939): 422-425.

36. Simatupang, T.M., and R. Sridharan. 2002. “The collaborative supply chain.” The International Journal of Logistics Management 13 (1): 15-30.

37. Simatupang, T.M., and R. Sridharan. 2005. “An integrative framework for supply chain collaboration.” The International Journal of Logistics Management, 16 (2): 257-274.

38. Simatupang, T.M., A.C. Wright, and R. Sridharan. 2004. “Applying the Theory of Constraints to supply chain collaboration.” Supply Chain Management: An International Journal, 9 (1): 57-70.

39. Sobek II, D.K., and A. Smalley. 2011. Understanding A3 thinking: a critical component of Toyota's PDCA management system. CRC Press.

40. Sterman, J.D. 1989. “Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment.” Management Science 35 (3): 321-339.

41. Takahashi, K., and N. Nakamura. 2002. “Comparing reactive Kanban and reactive CONWIP.” Production Planning & Control 13 (8): 702-714.

42. Watson, K. J., J.H. Blackstone, and S.C. Gardiner. 2007. “The evolution of a management philosophy: The theory of constraints.” Journal of Operations Management, 25 (2): 387-402.

43. Womack, J.P., and D.T. Jones. 1996. Lean thinking: Banish waste and create wealth in your organisation. Simon and Shuster.

44. Wu, H.H., C.P. Chen, C.H. Tsai, and T.P. Tsai. 2010. "A study of an enhanced simulation model for TOC supply chain replenishment system under capacity constraint." Expert Systems with Applications 37: 6435-6440.

45. Ye, T., and W. Han. 2008. "Determination of buffer sizes for drum–buffer–rope (DBR)-controlled production systems." International Journal of Production Research 46 (10): 2827-2844.

46. Youngman, K. 2009. "A Guide to Implementing the Theory of Constraints (TOC)." Accessed December 20 2014. http://www.dbrmfg.co.nz/

CHAPTER 5

APPLYING GOLDRATT'S THEORY OF CONSTRAINTS TO REDUCE THE BULLWHIP EFFECT THROUGH AGENT-BASED MODELING♦

♦ This article was published in *Expert Systems with Applications* (Elsevier), vol. 42, no. 4, pp. 2049-2060 in March 2015. DOI: 10.1016/j.eswa.2014.10.022. José Costas, Borja Ponte, Raúl Pino, David de la Fuente, Julio Puche.

Expert Systems with Applications, indexed in the ISI, is a Q1-journal in the category 'Operations Research and Management Science' (Science Edition). Impact Factor 2015: 2.981.

Abstract

In the current environment, Supply Chain Management (SCM) is a major concern for businesses. The Bullwhip Effect is a proven cause of significant inefficiencies in SCM. This paper applies Goldratt's Theory of Constraints (TOC) to reduce it. KAOS methodology has been used to devise the conceptual model for a multi-agent system, which is used to experiment with the well-known 'Beer Game' supply chain exercise. Our work brings evidence that TOC, with its bottleneck management strategy through the Drum–Buffer– Rope (DBR) methodology, induces significant improvements. Opposed to traditional management policies, linked to the mass production paradigm, TOC systemic approach generates large operational and financial advantages for each node in the supply chain, without any undesirable collateral effect.

Keywords

Bullwhip Effect; Drum–Buffer–Rope; KAOS modeling; Multi-agent Systems; Supply Chain Management; Theory of Constraints.

1. Introduction

The complexity and dynamism that characterize the context in which companies operate nowadays have drawn a new competitive environment. In it, the development of information technologies, the decrease in transport costs and the breaking down of barriers between markets, among other reasons, have led to the perception that competition between companies is no longer constrained to the product itself, but it goes much further. For this reason, the concept of Supply Chain Management (SCM) has gained a lot of strength to the point of having a strategic importance. The current global economic crisis, consequence of many relevant systemic factors due to the fact that globalization still has not been able to develop systemic dynamic properties to deal with a growing variety of requirements, is creating conditions which increase awareness to adopt new approaches to make business (among others, Schweitzer et al., 2009); hence, SCM is a boiling area for innovation.

Analyzing the supply chain, Forrester (1961) noted that changes in demand are significantly amplified along the system, as orders move away from the client. It was called the Bullwhip Effect. He studied the problem from the perspective of system dynamics. This amplification is also evidenced in the famous ‘Beer Game’ (Sterman, 1989), which shows the complexity of SCM. He concluded that the Bullwhip Effect is generated from local-optimal solutions adopted by supply chain members. This can be considered as a major cause of inefficiencies in the supply chain (Disney et al., 2005), because it tends to increase storage, labor, inventory, shortage and transport costs. Lee et al. (1997) identified four root causes in the generation of Bullwhip Effect in supply chains: (1) wrong demand forecasting; (2) grouping of orders into batches; (3) fluctuation in the products prices; and (4) corporate policies regarding shortage. The same idea underlies behind all of them: the transmission of faulty information to the supply chain. Therefore, the first approaches in the search for a solution to this problem were based on trying to coordinate the supply chain. Some practices that have been successfully implemented in companies are Vendor Managed Inventory (Andel, 1996), Efficient Consumer Response (McKinsey, 1992) and Collaborative Planning, Forecasting and Replenishment (DesMarteu, 1998). Nevertheless, the Bullwhip Effect is still a major concern around operations management in the supply chain. Chen and Lee (2012) discussed the linkage between the bullwhip measure and the supply chain cost performance, capturing the essence of most-real world scenarios.

The Theory of Constraints (TOC) was introduced by Goldratt (1984) in his best seller 'The Goal', representing a major innovation in the production approach. The author alleges that the sole purpose of an organization is to make money now and in the future. Hereupon, the author defines six variables as organizational measures to approach that goal. Three of them are operational: throughput, inventory and operating expense. The other three are financial: net profit, return on investment and cash flow. All these metrics are bound together through relationships. According to TOC, the most important thing to improve the overall system performance is to concentrate the whole improvement effort on its bottleneck. Goldratt proposes the Drum-Buffer-Rope (DBR) methodology to manage the system. Once the bottleneck is identified, it becomes the drum of the system. A buffer is used to protect against variability in replenishment time, because we aim to exploit the full capacity in the bottleneck. A rope is used to subordinate the system to the bottleneck.

The major contribution of this paper is to provide evidence via a multi-agent simulation model about the sound impact of TOC application to reduce the Bullwhip Effect in supply chains. TOC is compared against a traditional management alternative, typical in mass production paradigm: the order-up-to inventory policy. Our aim is to demonstrate that supply chains have plenty of reasons to operate according to the TOC systemic approach. Figure 1 depicts the structure of our work.

The conceptual multi-agent model has been worked out using KAOS methodology. Robust SW engineering and test driven development techniques have been applied to build and verify the model. A multi-agent system (MAS) is an optimal environment to address this issue, as it is a physically distributed problem, where each node has only a partial knowledge about the problem-world.

As shown in figure 1, our research method has been the following:

i. Definition of problem world (‘Beer Game’ supply chain) and problem statement (Bullwhip Effect).
ii. Clarification of the process. The ‘Beer Game’ is modeled as it is widely described in literature (among others, Kaminsky and Simchi-Levi, 1998): the unique source of noise is the variability in demand; the Bullwhip Effect emerges as a consequence of the agents’ behavior; the metrics considered are the shortage

penalties and the inventory costs. Once the material and the information flows are implemented, two engines are added: TOC and the order-up-to inventory policy. The experimenter chooses what engine the agents in the supply chain will use to make their purchasing decisions.

iii. Devise the conceptual model using KAOS methodology.

iv. ABMS development of the model using NetLogo, followed by verification using statistical tests.

v. Exploitation of the model: experimentation of different treatments.

vi. Problem analysis: descriptive and inferential statistics to derive conclusions.

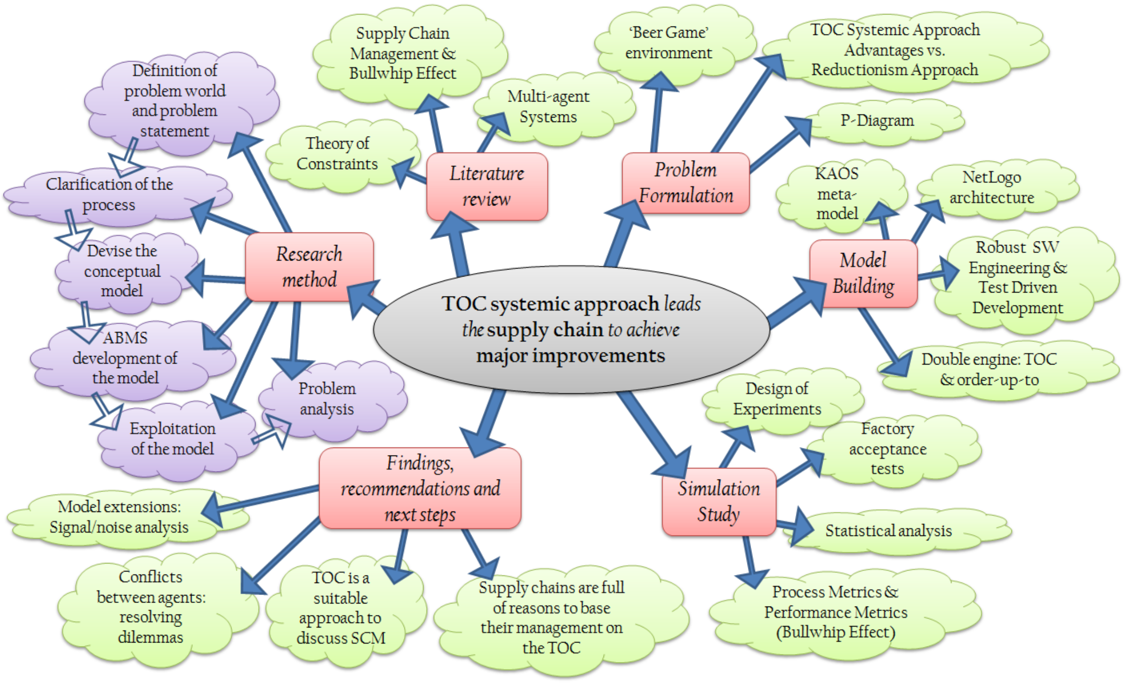


Figure 1. Structure of this work.

2. Literature Review

2.1. Theory of Constraints in Supply Chain Management

Elihayu M. Goldratt described in his book 'The Goal – A Process of Ongoing Improvement' (1984) his view about the best way to manage a company. He did it through fiction, telling how a troubled company managed to get over this situation. In a subsequent scientific work, Goldratt (1990) presented the Theory of Constraints (TOC) in more detail. This theory comprises three interrelated areas (Simatupang et al., 1997): logistics, logical thinking and performance measurement. In logistics, the methodology is based on the DBR scheduling method (Goldratt and Cox, 1984). The logical thinking is based on a continuous improvement cycle with five steps: (I) Identify the bottleneck; (II) Decide how to exploit the bottleneck; (III) Subordinate everything else in the system to the previous step; (IV) Elevate the bottleneck; and (V) Evaluate if the bottleneck has been broken, and return to the beginning. The performance measurement, which quantifies the application of this methodology, encompasses operational measures (throughput, inventory and operating expense) and financial measures (net profit, return on investment and cash flow), which obey to the same view: the only goal of the organization is to make money now and in the future.

Although TOC was initially oriented on the production system of the company, its application to other areas of the business has been proposed, such as marketing and sales (Goldratt, 1994), project management (Goldratt, 1997) or SCM (Goldratt et al., 2000). In this latter area, several authors have researched the application of the TOC. As an example, Umble et al. (2001) described the application of TOC in the implementation of an ERP system to manage the supply chain. Cox and Spencer (1998) proposed a method for SCM through TOC, valid when one company directs the entire chain. However, when this assumption does not apply and there are different companies in the same supply chain, the implementation of TOC is more complex. A dilemma rises because each company has to decide between gearing to the interests of the supply chain as a whole and pursuing only their own interests. Simatupang et al. (2004) showed that collaboration between different independent firms, according to the TOC, generates a much larger benefits to participants than the consideration of individual interests of each company.

Wu et al. (2010) developed an enhanced simulation replenishment model for TOC-SCRS (Theory of Constraints - Supply Chain Replenishment System) under capacity constraint in the different levels. The TOC-SCRS (Yuan et al., 2003) is a methodology widely used in businesses nowadays to improve the SCM and to reduce Bullwhip Effect. It is based on the use of two strategies (Cole and Jacob, 2002): (I) Each node holds enough stock to cover demand during the time it takes to replenish reliably; and (II) Each node orders only to replenish what was sold. The authors demonstrated the effectiveness of this system, in solving the conflict generated in determining the frequency and quantity of replenishment when the TOC- SCRS is applied in a plant or a central warehouse. In a later work (Wu et al., 2014), they proposed a two-level replenishment frequency model for the TOC-SCRS under the same constraints, which is especially suitable to a plan in which different products have a large sales volume variation. This methodology facilitates a plant or a central warehouse the implementation of TOC-SCRS.

2.2. Multi-Agent Systems in Bullwhip Effect reduction

MASs is a branch of Artificial Intelligence that proposes a model to represent a system based on the interaction of multiple intelligent agents (Wooldridge, 2000). Each agent evaluates different alternatives and makes decisions, in a clearly defined context, through local and external constraints. De la Fuente and Lozano (2007) defend this methodology in the study of SCM, based on its own characteristics: it is a physically distributed problem; it can be described a general pattern in decision-making; each agent can consider both individual and chain interests; and it is a highly complex problem, which is influenced by the interaction of many variables. For this reason, since the work of Fox et al. (1993), who were pioneers in representing the supply chain as a network of intelligent agents, many studies have followed this line.

Maturana et al. (1999) used the multi-agent architecture to create the Metamorph tool. It was aimed at facilitating the SCM in business through the introduction of intelligence in the design and manufacturing stage. Later Kimbrough et al. (2002) studied the agent's capability of managing their own supply chain. The authors concluded that they can determine the most appropriate policy for each level, achieving a large reduction in the Bullwhip Effect generated along the system. Some years later, Mangina and Vlachos (2005) designed a smart supply chain in the food sector. They demonstrated that agents increase the supply cain's flexibility, information access and efficiency. Liang and Huang (2006) developed a MAS to forecast the demand along a supply chain where each level has a different inventory policy. To calculate the forecast, they used a genetic algorithm. Fuzzy logic was introduced into the analysis by Zarandi et al. (2008). The authors constructed an agent-based system for SCM in dim environments. One of the latest studies on the subject is the one by Saberi et al. (2012), who analyzed the chain collaboration. In their work, the agents coordinate to make forecasts, to control the stock and to minimize total costs. Recently, Chatfield and Pritchard (2013) constructed a hybrid model of agents and discrete simulation in order to represent the supply chain. It was studied in several scenarios and they showed that returns of excess goods increase significantly the Bullwhip Effect.

The literature review leads us to conclude that multi-agent methodology is widely used to experiment around complex systems, such as supply chains. More specifically, it contains several works which apply these new technologies to analyze the well-known problem of the Bullwhip Effect. Likewise, the application of TOC has been studied to improve the management in complex systems, including supply chains. However, the authors are aware of multiple real supply chains and know it is not common to apply Goldratt's theory. The systemic thinking prompts the actors to solve a major dilemma, which consists on that the methods of measurement, linked to reward and punishment policies, in the supply chain are not usually defined from a systemic perspective, but from the relationships between each pair of nodes in the chain. Therefore, our aim is to compare the holistic TOC method against a traditional reductionist alternative –the ‘order-up-to’ inventory policy– from a multi-agent approach.

3. Problem Formulation

The Bullwhip Effect gained much importance when, in the early 90's, Procter & Gamble noticed that their demand for Pampers diapers suffered considerable variations throughout the year, which did not correspond to the relatively constant demands of its distributors –in addition, the swings of its suppliers were greater (Lee et al., 1997). Since then, this phenomenon has been a fruitful research area within logistics studies. Nevertheless, at present, it is one of the main concerns for business regarding to SCM. As way of example, Buchmeister et al. (2012) illustrate this phenomenon using real data in three simulation cases of a supply chain with different level constraints (production and inventory capacities).

In our study, we have considered a traditional single-product supply chain with a linear structure, composed of five levels: client, shop retailer, retailer, wholesaler and factory, as the one used in the ‘Beer Game’. Among the levels, there are two main flows: the material flow (related to the shipping of the product) from the factory to the client, and the information flow (related to sending the orders) from the client to the factory. Thus, there are five main actors. Four of them (shop retailer, retailer, wholesaler and factory) are responsible for managing the supply chain, in order to meet the other’s (customer) needs.

The only purpose of the supply chain is, according to TOC, to make money, now and in the future. To assess the approximation of a company to this goal, the author proposes three financial metrics: net profit, return on investment (ROI) and cash flow. These metrics must be understood as complementary indicators. Thereby, improving the SCM requires the simultaneous increase of the three values. The next question is: how can the supply chain achieve it? Then, a second level of goals appears: (I) improve customer satisfaction; (II) improve the efficiency of the supply chain; and (III) improve the utilization of the capacity.

Here, we can link our analysis with the TOC, considering three operational metrics: throughput (the rate at which system generates money through sales), inventory (money invested in purchasing items intended to be sold) and operating expense (money spent in order to turn inventory into throughput). Customer satisfaction is a big contributor to throughput; increased efficiency means a decrease in operating expense; and improving capacity usage implies achieving good results in the inventory. This operational metrics can also be used to quantify the results of the supply chain, as the financial ones can be understood as a direct consequence of these.

How do we attain these three goals of the second level? To increase customer satisfaction, the key element is minimizing missing sales. Our model does not consider the effect of other factors, such as marketing. The client will be satisfied if he finds what he needs in the shop retailer when he needs. To improve supply efficiency and capacity utilization, the chain needs to reduce the Bullwhip Effect that causes an amplification of the demands variability of levels upstream, which hinders both transportation and inventory management. Thus, the decrease of the Bullwhip Effect brings the system to improve its operational, and consequently, financial metrics.

Many authors quantify the Bullwhip Effect in a level *n* of the supply chain as the quotient between the variance of the purchase orders launched (${\sigma_{POE}^2}^n$) and the variance of the purchase orders received (${\sigma_{POR}^2}^n$), adjusted both the numerator and denominator by the mean value (${\mu_{POE}}^n$, ${\mu_{POR}}^n$), according to equation 1. For stationary random signal, in a linear supply chain, over longs periods of time, both means values are the same. It should be noted that the purchase orders received by the shop retailer are the sales orders, which meet the demand of the customer, and that purchase orders emitted by the upper level of the supply chain (factory) translate in their own production. As the purchase orders launched by each level are the sale orders received by the next one, the total Bullwhip Effect generated in the supply chain (${BE_{orders}}^{sc}$) can be expressed as the product of the

Bullwhip Effect in the four different levels, by equation 2. When this ratio is higher than 1, there is Bullwhip Effect in the supply chain.

$$BE_{orders}{}^{n} = \frac{\sigma_{POE}^{2}{}^{n} / \mu_{POE}{}^{n}}{\sigma_{POR}^{2}{}^{n} / \mu_{POR}{}^{n}} = \frac{\sigma_{POE}^{2}{}^{n}}{\sigma_{POR}^{2}{}^{n}} \qquad (1)$$

$$BE_{orders}{}^{sc} = \prod_{n=1}^{4} BE_{orders}{}^{n} \qquad (2)$$

This is a useful measure to quantify the evolution of orders, but only compares output variance with input variance, and does not describe the structure that causes the variation increase. For this reason, some authors (among others, Disney and Towill, 2003) also recommend the use of an alternative measure of the Bullwhip Effect at each level *n* of the supply chain ($BE_{inventory}{}^{n}$), which quantifies fluctuations in actual inventory. It can be expressed as the quotient of the variance of the stock ($\sigma_{STOCK}^{2}{}^{n}$) and the variance of the demand ($\sigma_{POR}^{2}{}^{n}$), by means of equation 3. It is important to note that they are complementary measures. That is to say, to improve the SCM is necessary to reduce the two of them, and not just one at the expense of the other.

$$BE_{inventory}{}^{n} = \frac{\sigma_{STOCK}^{2}{}^{n}}{\sigma_{POR}^{2}{}^{n}} \qquad (3)$$

The goals of this level face two major obstacles of the SCM: uncertainty in both demand and lead time. Uncertainty in the final customer demand is modeled through various statistical distributions. Lead time is modeled constant, as stated in the ‘Beer Game’. Obviously, if orders lead time and material lead time were both null, the supply from the factory would instantly respond to customer requirements and Bullwhip Effect would not rise. The only relevant controllable factor (parameter) in our model is the engine to be used by agents to make their purchasing decisions. For the sake of simplicity, we have not considered other causes of the Bullwhip Effect, as the uncertainty in the lead time or variation in prices.

Figure 2 points out the p-diagram (parameter diagram –a widely used tool in robust engineering) that we have used to establish the perimeter of our study. In it, we can see the overall supply chain function, the noise sources that threaten the system function, and the parametric space, which are controllable factors either at engineering stage or manufacturing stage.

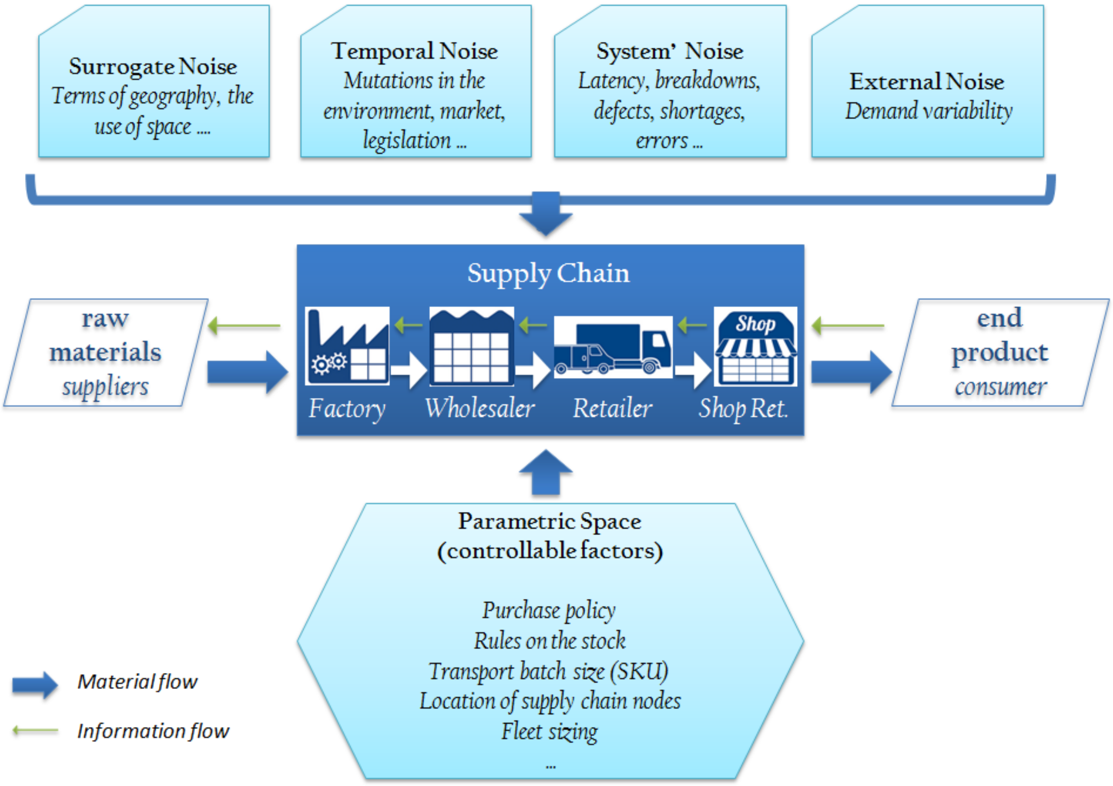


Figure 2. P-diagram of the system that we have developed.

4. Description of the Multi-Agent System

We have used KAOS methodology (Dardenne et al., 1993) for the conceptual design. It is an engineering methodology that joins, in the development of a software application, the overall objective that should be met and the specific requirements that should be considered. This methodology relies on the construction of a requirement model, whose graphical part can be represented by means of the KAOS Goal Diagram. Figure 3 shows the KAOS Goal Diagram that we have created and used in the development of the system.

TOC approach consists on managing the supply chain based on the bottleneck. This is one of the foundations of the TOC: any improvement that is deployed away from the bottleneck of a system represents a waste of resources. Therefore, this fact leads to a new question: Where is the bottleneck in this supply chain? The factory would be the bottleneck if its production rate cannot cover the customer demand. But the factory has not a capacity constraint in the 'Beer Game'. The intermediate nodes, wholesaler and retailer, could be the bottleneck if its storage or transport capacity did not allow the supply chain to meet the final demand, but this is not the situation that we have considered. So, the bottleneck is the final customer demand. To maximize the flow at the bottleneck means to have zero missing sales at the shop retailer. Therefore, the drum is placed at the shop retailer.

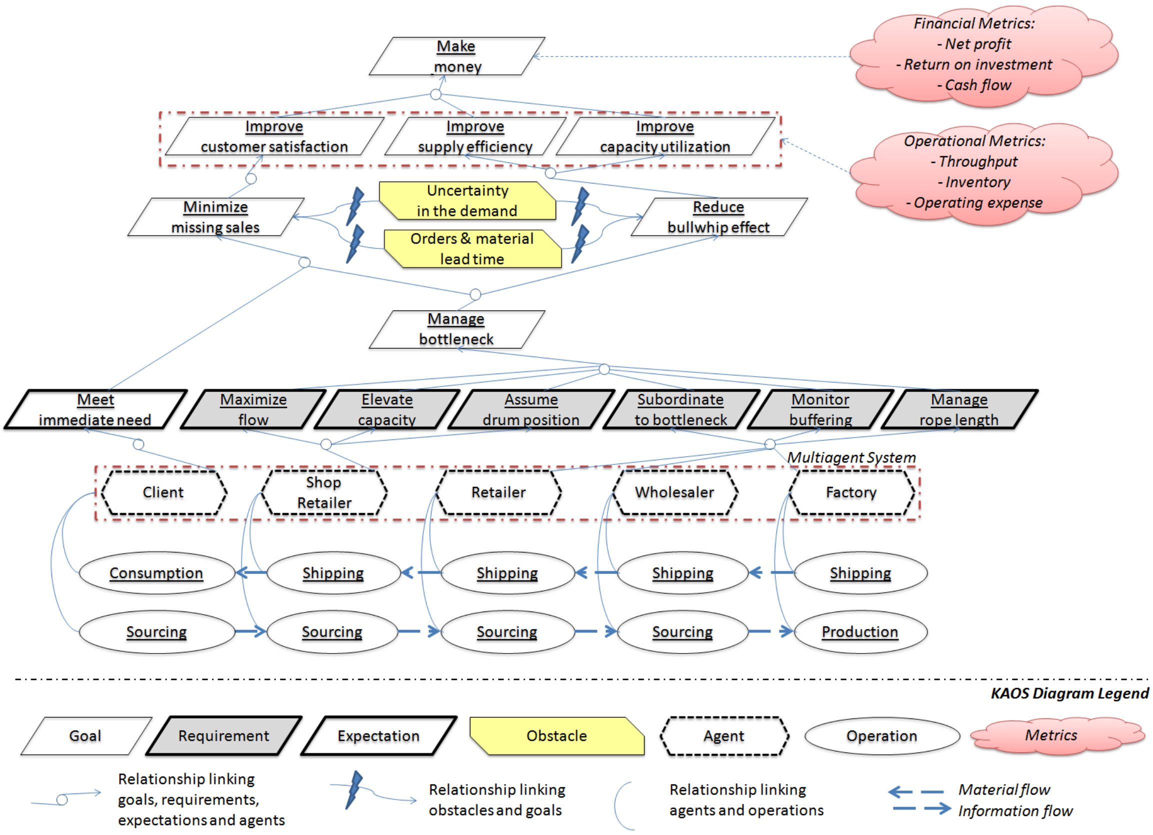


Figure 3. KAOS Goal Diagram of our MAS.

Each time that a demand event is triggered to the system, the drum makes all the agents react. Each agent (node) calculates its rope length to the drum position and makes the order decision based on its downstream buffer to the bottleneck. Instead of traditional safety stock based on material quantities, TOC-based buffers are a function of the lead time. Buffer management consists on moving the flow so that arrival happens on time at the bottleneck. Because the shop retailer is the drum, this agent looks for maximizing flow; which means preventing missing sales by linking the final customer demand forecast straight to the factory. All other nodes work subordinated to the drum with a shipping rope.

Each node works using a finite state machine schema. The agent is idle until the drum triggers it. From the idle state it switches to serve backorders state. Then, it flows to the shipping orders state. Once the agent has moved material downstream, it moves to the sourcing state (take care of information flow). Finally, the agent moves to the reporting state, when it cares about updating and exporting information. And then the agent switches now to the idle state to reiterate the loop. The state transition diagram is represented in figure 4.

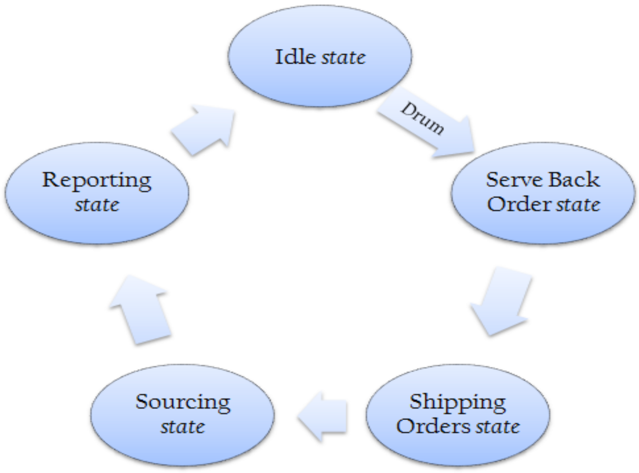


Figure 4. State transition diagram (local for each agent).

Some details about our simulation engine should be commented. The simulation clock advances based on a FEL (future event list). Events are scheduled in the future and the clock advance will move to the event which is sooner due. Every takt (block of time between two consecutive arrivals of customers to the shop retailer) schedules the next one. Each customer arrival schedules new events in the FEL so to divide each time bucket into small time windows. Synchronizing mechanisms are used to force nodes to follow a downstream sequence for material flow and an upstream sequence for the orders flow.

During these sequences agents transition their states to perform all the activities: move material downstream, move orders upstream, serve backorders just in case, serve the current order, place backorder if needed, place its purchase order upstream (according to the settings for the order policy), and report data into the export file. Of course the system behaves polymorphous depending on the setting of the experiment. This means that details of what each node does at each state follows the appropriate rules linked to the parameters given at the setup stage.

We have used robust SW engineering techniques (Taguchi, 2000) to build the model and NetLogo 5.0.5 to implement it. Figure 5 shows a screenshot of the interface window of the implemented model. The interface window provides the experimenter with the animation frame, the controls to setup parameters and to run each experiment, and the graphics and monitoring stuff to track what the system is doing. NetLogo provides two additional windows, one for the model documentation and another for the model code.

In the next paragraphs we will clarify some relevant details about what the system does when operating under TOC parameters and when the order-up-to policy is the selection made by the experimenter.

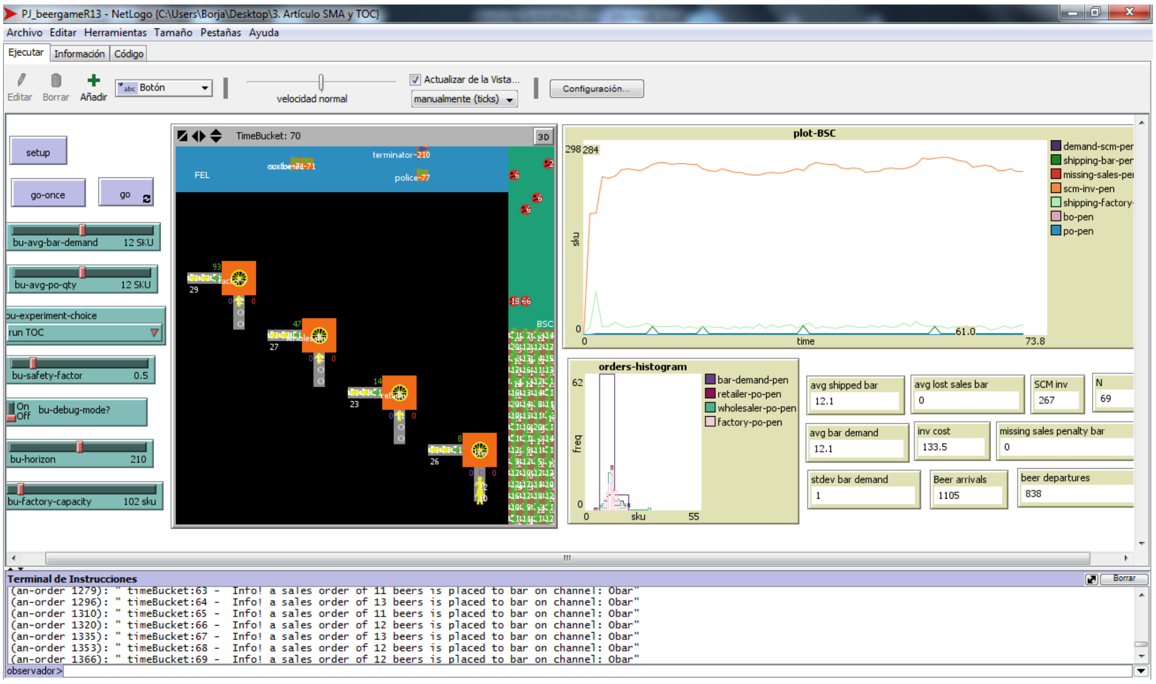


Figure 5. Screenshot of the system interface at one particular moment of the simulation.

4.1. Order-up-to inventory policy

This policy is implemented as follows: at the end of each period *t*, the shop retailer, retailer, wholesaler and factory update the forecast $(\widehat{D_t})$ based on the demand or order received, by means of a moving average of the last three observations (D_{t-i}), according to equation 4.

In this policy, under the assumption of normal demand, the order-up-to point (y_t) is estimated as the product of the forecast and the lead time (L), plus a term related to the safety stock (equation 5). It depends on a parameter (Z) that is a function of the security level and the standard deviation of the error (S_t). We have used $Z = 1.64$ in order to work with a confidence level of 95%. The purchase order quantity for each period is the difference between the order-up-to point of this period and the previous one, plus the demand of the previous period, by equation 6. Note that the purchase order arrives at the start of period $t+L$ and sales orders are filled at the end of each period. More information about this management policy can be found in Chen et al. (2003). In our case, we have used a three period moving average to calculate the forecast.

$$\widehat{D_t} = \frac{1}{n} \cdot \sum_{i=1}^{n} D_{t-i} \tag{4}$$

$$y_t = L \cdot \widehat{D_t} + Z \cdot \sqrt{L} \cdot S_t = L \cdot \widehat{D_t} + Z \cdot \sqrt{L} \cdot \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (D_{t-i} - \widehat{D_{t-\iota}})^2} \tag{5}$$

$$q_t = y_t - y_{t-1} + D_{t-1} = \left(1 + \frac{L}{n}\right) \cdot D_{t-1} - \left(\frac{L}{n}\right) \cdot D_{t-(n+1)} + Z \cdot \sqrt{L} \cdot (S_t - S_{t-1}) \tag{6}$$

4.2. DBR methodology — Goldratt's TOC policy

The DBR methodology has been implemented according to the Goldratt's TOC, summarized in section 2 and following to the meta-model explained above. We should remember that, in the context we are considering, the shop retailer is the constraint in the system, so it must be the drum. The aim of the solution is to protect it, and therefore the supply chain as a whole, against process dependency and variation, and thus to optimize the system. In these circumstances, the other levels must be subordinated to the shop retailer. The buffer is the material release duration and the rope is the release timing. Kelvyn Youngman (2009) has developed an outstanding guide for the implementation of the TOC in systems of very different kinds, which can be consulted to get further detail in the process described below.

In the TOC mode, the system operates in two stages. In the first one, the systemic condition to tie the different levels of the supply chain through time (and not by product) is established. It is the planning stage and it is orientated to operate the system as a whole. In the second one, the buffer is administered along the intermediate stations, to guide the way in which the motor is tuned for peak performance. It is the control stage that allows us to keep a running check on the system performance. The idea is summarized in figure 6.

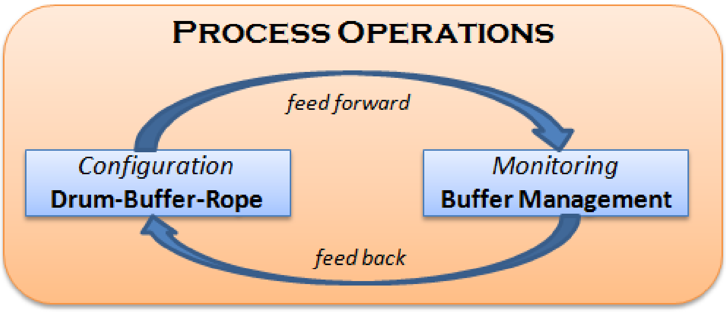


Figure 6. Two-stage based operation system.

With the previous objective, at each time unit, the factory uses the history of the demand in the shop retailer (the time interval defined by the rope, which is the period of time to protect), in order to decide the production orders that must be placed in the channel (the manufacturing time is equal to the lead time in the remaining levels: 3 periods). Subsequently, each node of the supply chain, except the shop retailer (as no other level can be found downstream) manages the buffer. The horizontal channels are the buffer of the model. The buffer is time and material flow, but not the order flow. Manage it means compensating in each *takt* the flow dissipated downstream after shipping. Therefore, for example, in the case of the factory, the buffer is 9 time units (lead time of 3 units in the previous three levels). Unlike classical policies, the TOC orders are dosage orders into the buffer and they are dissipative. They have no lead time, because each agent decides what to dose subordinated to the bottleneck. They do not generate backorders, as the next dosage again obey the bottleneck. Figure 7 graphically represents this idea, showing the drum, the buffer and the rope.

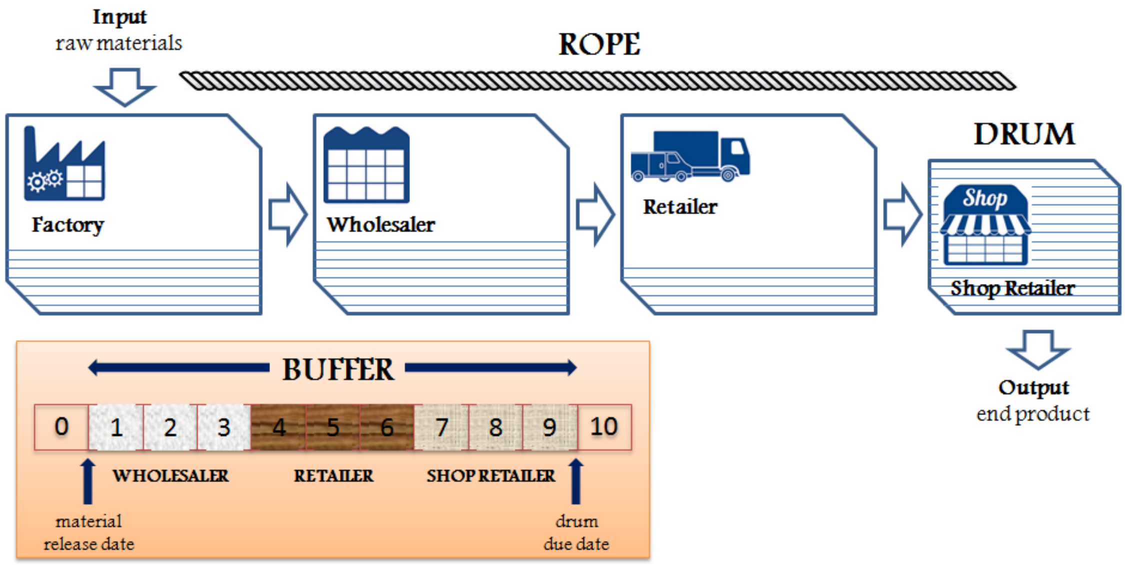


Figure 7. Schematic representation of the MAS when it works according to the TOC.

5. Simulation Study and Conclusions

As the equations related to the inventory policy that we have used to contrast the results are based on the assumption of normal demand, we have simulated the customer demand through a normal distribution with a mean of 12. We have performed treatments on three different scenarios: when the variability is low (standard deviation of 1; coefficient of variation 8.3%), when the variability is moderate (standard deviation of 3; coefficient of variation 25.0%), and when the variability is high (standard deviation of 5; coefficient of variation 41.7%), in order to extend the conclusions considering the effect of the demand variability in the SCM. Thus, our experimentation approach, can be written as shown in equation 7, where Y is a vector of the key performance indicators (in terms of Bullwhip Effect); X is the policy management, which is a nominal attribute variable (order-up-to inventory policy or DBR methodology); Z is an external noise condition, which is characterized for de experiment as $N(12, \sigma)$, where σ is set to three different levels in order to represent different levels of variability with respect to the average demand; and ξ represents the residuals –the unexplained part of the system response.

$$Y = f(X, Z) + \xi \tag{7}$$

So, it is a full DoE (Design of Experiments) with two factors. One factor (order policy) is controllable and is taken at two levels; while the other factor (demand law) is noise and enters the simulated experiment at three levels. This idea is shown in table 1.

Table 1. DoE (Design of Experiments) table.

Factor	*Level*	*Treatment*	*Demand Law (Z)*	*Order Policy (X)*
Demand Law	Normal(12,1)	1	Normal(12,1)	Order-up-to i. p.
(Z)	Normal(12,3)	2	Normal(12,3)	Order-up-to i. p.
	Normal(12,5)	3	Normal(12,5)	Order-up-to i. p.
Order Policy	Order-up-to i. p.	4	Normal(12,1)	DBR methodol.
(X)	DBR methodol.	5	Normal(12,3)	DBR methodol.
		6	Normal(12,5)	DBR methodol.

A time horizon of 330 periods was used for each treatment. The first 30 are discarded as warm-up period, so to avoid the initial transitory that can alter the results. On the other hand, the 300 remainder periods is a large enough time interval to check stability according to the common practices.

5.1. Model verification and validation

A fundamental step in any modeling process is the verification of the model, with the aim of checking its cohesion and consistency; that is, to check that the development matches the logic of the conceptual design. This model was created following strict rules of clean code, test driven development focus, versioning for continuous functionality increments, and it uses failure modal analysis in order to prevent failures. Although these good practices of software engineering reduce the probability of error, they do not eliminate it completely. Therefore, we have complemented it with mechanics (exception handling, cross checking, police agents for system audits) for early detection of any system malfunction.

Another essential step in simulation process is the validation phase. The experimenter wants model predictions to match reasonably well the reality, so that the simulation model is useful to devise changes and apply them to improve the real system. To validate our model, we have used factory acceptance test (FATs), so to confirm that the model exhibits a well-known behavior when exposed to controlled conditions. As an example, we include one this kind of tests that are implemented in the model.

Test conditions:	(I) Constant demand in the shop retailer: 12 sku / period.
	(II) Damaged equipment on the factory: zero production.
Expected behavior:	(I) It only serves customers until the initial stock is depleted.
	(II) Cumulative backorders are generated at each node.
Acceptance criteria:	(I) Demand turns into missing sales (12 sku / period) in steady state.
	(II) Storage costs are zero in steady state.

Once the FAT tests were satisfactory, the standard approach was used when comparing treatments under stochastic conditions: each treatment is replicated (it was run three times) so that the statistical analysis takes into account the experimental error. An overall stability study (run several trajectories –replicas– of each experimental treatment) about the key output metrics (lost sales, stocks) was also conducted. And, of course, we did care about the experimental error (using replicas and hypothesis testing).

The model statistically probed to be valid: matched expected outputs under controlled scenarios, reached stability and have repeatability.

5.2. Analysis of the treatments

Tables 2, 3, 4 and 5 report the final results of the treatments, both the outcomes exported from the simulation (process metrics) and the results of the simulations in terms of Bullwhip Effect and missing sales (performance metrics).

Tables 2 and 3 demonstrate the huge generation of Bullwhip Effect along the supply chain when using the order-up-to inventory policy. Whilst the quantity order average remains constant along the supply chain nodes (it only varies slightly due to missing sales and inventory accumulation), the quantity order variance increases greatly as we move upstream. It is interesting to see that the average inventory increases dramatically upstream the chain. Nevertheless, the amount of missing sales is noteworthy. As a conclusion, with the order-up-to policy the service level to customers is not extremely bad (still, it is not excellent), and the weak point is that this bad service is obtained at a huge cost in terms of inventory. The lesson learnt, and it is very usual in the marketplace, is that the customer service is protected with huge inventory and this policy is not effective, because the root cause of the problems is not being considered. According to the industrial experience of the authors, this is a very common finding in ailing processes.

Table 2. Results of the tests when the order-up-to inventory policy is used (I): Mean (left) and variance (right) of the consumer demand, purchase orders (PO), factory production and inventory in the different levels of the supply chain (without warm-up time).

Process Metrics	*Scenario 1* *Low variability* *[Treatment 1]*	*Scenario 2* *Mid variability* *[Treatment 2]*	*Scenario 3* *High variability* *[Treatment 3]*
Consumer Demand	11.98 – 1.04	11.97 – 7.97	11.91 – 27.61
Shop Retailer PO	11.47 – 98.39	11.49 – 133.53	11.64 – 232.13
Retailer PO	12.04 – 380.20	11.79 – 715.74	12.50 – 1008.79
Wholesaler PO	11.79 – 1405.58	13.17 – 1994.30	13.47 – 3304.94
Factory Production	12.08 – 4247.31	14.15 – 4162.65	13.03 – 7228.66
Shop Retailer Inv.	12.0 – 101.1	19.2 – 215.9	34.9 – 613.6
Retailer Inv.	67.9 – 1011.38	105.1 – 4429.3	154.5 – 8362.3
Wholesaler Inv.	218.9 – 13471.1	384.1 – 22900.2	559.9 – 51286.0
Factory Inv.	577.7 – 32599.2	593.1 – 13674.0	1057.0 – 137635.3

Table 3. Results of the tests when the order-up-to inventory policy is used (II): Orders Bullwhip Effect (Bw_O) and Inventory Bullwhip Effect (Bw_I) generated along the different nodes, in addition to lost sales to evaluate the performance of the supply chain (without warm-up time).

Process Metrics	*Scenario 1 Low variability [Treatment 1]*	*Scenario 2 Mid variability [Treatment 2]*	*Scenario 3 High variability [Treatment 3]*
Shop Retailer Bw_O	99.13	17,47	8.60
Retailer Bw_O	3.68	5,22	4.05
Wholesaler Bw_O	3.78	2,49	3.04
Factory Bw_O	2.95	1,94	2.26
Supply Chain Bw_O	4063.14	442.07	239.33
Supply Chain Lost Sales	163	124	86
Shop Retailer Bw_I	97.58	27,10	22.22
Retailer Bw_I	10.28	33,17	36.02
Wholesaler Bw_I	35.43	32,00	50.84
Factory Bw_I	23.19	6,86	41.65

Looking at these tables, it can be seen that the greatest Bullwhip Effect is generated, according to the classical formulation, in the scenario of low variability. Obviously, the greater the variability in consumer demand, the greater the variability in the rate of production of the factory. However, the relationship between the two variances is much larger when the variability in consumer demand is low. Moreover, this classic inventory management policy generates more missing sales when the variability of consumer demand is low. At first glance, this result might seem surprising, but it is not, as the explanation lies in the level of inventories: when the variability is very high, the levels of the supply chain tend to be overprotective. For this reason, the missing sales are reduced at the expense of increasing the inventory far from the customer.

Tables 4 and 5 point out that the TOC also causes Bullwhip Effect in the supply system, since variability in purchase orders increases and both the mean and the variance of the inventory level increment as they move away from the consumer. However, a simple comparison of these tables with respect to tables 1 and 2 makes clear the enormous effectiveness of DBR methodology in managing the supply chain. The amplification of the variability of orders is much lower when the supply chain is managed according to the practices proposed by Goldratt. Likewise, the TOC gets to manage the supply chain with minor inventories at all levels. Moreover, despite that, the amount of missing sales decreases meaningfully. Hence, the important findings using TOC approach is that both negative effects (Bullwhip Effect and missing sales) reduce at the same time when compared to the order-up-to policy.

Table 4. Results of the tests when the DBR methodology is used (I): Mean (left) and variance (right) of the consumer demand, purchase orders (PO), factory production and inventory in the different levels of the supply chain (without warm-up time).

Process Metrics	*Scenario 1 Low variability [Treatment 4]*	*Scenario 2 Mid variability [Treatment 5]*	*Scenario 3 High variability [Treatment 6]*
Consumer Demand	12.07 – 1.13	12.47 – 11.03	11.79 – 24.43
Shop Retailer PO	12.10 – 9.11	13.04 – 75.82	12.83 – 134.10
Retailer PO	12.10 – 7.32	12.33 – 58.37	11.66 – 101.48
Wholesaler PO	12.09 – 5.63	12.36 – 53.60	11.47 – 110.75
Factory Production	12.09 – 7.98	12.47 – 76.48	11.39 – 145.03
Shop Retailer Inv.	9.2 – 12.5	16.8 – 74.1	21.9 – 142.9
Retailer Inv.	14.0 – 23.8	18.6 – 140.4	20.6 – 209.7
Wholesaler Inv.	50.7 – 17.2	56.5 – 190.7	59.3 – 523.7
Factory Inv.	97.1 – 18.0	113.6 – 162.0	121.0 – 441.1

Table 5. Results of the tests when the DBR methodology is used (II): Orders Bullwhip Effect (Bw_O) and Inventory Bullwhip Effect (Bw_I) generated along the different nodes, in addition to lost sales to evaluate the performance of the supply chain (without warm-up time).

Process Metrics	*Scenario 1 Low variability [Treatment 1]*	*Scenario 2 Mid variability [Treatment 2]*	*Scenario 3 High variability [Treatment 3]*
Shop Retailer Bw_O	8.02	6.57	5.05
Retailer Bw_O	0.80	0.81	0.83
Wholesaler Bw_O	0.77	0.92	1.11
Factory Bw_O	1.42	1.42	1.32
Supply Chain Bw_O	7.03	6.94	6.15
Supply Chain Lost Sales	1	54	82
Shop Retailer Bw_I	11.01	6.72	5.85
Retailer Bw_I	2.61	1.85	1.56
Wholesaler Bw_I	2.34	3.27	5.16
Factory Bw_I	3.19	3.02	3.98

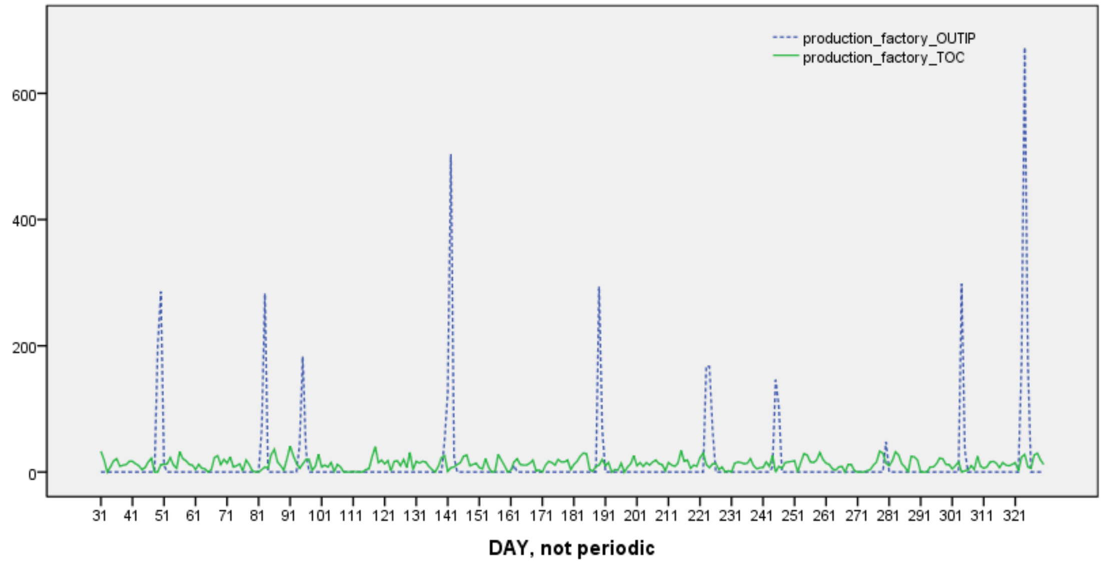


Figure 8. Factory production in the two tests (order-up-to inventory policy and DBR methodology) carried out with a N(12,3).

The generation of the Bullwhip Effect in the supply chain and the improvements introduced by Goldratt's practices in comparison with the traditional management policies can be shown graphically in many different ways. For example, figure 8 exhibits the production rate of the factory throughout the time horizon for the two tests assuming normal with mean 12 and standard deviation 3 in the final consumer. When the system works according to the order-up-to inventory policy, the factory production varies greatly: in most periods, it has no production needs while in some specific moments it must manufacture very high amounts of product. With the DBR methodology, however, variability in the factory production is much lower, which translates in cost savings from different perspective (among others, labor, inventory, and transportation costs).

Why does such amplification occur? When the supply chain is managed according to the order-up-to inventory policy, the peaks in orders received for each level translate into an even bigger peak in orders placed by that level. The time difference is the lead time. That is to say, each level contributes increasing the distortion in the supply chain, and so decreasing the reliability of the transmitted information. When using TOC, the supply chain performs dramatically better.

The other way to observe the Bullwhip Effect is through the inventory of the various levels. It is possible to see it, for example, by means of box plots. Figure 9 shows these graphs, with the average, the indicators of the first and third quartile and the upper and lower limits, for the stock of the different members of the supply chain in tests with mean 12 and standard deviation 5. It should be noted that the values lower than 0 are related to inventory backorders that will be met the following periods. It is enough to compare the vertical scale of the two graphs to observe the improvements introduced by TOC, both in mean and in variance.

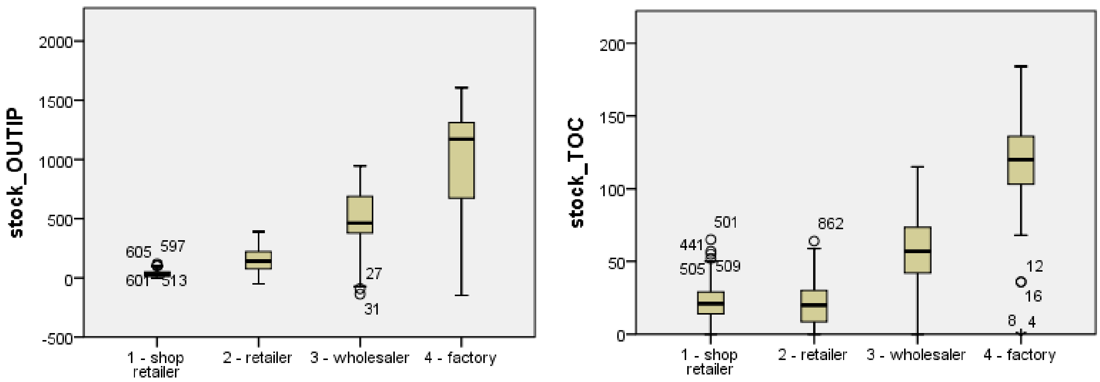


Figure 9. Box plots of the inventory level in the different members of the supply chain in the two tests (Order-up-to inventory policy and DBR methodology) carried out with a N(12,5).

5.3. Statistical significance of results

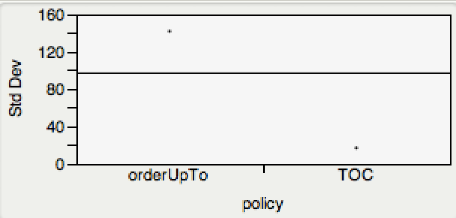
By looking at the plots shown above, we have visual evidence that the supply chain performs much better when using TOC, as commented. Nevertheless, it should be formally verified. The statistical tests were conducted for the different treatments, although they are only shown in one case, by way of example.

First, we concentrate on missing sales at the shop retailer, which is the only point where the fact of missing sales is really a critical concern. When the standard deviation of the demand is 5, we have the distribution for the missing sales penalty in each time bucket (sample size N > 150, once excluded the warm-up period). We have tested the null hypothesis "*H0: missing sales mean = 0*". For the order-up-to inventory policy, using 1-sample t test has a pValue less than 5%, which rejects null hypothesis. So, the penalty for missing sales is significantly different from zero. On the other hand, running a same length trajectory with TOC, all time buckets, after the warm-up period, have zero lost sales. The conclusion is that TOC policy effectively protects the supply chain against losing sales, whilst this does not happen with the order-up-to policy.

Once we have got formal evidence that the supply chain performance significantly improves when applying TOC in terms of external customer satisfaction (here, maximizing sales by exploiting the bottleneck), we now take care of getting also formal evidence that this achievement is not at the expense of increasing inventory cost in the overall supply chain. The inventory total cost has been collected during a long (for example, 200 time buckets) period of time after the system warm-up, and proceed first to check is the variance of this metric is unequal when using TOC versus when using order-up-to policy. We check, using a 2-variance test, the null hypothesis "*H0: variance (total inventory cost in the supply chain) | policy = TOC) = variance (total inventory cost in the supply chain) | policy = order-up-to)*". Figure 10 shows that in the sample, the standard deviation statistic of the metric at TOC level is less than at order-up-to level; the Levene test shows a p-value lower than 5%; so we reject null hypothesis. Therefore, TOC policy induces less variance in the inventory cost (so, to the goal stock in the system).

Figure 10 also displays the Welch's test to compare the means. Again, we reject the null hypothesis "*H0: mean (total inventory cost in the supply chain) | policy = TOC) = mean (total inventory cost in the supply chain) | policy = order-up-to)*". And, we take the alternative hypothesis: the total inventory cost in the supply chain is less when we use TOC policy. In conclusion, as expected, TOC not only gives a full protection against missing sales (while order-up-to does not), but besides, TOC achieve this result even reducing the total inventory cost (less variance and lower mean).

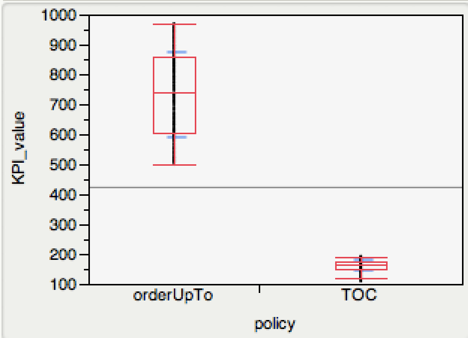
Tests that the Variances are Equal



Level	Count	Std Dev	MeanAbsDif to Mean	MeanAbsDif to Median
orderUpTo	159	142.1137	125.4534	125.4088
TOC	189	17.0923	13.8858	13.6825

Test	F Ratio	DFNum	DFDen	p-Value
O'Brien[.5]	267.7306	1	346	<.0001*
Brown-Forsythe	520.6055	1	346	<.0001*
Levene	525.9818	1	346	<.0001*
Bartlett	529.5158	1	.	<.0001*
F Test 2-sided	69.1304	158	188	<.0001*

Inv cost for the overall SCM



Welch's Test

Welch Anova testing Means Equal, allowing Std Devs Not Equal

F Ratio	DFNum	DFDen	Prob > F
2538.7693	1	161.85	<.0001*

t Test
50.3862

Figure 10. Hypothesis contrast to the significant difference between the inventory costs and averages of both policies.

6. Findings, Recommendations, and Next Steps

The new competitive environment has granted the Supply Chain Management a strategic role in the search for competitive advantage. For this reason, the orders variance amplification along the supply chain, known as the Bullwhip Effect, is an important concern for businesses, as it is a major cause of inefficiencies. Traditional management policies linked to the mass production paradigm, such as order-up-to inventory policy, are unsuccessful –as already shown in literature– in terms of fighting the Bullwhip Effect.

KAOS methodology was used to devise the multi-agent simulation model carried out on this research. The Gall's incremental principle (a complex system that works properly has evolved from a simple system which was effective) has been applied to end up with a highly reliable, self-controlled, tested and flexible model so to experiment TOC approach versus order-up-to policies for managing a multi-echelon supply chain and collect data evidence of system behavior. Statistical analysis has been applied to these data blocks taking into account the warm-up period, stability study and the final hypothesis testing to raise our conclusions.

Our first finding was that the higher the final customer demand variability, the higher is the amplification upstream the supply chain, because each node tends to overprotect itself due to the fear of breaking stock.

TOC philosophy has demonstrated in this work that is highly effective in remedying this issue. A dramatic improvement in the overall supply chain has been reached in several explored levels of external demand variability, but the more important point is that every level has improved its own performance by subordinating to the bottleneck. Hence, the best solution for the system is the best solution for each individual member.

The major contribution of this work has been to demonstrate that considering only the main effects, there are enough reasons to manage the supply chain according to Goldratt's philosophy.

There are plenty of model extensions and future works that this research group is planning as next steps on this fascinating topic.

i. To analyze why, provided that TOC is a mature and validated theory, it is not yet widely used. We wonder that moving the agents away from their natural egoist behavior needs some educational phases, and simulation can play an important role here.

ii. To extend this model to a larger noise conditions scenario. Now the noise factors have been limited in the model to include only different levels of variability in the external demand and to keep constant the delays in the material and in the information flows. Of course, considering other disturbance factors like scrap, variability in transportation delays, errors in the information flow and other sources of waste in the supply chain, a comparison of system robustness using TOC versus other management policies can provide insights to other relevant findings.

iii. To place SCM rules and controls to prevent selfish behavior of agents that could operate against the supply chain major interests. We also plan to explore to what extent agents applying fuzzy logic decision in their quest of local optima compares against applying holistic fuzzy logic decision making engines. Thereby, the concept of the Nash Equilibrium in supply chains must be introduced.

iv. To model adaptive mechanisms on the supply chain in order to detect and react to bottleneck displacements; for instance, due to changes in the storage technology, storage policies, multimodal transportations costs and so forth.

Even though the shift in our production and management systems was initiated after World War II, with lean manufacturing taking over the mass production paradigm, the systemic approach has spread in a very irregular way. Agent-based modeling and simulation is an important tool to educate people, and to contribute to create critical mass for a large deployment of the systemic approach, which in the end translates in a better skilled population to deal with complex systems like supply chains.

References

1. Andel, T. (1996). Manage inventory, own information. Transportation & Distribution, 37(5), 54-58.

2. Buchmeister, B., Friscic, D., Lalic, B. & Palcic, I. (2012) Analysis of a Three-Stage Supply Chain with Level Constraints. International Journal of Simulation Modelling, 11(4), 196-210.

3. Chatfield, D. C., & Pritchard, A. M. (2013). Returns and the bullwhip effect. Transportation Research Part E: Logistics and Transportation Review, 49(1), 159-175.

4. Chen, F., Drezner, Z., Ryan, J. K., & Simchi-Levi, D. (2001). Quantifying the Bullwhip Effect in a Simple Supply Chain: The Impact of Forecasting, Lead Times, and Information. Management Science, 46(3), 436-443.

5. Chen, L., Lee, H. L. (2012). Bullwhip Effect Measurement and its Implications. Operations Research, 60(4), 771-784.

6. Cole, H., & Jacob, D. (2003). Introduction to TOC supply Chain. AGI Institute.

7. Cox, J. F., & Spencer, M. S. (1998). The Constraints Management Handbook, Lucie Press, Boca Raton, FL.

8. Dardenne, A., Lamsweerde, A., & Fichas, S. (1993). Goal-directed requirements acquisition. Science of Computer Programming, 20, 3-50.

9. De la Fuente, D., & Lozano, J. (2007). Application of distributed intelligence to reduce the bullwhip effect. International Journal of Production Research, 44(8), 1815-1833.

10. DesMarteu, K. (1998). New VICS publication provides step-by-step guide to CPFR. Bobbin, 40(3), 10.

11. Disney, S. M., & Towill, D. R. (2003). On the bullwhip and inventory variance produced by an ordering policy. Omega – The International Journal on Management Science, 31, 157-167.

12. Disney, S. M., Farasyn, I., Lambrecht, M., Towill, D. R., & Van de Velde, W. (2005). Taming the Bullwhip Effect whilst watching customer service in a single supply chain echelon. European Journal of Operational Research, 173(1), 151-172.

13. Forrester, J. W. (1961). Industrial dynamics, MIT Press. Cambridge, MA.

14. Fox, M. S., Chionglo, J. F., & Barbuceanu, M. (1993). The Integrated Supply Chain Management System. Internal Report, Dept. of Industrial Engineering, University of Toronto.

15. Goldratt, E. M. (1990). Theory of Constraints. North River Press, Croton-on-Hudson, NY.

16. Goldratt, E. M. (1994). It's not Luck. North River Press, Great Barrington, MA.

17. Goldratt, E. M. (1997). Critical Chain. North River Press, Great Barrington, MA.

18. Goldratt, E. M., & Cox, J. (1984). The goal – A process of ongoing improvement. North River Press. Croton-on-Hudson, NY.

19. Goldratt, E. M., Schragenheim, E., & Ptak, C. A. (2000). Necessary but not sufficient. North River Press. Croton-on-Hudson, NY.

20. Kaminsky, P., & Simchi-Levi, D. (1998). A new computerized beer game: a tool for teaching the value of integrated supply chain management. Supply Chain and Technology Management. Hau Lee and Shu Ming Ng Edts. The Production and Operations Management Society. Miami, Florida.

21. Kimbrough, S. O., Wu, D. J., & Zhong, F. (2002). Computers play the beer game: can artificial manage supply chains? Decision Support Sytems. 33, 323-333.

22. Lee, H. L., Padmanabhan, V., & Whang, S. (1997). The Bullwhip Effect in supply chains. MIT Sloan Management Review, 38(3), 93-102.

23. Liang, W. Y., & Huang, C. C. (2006). Agent-based demand forecast in multi-echelon supply chain. Decision Support Systems, 42(1), 390-407.

24. Mangina, E., & Vlachos, I. P. (2005). The changing role of information technology in food and beverage logistics management: Beverage network optimization using intelligent agent technology. Journal of Food Engineering, 70(3), 403-420.

25. Maturana, F., Shen, W., & Norrie, D. H. (1999). MetaMorph: An adaptive agent-based architecture for intelligent manufacturing. International Journal of Production Research, 37(10), 2159-2173.

26. McKinsey, J. (1992). Evaluating the impact of alternative store formats. Supermarket Industry Convention. Food Marketing Institute, Chicago.

27. Saberi, S., Nookabadi, A. S., & Hejazi, S. R. (2012). Applying Agent-Based System and Negotiation Mechanism in Improvement of Inventory Management and Customer Order Fulfillment in Multi Echelon Supply Chain. Arabian Journal for Science and Engineering, 37(3), 851-861.

28. Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A., & White, D. R. (2009). Economic Networks: The New Challenges. Science, 325 (5939), 422-425.

29. Simatupang, T. M., Wright, A. C., & Sridharan, R. (2004). Applying the Theory of Constraints to supply chain colaboration. Supply Chain Management: An International Journal, 9(1), 57-70.

30. Simatupang, T. M., Hurley, S. F., & Evans A. N. (1997). Revitalizing TQM efforts: a self-reflection diagnosis based on the Theory of Constraints. Management Decision, 35(10), 746-752.

31. Sterman, J. D. (1989). Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. Management Science, 35(3), 321-339.

32. Taguchi, G., Chowdhury, S., & Taguchi, S. (2000). Robust Engineering. Mc Graw – Hill, New York.

33. Umble, M., Umble, E., & von Deylen, L. (2001). Integrating enterprise resources planning and Theory of Constraints: a case study. Production and Inventory Management Journal, 42(2), 43-48.

34. Wilensky, U. (1999). NetLogo. Northwestern University, Evanston, IL: The Center for Connected Learning and Computer - Based Modeling. http://ccl.northwestern.edu/netlogo/ Last access 10 April 2014.

35. Wooldridge, M. (2000). Reasoning about Rational Agents. MIT Press, Cambridge, Mass.

36. Wu, H. H., Chen, C. P., Tsai, C. H., & Tsai, T. P. (2010). A study of an enhanced simulation model for TOC supply chain replenishment system under capacity constraint. Expert Systems with Applications, 37, 6435-6440.

37. Wu, H. H., Lee, A. H. I., Tsai, T. P. (2014). A two-level replenishment frequency model for TOC supply chain replenishment systems under capacity constraint. Computers & Industrial Engineering, 72, 152-159.

38. Youngman, K. (2009). A Guide to Implementing the Theory of Constraints (TOC). http://www.dbrmfg.co.nz/ Last access 9 July 2014.

39. Yuan, K. J., Chang, S. H., & Li, R. K. (2003). Enhancement of theory of constraints replenishment using a novel generic buffer management procedure. International Journal of Production Research, 41(4), 725-740.

40. Zarandi, M. H., Fazel, Pourakbar, M., & Turksen, I. B. (2008). A fuzzy agent-based model for reduction of bullwhip effect in supply chain systems. Expert systems with applications, 34(3), 1680-1691.

CHAPTER 6

HOLISM VERSUS REDUCTIONISM IN SUPPLY CHAIN MANAGEMENT: AN ECONOMIC ANALYSIS♦

♦ This article was published in *Decision Support Systems* (Elsevier), vol. 86, no. 1, pp. 83-94 in June 2016. DOI: 10.1016/j.dss.2016.03.010. Authors (in order): Borja Ponte, José Costas, Julio Puche, David de la Fuente, Raúl Pino.

Decision Support Systems, indexed in the ISI, is a Q1-journal in the category 'Operations Research and Management Science' (Science Edition). Impact Factor 2015: 2.604.

Abstract

Since supply chains are increasingly built on complex interdependences, concerns to adopt new managerial approaches based on collaboration have surged. Nonetheless, implementing an efficient collaborative solution is a wide process where several obstacles must be faced. This work explores the key role of experimentation as a model-driven decision support system for managers in the convoluted decision-making process required to evolve from a reductionist approach (where the overall strategy is the sum of individual strategies) to a holistic approach (where global optimization is sought through collaboration). We simulate a four-echelon supply chain within a large noise scenario, while a fractional factorial design of experiments (DoE) with eleven factors was used to explore cause-effect relationships. By providing evidence in a wide range of conditions of the superiority of the holistic approach, supply chain participants can be certain to move away from their natural reductionist behavior. Thereupon, practitioners focus on implementing the solution. The theory of constraints (TOC) defines an appropriate framework, where the Drum–Buffer–Rope (DBR) method integrates supply chain processes and synchronizes decisions. In addition, this work provides evidence of the need for aligning incentives in order to eliminate the risk to deviate. Modeling and simulation, especially agent-based techniques, allows practitioners to develop awareness of complex organizational problems. Hence, these prototypes can be interpreted as forceful laboratories for decision making and business transformation.

Keywords

Drum–Buffer–Rope; Model-driven decision support systems; OUT policy; Theory of constraints; Throughput accounting.

1. Introduction

Reductionism and holism represent two opposite philosophical approaches to problem solving. While the former is based on the “divide and conquer” paradigm (breaking down the problem into simpler and smaller parts), the latter underscores the idea that systems must be viewed as a whole and not as collection of parts. In this sense, when reductionism is applied to Supply Chain Management (SCM), the overall strategy is obtained as a sum of the individual strategies of the companies that conform the supply chain (*i.e.* local optimization). On the contrary, in a holistic context, these individual strategies are the result of an overall strategy defined by collaboration (*i.e.* global optimization). Therefore, supply chain members must tackle a dilemma [1]: deciding between favoring decisions which go in their own interest and accommodating those that consider the interest of the system as a whole.

Given that supply chains are growingly built on interrelationships, practitioners widely accept that holistic approaches play a crucial role in improving overall performance. Nevertheless, even though it might seem counterintuitive, reductionism is still widespread in real systems [2]. This approach results in a Nash equilibrium, which brings lower overall performance [3]. If each member ignores the impact of its actions on the other echelons, the maximization of individual metrics often occurs at the expense of the entire supply chain performance [4]. In this sense, local optimization has shown to be a major source of inefficiencies, such as the well-known Bullwhip Effect [5], that define a set of common issues faced by real supply chains –e.g. excessive inventories, low customer service level, and high production variability.

Under these circumstances, collaboration stands out as a key source of competitive advantages. This research work explores the key role of experimentation through Agent-Based Modeling (ABM) [6] as a powerful model-driven decision support system [7] for managers in the complex decision-making process of adopting a collaborative solution within the supply chain. These computer-based prototypes can be interpreted as forceful and risk-free laboratories for business transformation.

Firstly, this article aims to provide evidence of the fact that holistic approaches clearly outperform reductionist ones from an economic perspective. In addition, it uses experimentation techniques to define the sources where the upgrade is based on. It is only by understanding the improvement that supply chain actors can be certain when moving away from their natural reductionist behavior.

Once developed the awareness of practitioners through the economic comparison, they focus on adopting holistic solutions. The implementation of efficient solutions is a wide process that requires an integrative schema, where information sharing must be understood as an enabler. In this sense, experimentation aims to motivate the confidence between supply chain echelons. It is relevant to underline that the efficiency of the collaborative solution depends not only on the technical component but also on the acceptance level of the decision makers.

From that point on, adopting holistic approaches requires to integrate processes, synchronize decisions, and set up systemic performance indicators. This paper shows how to use the Theory of Constraints (TOC) [8] with that goal through modeling techniques, which has been compared to the inventory cost-optimal Order-Up-To (OUT) policy [9]. That is, TOC allows participants to define a systemic methodology to tackle the previously identified issues. This production paradigm manages the supply chain flows through the Drum-Buffer-Rope (DBR) method with a focus on the bottleneck and defines a systemic scorecard through the Throughput Accounting (TA).

However, this is not enough. If there are huge differences in how the echelons benefit from collaboration, the holistic approach would not be viable. That is, collaboration must be aimed at achieving the optimal through a Nash equilibrium where the incentives to deviate are eliminated. For this reason, we last focus on the concept of “incentive alignment” [10]. The profit increase must reward the contribution of each node in order to avoid opportunistic behaviors. Agent-based prototypes can lead the supply chain to modify its costs structure with this aim.

By way of summary, *Figure 1* describes the decision map of this research and shows how experimentation conducts, in each step, the adoption of a holistic management. This graph, assuming the reductionist approach is the baseline, underscores the need for five connecting features in order to implement efficient collaborative solutions [10].

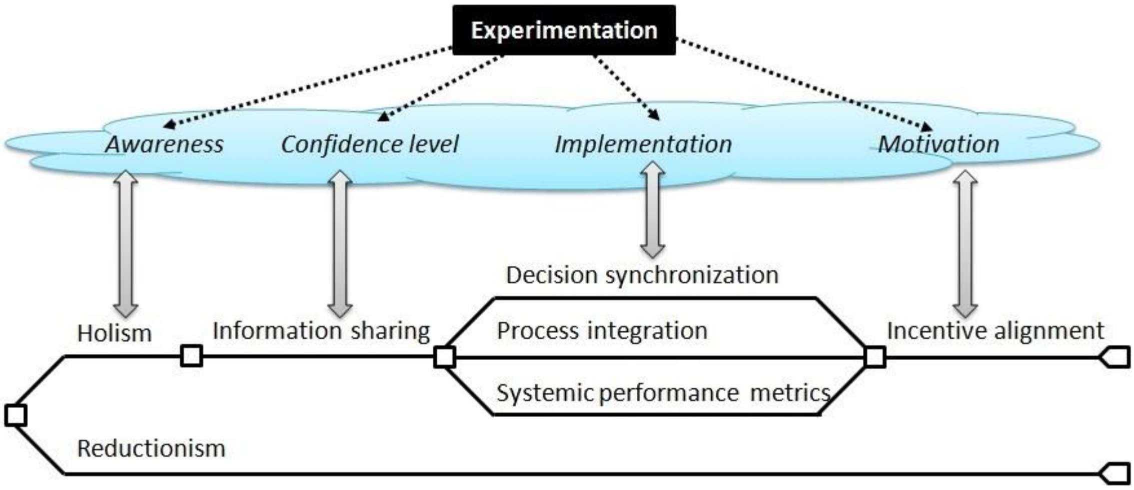


Figure 1. Decision map of this research, which highlights the role of experimentation.

2. Background: Literature Review

Some key points of the research context are described below. We first present the main dilemma faced by supply chain members. Next, we introduce an integrative framework to take advantage of the collaborative approach. Lastly, the TOC is detailed as a systemic method to combat the issues derived from the reductionist approach.

2.1. The dilemma of supply chain practitioners

The main goal of companies is to make money now and in future [8]. To accomplish this objective, they can adopt two positions. This is the so-called dilemma [1]. This inherent decision for supply chain members can be expressed by *Figure 2*.

The traditional approach to face this dilemma (to maximize individual performance) consists in seeking for protecting their individual profitability [4]. This reductionist behavior provokes win-lose games, in which each member looks for its own bargain at the expense of its partners [11]. Local optimization results in multiple forecasting, price fluctuations, and rationing games that (strengthened by order batching and lead times) translate into information distortion along the supply chain. This causes dramatic inefficiencies within the system, through the well-known Bullwhip Effect [12], *e.g.* low service levels and excessive fluctuations in inventories and orders.

From a holistic perspective, the various supply chain echelons understand that the best solution for the whole system leads to the best solution for them. Therefore, in order to maximize profit, they make decisions considering the global profitability, namely they use systemic performance metrics [1]. In a collaborative way, supply chain participants coordinate their processes and synchronize their decisions aimed at revenues from final customers instead of their own sales [4].

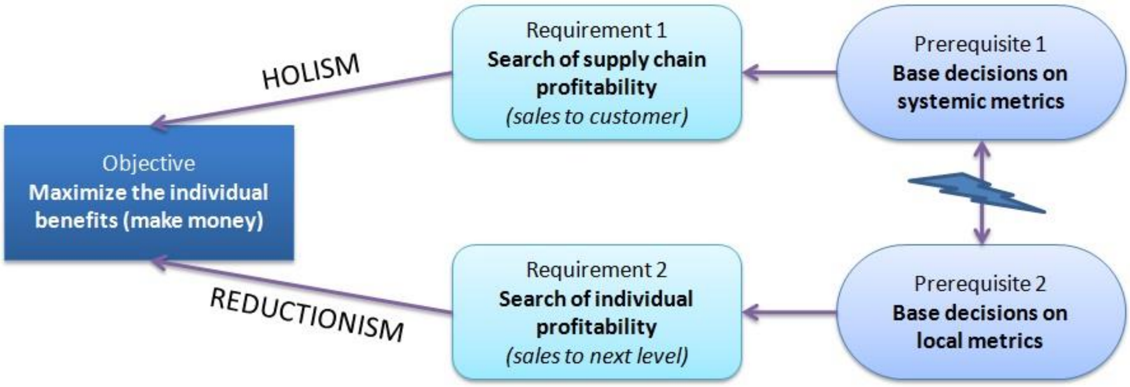


Figure 2. The dilemma of supply chain nodes through an evaporating cloud (adapted from [1]).

2.2. An integrative framework for supply chain collaboration

The literature on the subject mostly assumes holistic methods to outperform traditional management policies (based on reductionist principles) in terms of overall supply chain performance. However, practitioners find it difficult to address the issue of supply chain collaboration, which justifies why, even shown it superiority, it is not widespread in practice [13]. If a robust solution is not found, the menace of opportunistic behaviors arises, which creates an environment of uncertainty and complexity such that the cost of transacting under this context involves additional risk and expense [14]. Namely, the effectiveness of collaboration relies not only upon the integration of operations, but also upon the level to which efforts are aligned [13].

This fact highlights the relevance of defining an appropriate framework for supply chain collaboration where to obtain competitive advantages by working together [4]. This must be integrative, *i.e.* connecting different features of collaboration, such as the one proposed by Simatupang and Sridharan [10], which considers five edges:

i. Information sharing, defined as the access to private data in all members' systems creating visibility at the different nodes on the overall system state.

ii. Decision synchronization, which refers to the extent to which the various echelons can orchestrate critical decisions at planning and execution levels.

iii. Incentive alignment, achieved through the process of sharing costs, risks, and benefits among the various participants in an equitable manner.

iv. Integrated processes, *i.e.* the design of the efficient supply chain flow that delivers products to end customers in a timely manner at lower costs.
v. Systemic performance indicators, understood as the process of devising and implementing metrics that guide members to improve overall performance.

2.3. The Theory of Constraints (TOC) in Supply Chain Management (SCM)

The TOC, presented by E. M. Goldratt [8], meant a major innovation in the production field. This management philosophy views any system as being limited in reaching a higher performance only by its bottleneck. Although it was first oriented on the manufacturing system, further development incorporates solutions for other business areas, such as SCM [15]. The TOC holistic paradigm has been shown to achieve breakthrough improvements in comparison with mass production alternatives in terms of lead time reduction, customer service level increase, and throughput growth [16].

For this reason, the TOC can be presented as the core of the collaborative solution, namely, it allows managers to integrate processes, synchronize decisions and set up a performance system [1]. Its logical thinking is expressed as a continuous improvement cycle with five steps [8]:

i. To identify the bottleneck.
ii. To decide how to exploit the bottleneck.
iii. To subordinate everything else in the system to the previous step.
iv. To elevate the bottleneck.
v. To evaluate if the bottleneck has been broken, and return to the beginning.

To manage the system, the TOC proposes the DBR methodology [8]. This pull-oriented strategy (*i.e.* replenishment is based on actual demand instead of on the forecast) aims to manage properly the bottleneck (ensuring its steady supply) through suitable coordination. It is named by its three main components. The drum, placed at the bottleneck, is a system pacemaker. The rest of the nodes follow its beat (production rate). The buffer protects the drum against variability, so that the full capacity in the bottleneck is exploited. The rope is the release mechanism that subordinates the entire system to the drum. The DBR configuration (planning state) is complemented with the buffer management (monitoring stage), which implies administrating the buffer along the different nodes in order to guide how the system is tuned for peak performance. Applying this method along the supply chain warrants the concentration of all members to what matters for the system as a whole [2].

In addition, the TOC defines a collaborative performance system to measure how the company performs. Unlike cost accounting (aimed at cost reduction), the TA [17] seeks to maximize the efficiency of the flow of value. Hence inventory is not considered as an asset, but a liability. The TA aims to enable managers to examine the link between process constraints and financial performance in decision making, *i.e.* to determine the real impact of their decisions. Three financial measures are proposed as complementary indicators: net profit (absolute terms), return on investment (relative terms), and cash flow (survival terms). The operational decisions are related to overall system success through the "cost bridge", defined by three metrics [17]:

i. Throughput: the rate at which the system generates money through sales, *i.e.* the difference between the revenue and the total variable costs.
ii. Inventory: it includes not only raw material, work-in-progress, and finished goods stock but also all other invested money in the supply chain.
iii. Operating Expense: all the money the system spends in order to turn inventory into throughput, *e.g.* transformation and shipping costs.

3. Problem Formulation: Supply Chain Model

This section is devoted to detail the conceptual model of the agent-based supply chain that has been developed in this research, as well as the wide context where to confirm economic robustness of the results.

3.1. Supply chain scenario: Assumptions and scope

In the same line as other relevant and recent studies [18], the supply chain has been analyzed under the Beer Game environment [19]. This is a traditional single-product supply chain with a serial structure formed by four echelons (factory, distributor, wholesaler, and retailer). With the aim of bringing it closer to reality, the noise sources have been expanded in order to consider common hurdles in real supply chains. It can be called the noise-extended Beer Game environment. The assumptions are as follows:

i. Stochastic customer demand. Specifically, a normal distribution simulates demand. Both mean and standard deviation are selected by the experimenter.

ii. Stochastic lead time. Each node receives both product and orders within a time range (set by the user) after sent, defined by a continuous uniform distribution.

iii. Stochastic failure of products. In each product's action (including storage and transport) along the supply chain, there is a probability of failure, defined by the defective products rate, which is set by the experimenter.

iv. Constrained production (factory) and transportation (between the various echelons) capacity. The user defines these quantity limitations. However, unconstrained storage capacity has been considered.

v. Non-negative condition of the order quantity. Each member cannot return the product to its supplier.

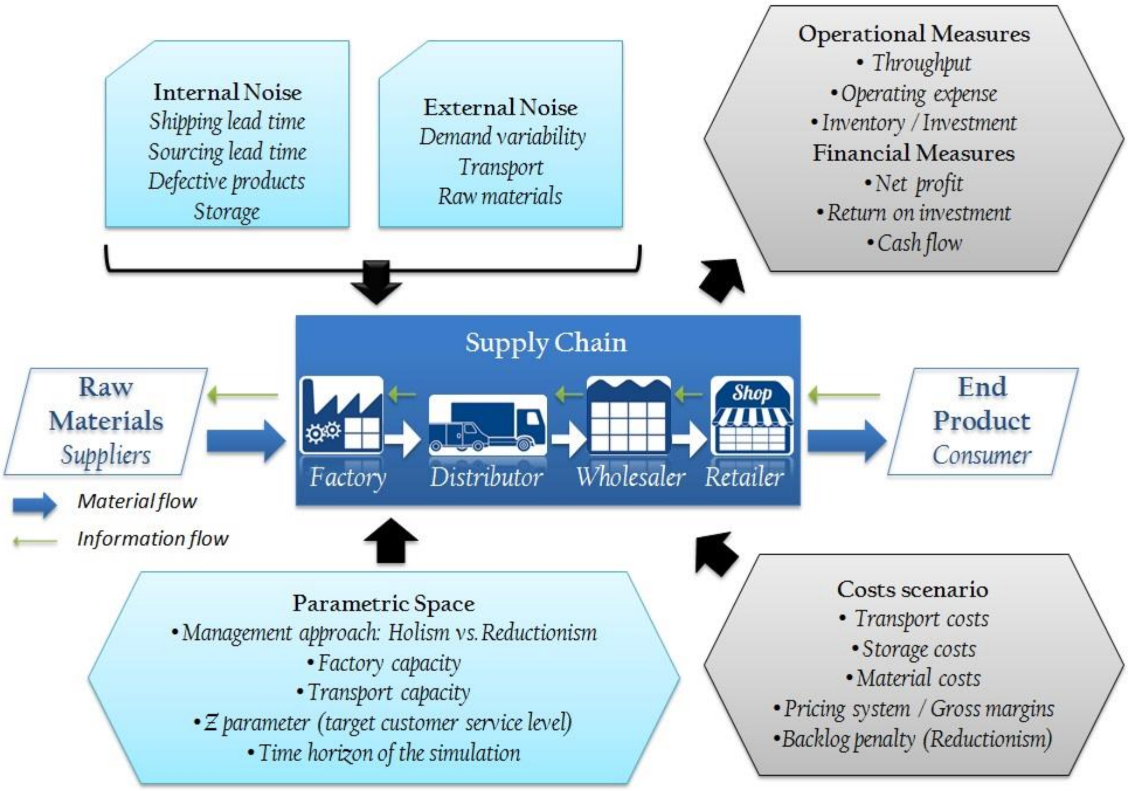


Figure 3. Scope of the research, by means of a parameter diagram.

Figure 3 displays the parameter diagram that describes the scope of this study. In the center, it shows the overall system function responsible for transforming raw materials into finished products. Among the nodes, the material flow (from the factory to the retailer) refers to the shipping of the product, while the information flow (in the opposite direction) represents the orders. At the top part, the noise sources (uncontrollable factors) that threaten the supply chain can be seen divided into internal (lead times, defective products, and storage) and external (demand variability, transport, and raw materials). At the bottom part, the parametric space (controllable factors) highlights the factors to be modified. The extensive costs scenario and the performance (both financial and operational) metrics explained below are also shown.

3.2. Economic model and performance metrics

The economic model seeks to imitate the main revenues and costs faced by real supply chain members. Income in the overall system is only generated through selling the product (to customer). In each node, money is made by sales to the next echelon. Expenditure is incurred in three ways: storage, transport, and provisioning. All of them have been considered to be proportional. Obviously, the provisioning cost of each echelon means the income for the previous one. Hence, a pricing system must be defined in the supply chain. All these economic parameters are set by the user.

As explained, the TA is supported on three operational measures. According to TOC principles, the throughput is the difference between the revenue through sales (selling price times sales quantity) and the variable costs related to purchases (buying price times purchase quantity). Note that the gap between sales and purchase quantities is due to defective products and storage. The operating expense is calculated by adding storage and transport costs, as both are assumed to transform inventory into throughput. In the reductionist approach, this expense is adjusted by the difference between the money paid and received due to backorder penalty (as it is a usual practice in reductionist systems). Finally, the inventory in terms of TA is obtained by estimating the economic value of the products that are stored in each node.

From that point on, the key financial indicators can be easily obtained [17]. The net profit is expressed as the difference between throughput and operating expense, the cash flow considers, besides the above difference, the change of investment in the same horizon, and the return on investment is the net profit divided by the inventory. These metrics can be obtained for the overall supply chain (in the holistic supply chain) or node by node (in the reductionist system).

Figure 4 outlines the economic model of the supply chain, including the three operating metrics and the main financial measure (net profit).

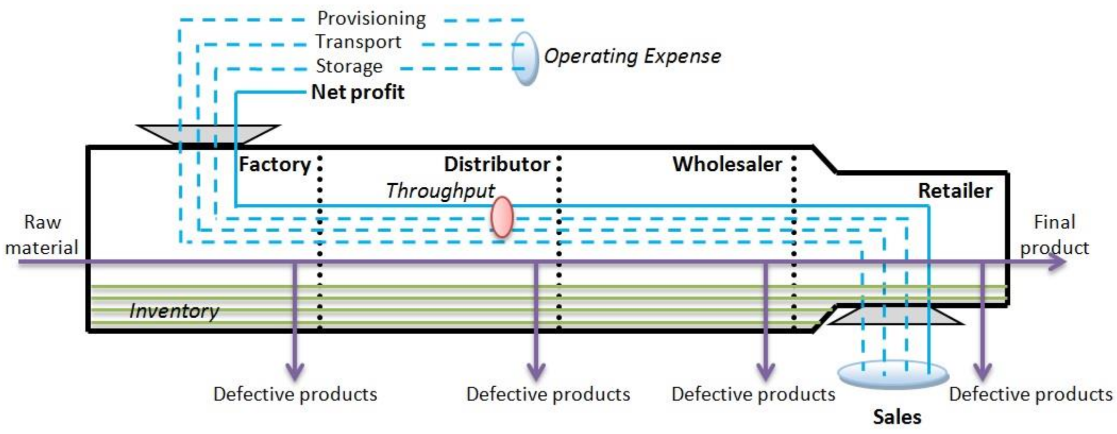


Figure 4. Overview of the economic model of the supply chain, with the operating metrics.

3.3. Reductionist approach: The Order-Up-To (OUT) inventory policy

In the non-collaborative management, each supply chain echelon communicates only with the previous one (to receive an order and to send the product) and with the next one (to place an order and to receive the product). Therefore, customer demand is only known by the retailer. In this mode, orders not fulfilled in time are backlogged, as usual in these replenishment models [9]. In each node, these backorders (involving an economic penalty) are fulfilled as soon as on-hand inventory becomes available.

This approach has been implemented through the OUT method. These policies are often used in the real world, given the usual practice in retailing to replenish very frequently [5] and because it is optimal in terms of inventory and shortage costs [20].

The classic OUT method is a periodic review system for issuing orders depending on demand forecasting and (both on-hand and on-order) inventory position, in order to bring the inventory position up to a defined level. That is, the order rate is the sum of the forecast, the gap between actual and target net stock (on-hand inventory), and the discrepancy between actual and target work-in-progress (on-order inventory). In this research, the demand has been forecast using a three-period moving average.

Both the target work-in-progress and the target net stock (a safety stock) are considered to be variables. Its sum to the forecast defines the OUT point. The former aims to cover the lead time between nodes, so it is easily estimated as the forecast times the average shipping lead time. The latter is aimed to protect supply chain nodes from demand variability and from internal noise sources (variability in lead times and defective in products). It is proportional to the demand standard deviation, to the lead time range, and to the estimated defective products. Each term in multiplied by the parameter Z, set by the user and related to the desired service level.

In the reductionist system, the discrete operation (sequence of events) of each node is summarized by *Figure 5* for the distributor. At the beginning of each period, the product is received from the factory and it is stocked up. Then, the order is received from the wholesaler, and the net stock is checked in order to prepare the shipping. If the order can be fulfilled (besides considering previous backlog), the required quantity is sent; otherwise, backlog is generated and the available product is shipped. The next steps depend on the previously explained OUT policy. Notice both flows are delayed due to the lead time. The operation is similar for the other nodes.

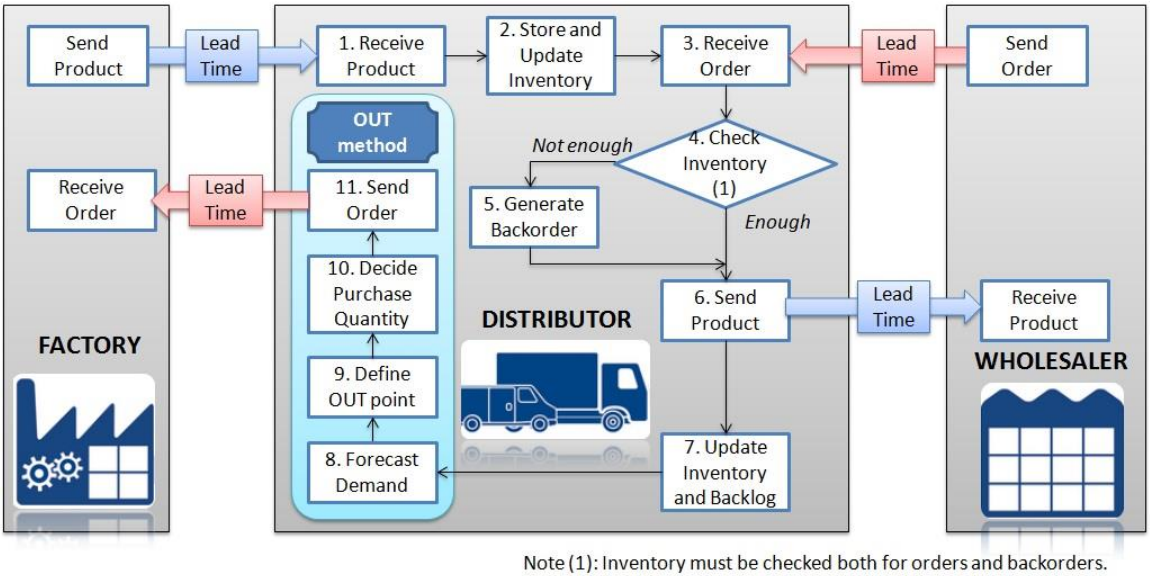


Figure 5. Sequence of events for the distributor when applying the OUT policy.

3.4. Holistic approach: The Drum-Buffer-Rope (DBR) methodology

Under the collaborative approach, the system is ruled by a kind of headquarters that accounts for the interest of the whole supply chain, taking decisions on the basis of greater visibility (supported by information sharing). Accordingly, the various nodes behave as required to protect the overall supply chain function.

From that point on, the holistic approach has been implemented through TOC principles, which are based on prioritizing the system bottleneck. In particular, the DBR method leads to both synchronize the entire sequence of integrated activities required to deliver products and to create effective processes aimed at achieving breakthrough improvements in system performance and reliability [21].

According to TOC logical thinking, the first step is to identify the bottleneck. In supply chains, the bottleneck tends to be the sales constraint as (production, transport, and storage) capacities are usually higher than demand [22]. Under this scenario, the demand is an external constraint beyond the supply chain sphere of influence. Hence, the bottleneck cannot be elevated or broken, and consequently the fourth and fifth steps of the improvement cycle are not required.

That is, the key points are the second and third steps, which define the sequence of events through the DBR method. On the one hand, the supply chain must efficiently exploit the bottleneck. This means to sell the product at the retailer, *i.e.* to minimize lost sales. To this end, the *drum* is placed at the retailer. This must beat out (define) the production and distribution rate for the whole system according to the actual demand. On the other hand, the other nodes must be subordinated to the bottleneck. In this sense, the retailer is protected from shortages, and thus the supply chain as a whole, against variation.

To subordinate the factory, the distributor, and the wholesaler to the bottleneck, we need the buffer and the rope [23]. The *buffer* is aimed at protecting the bottleneck through time. Uncertainty in the supply chain (demand, lead times, and defective products) must not increase lost sales. Thus, the buffer refers to the time period between releasing the material and the drum due date. For each node, the buffer considers the maximum lead time between itself and the customer. The *rope* is the release timing. It can be understood as a real-time feedback between the drum and the node operation. It should be noted that the rope length covers the same as the buffer duration. Tying the rope ensures that excess flow cannot be admitted. In this sense, the rope defines how much to order: the difference between the desired (considering the drum rate along the buffer time, and a safety stock to protect against demand variability and defective products) and the actual (sum of the net stock and shipping product) supply chain inventory. As the overall inventory is constrained, the Bullwhip Effect is dramatically reduced.

This DBR configuration, where the systemic condition to tie the different members through time (not by product) is established, is the planning stage. It is aimed to operate the system. Subsequently, a second stage is required each time period [2]. In the control stage (aimed at keeping a running check of the system efficiency), the buffer is managed along the intermediate members. Buffer management consists in moving the flow so that arrival happens on time at the bottleneck.

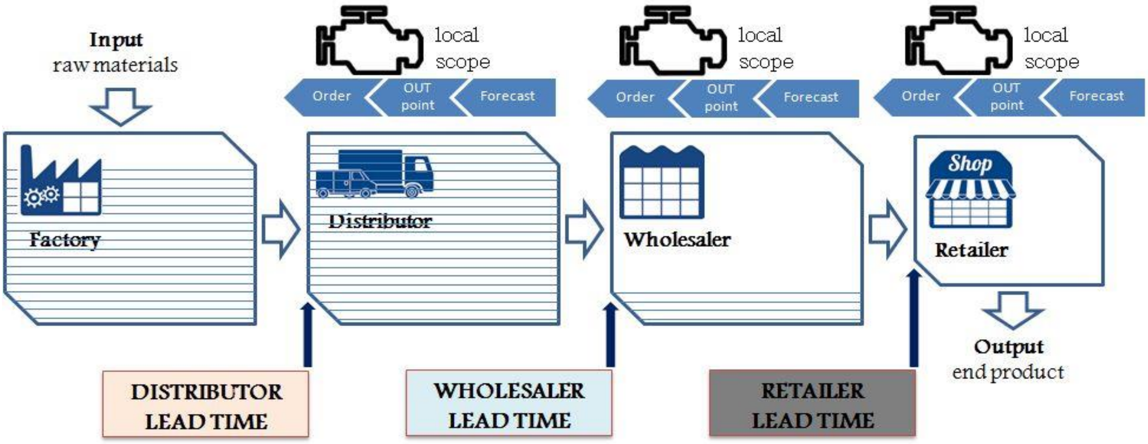


Figure 6a. Overview of the supply chain, when working according to the OUT policy.

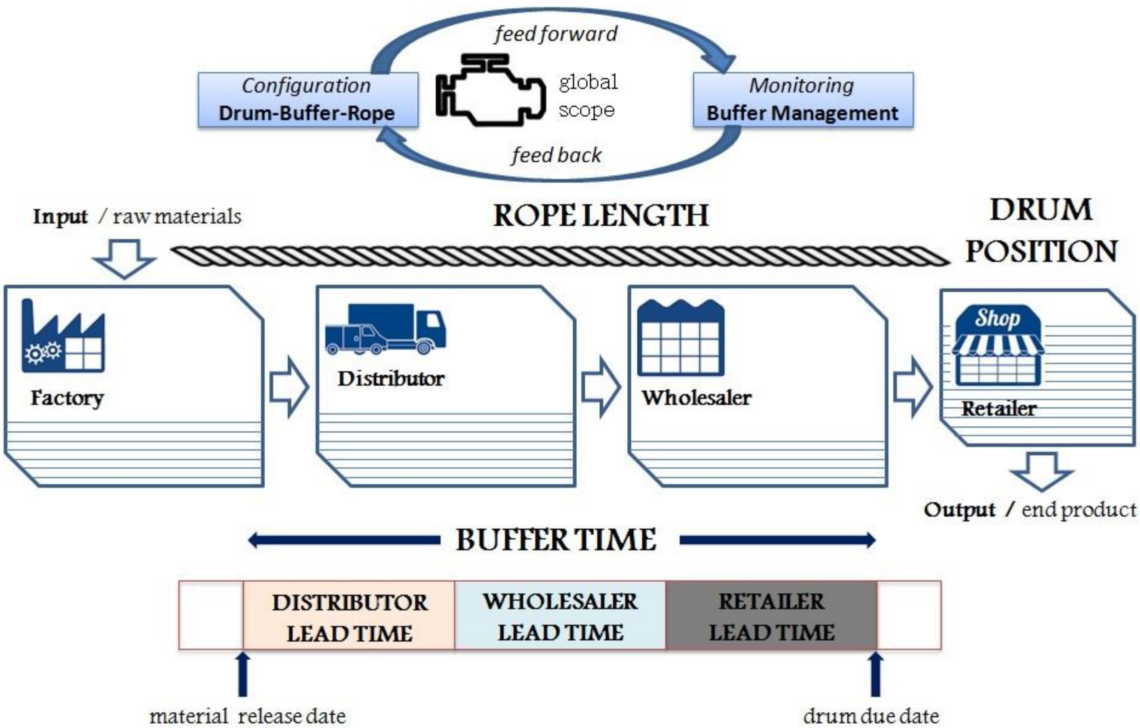


Figure 6b. Overview of the supply chain, when working according to the DBR methodology.

The main ideas explained above are displayed in *Figure 6b* (it highlights the role of the three main components of the DBR method) in contrast to *Figure 6a*, which shows the basic ideas of the OUT policy. It should be noted that the factory decides the production orders that are placed based on the recent demand, while the rest of the nodes compensate the flow dissipated downstream after shipping. They calculate the rope length to the drum position, and make the order decision based on its downstream buffer to the bottleneck. This way, each supply chain member decides the quantity to dose subordinated to the bottleneck, so these dissipative orders do not have lead time nor generate backorders since the next dosage again obey the bottleneck [24]. On the contrary, note that in the reductionist system the nodes consider the upcoming demand and local inventories to order. Youngman [22] has developed an outstanding guide for TOC implementation in production and distribution systems, which can be consulted to get further detail.

4. Agent-Based Development of the System

In order to carry out the experimental approach aimed at comparing a supply chain managed using the OUT policy versus the same system ruled by TOC principles, the noise-extended Beer Game environment was required to be modeled. From the different available alternatives to create this model, we chose ABM [6].

ABM is a decentralized approach to model design emerging analytical method for social sciences, aimed at simulating the actions and interactions of autonomous agents (between them and with the environment) with a view to assessing their effects on the system as a whole [6]. This modeling approach follows the underlying notion that complex systems are built bottom-up. ABM fits in computational science [25] and is a very suitable approach when the problem is intractable by analytical tools, when the theoretical approach might be not reliable, or the experimentation with a real system is unfeasible or costly; all of which apply in our case. Actually, ABM is largely used to analyze the complex behavior of supply chains [26].

As ABM has its roots in Complex Adaptive Systems [27], we have extensively used these mechanisms to build our model: the agents are tagged (can be distinguished), they have internal and polymorphous rules to represent decision making, and the model is created by using building-blocks (aggregating simpler reusable components). The system has been implemented using NetLogo 5.1.0 [28]. NetLogo is a multi-agent programmable modeling environment continuously developed by the Center for Connected Learning and Computer-Based Model (Northwestern University).

We have used different breeds (types) of agents to represent the system, such as actors (supply chain echelons), events (that trigger the action), entities (representing material and orders), records (performance metrics), and police (for controlling and debugging). Each one has its own attributes and methods. Thus, agents are heterogeneous [29].

The engineering of the agent-based model consists in making the agents to follow discrete-event cues and make them behave as finite-state machines. For this reason, a Future Event List (FEL) artifact is a core feature in the model, as it cares about cueing future events to deploy action. In addition, the model is based on a finite-state engine, which makes actors roam through a cyclic map of states to perform the previously defined sequence of events (see *Figures 5 and 6*). At the beginning of each cycle (local for each agent), the agent is idle. At the end, it reports the main results. Therefore, agents are autonomous [30] in terms of their decision making.

Two essential phases in modeling are verification (checking cohesion and consistency) and validation (predictions must match the reality). In this regard, the model was developed following strict rules of clean code, test-driven development, and robust engineering. We used anti-error mechanisms (*e.g.* cross checking) for early detection of system malfunctions. In addition, several acceptance tests have been used to confirm that the model exhibits a known behavior when exposed to controlled conditions.

5. Simulation Study and Discussion of Results

This section presents the Design of Experiments (DoE), shows the results obtained in this research work, and discusses them based on the stated objectives.

5.1. Design of Experiments (DoE)

This DoE aims to assess the impact of moving from the OUT policy to a DBR-managed supply chain in a wide variety of scenarios, both from an overall and a node-by-node perspective. Using Goldratt's principles, results (Y) are expressed in terms of net profit (NP) in the entire supply chain (i=0) and in the four members ($i=1,\dots,4$). This larger-the-better indicator represents the critical concern the various supply chain members.

The experimentation approach shown in *Eq. (1)* is defined as a function of eleven variables. In other words, treatments have been performed on different scenarios defined by the combination of eleven factors; see the parameter diagram in *Figure 3*. Four of them are controllable: management policy (*X1*), production capacity (*X2*), transport capacity (*X3*), and Z safety parameter (*X4*). The remaining seven factors are noise: standard deviation of the demand (*Z5*), transport cost (*Z6*), storage cost (*Z7*), defective product rate (*Z8*), gross margin of the supply chain echelons (*Z9*), range of the order lead time (*Z10*), and range of the product lead time (*Z11*). All of them are real factors except *X1* that is a categorical variable. It should also be considered the unexplained part of the system response, *i.e.* the residuals (ξ). Therefore, it is a fractional factorial DoE with eleven factors.

$$Y = [NP_i]_{i=0}^{i=4} = f(X1, X2, X3, X4, Z5, Z6, Z7, Z8, Z9, Z10, Z11) + \xi \qquad (1)$$

Each factor in the DoE has two levels. *Table 1* outlines the levels that have been defined. We have sought for wide enough ranges in these variables where to derive conclusions with general implications for real supply chains. Note that we have selected as fixed those factors that act like an anchor for the others, see *Table 1*.

We have employed 10 and 30 as values for the standard deviation of the demand, since the coefficient of variation (*i.e.* the ratio of the standard deviation to the mean of the demand, which is 100) of retail series is usually lower than 50% [5]. Regarding the capacities, the lower level introduces a significant constraint in the system (40% greater than the average demand, which reduces the supply chain ability to react when backlog occurs especially when suffering from the Bullwhip Effect), while the higher level creates an unconstrained environment. In terms of the safety stock, the use of 90% (it is considered that lower values provoke a high number of lost sales) and 99% (it is considered that greater values result in excessive storage costs) to define the interval of target customer service levels is common both in research studies and in practice [31].

Although the range of the order lead time is usually considered as null (*i.e.* fixed lead time, level 1 in our DoE), we have also chosen a range of 1 (level 2) to analyze the impact of this variable. At the same time and since the minimum value is 4, the range of the product (shipping) lead time varies between 1 (low variability; lead time between 4 and 5) and 4 (high variability; lead time between 4 and 8). These ranges can be understood as common in practice; e.g. see [32]. In terms of the defective product rate per time unit, we have selected an interval from 500 to 6,000 parts per million, since the model have been designed to usually operate within the industry-average area in the six-sigma scale [33]. From this point on, the model may explore other points either in the best-in-class area or in the non-competitive area of the six-sigma scale.

Regarding the economic factors, the unit material cost ($0.40) sets the economic scale. The unit gross margin per node has been decided to cover an interval from the 50% to the 150% of the material cost, while the ratio of the transport and storage costs per period have been chosen to be between 0.5% and 2.5%. Nonetheless, the economic values are meaningless in their selves, but the relevant point is their financial implications on the supply chain. For example, the Return on Sales (ratio of the operating profit to the sales revenue) in the tests performed varies from -4% to 70%, which can be assume to largely cover the usual financial situation of real systems.

Table 1. DoE: Definition of the factors and levels.

Factor	*Role*	*Level 1 (Low*)*	*Level 2 (High*)*
Management policy (X1)	Controllable	Holism - DBR	Reduct. - OUT
Production capacity (X2)	Controllable	140 u	9876 u
Transport capacity (X3)	Controllable	140 u	9876 u
Z safety parameter (X4)	Controllable	1.282 (90%)	2.326 (99%)
St. Dev. of the demand (Z5)	External Noise	10 u	30 u
Transport cost (Z6)	External Noise	0.002 $/u/period	0.01 $/u/period
Storage cost (Z7)	Internal Noise	0.002 $/u/period	0.01 $/u/period
Defective products rate (Z8)	Internal Noise	500 ppm	6000 ppm
Gross margin (Z9)	Internal Noise	0.20 $/u	0.60 $/u
Order lead time: Range (Z10)	Internal Noise	0 periods	1 period
Product lead time: Range (Z11)	Internal Noise	1 period	4 period
Mean of the demand (Fixed)	External Noise	100 u	
Material cost (Fixed)	External Noise	0.40 $/u	
Order lead time: Min (Fixed)	Internal Noise	1 period	
Product lead time: Min (Fixed)	Internal Noise	4 periods**	
Backorder penalty (Fixed)	Internal Noise	0.04 $/u/period	

Note: (*): "Low" and "high" refers only to the categorical variables; (**): Except in the factory, where the product lead time is 10; (1): The fractional factorial DoE requires the use of mid levels in numerical factors. According to the standard logic, we have chosen the following values: X2 - 180 u; X3 - 180 u; X4 - 1.64 (95%); Z5 - 22 u; Z6 - 0.006 $/u/period; Z7 - 0.006 $/u/period; Z8 - 2000 ppm; Z9 - 0.60 $/u; Z10 - 0 periods; Z11 - 2 periods.

5.2. Layout and results

Following Fisher's strategy [34], an 18-row orthogonal inner array (block 1) has been created. Each row represents a treatment defined by a different combination of factors. This technique allows one to draw conclusions from a broad design space exploring some strategic points. In addition, 6 additional runs have been carried out (block 2). These are the same intermediate treatment replicated three times for both management approaches, with the aim of checking consistency of results and system stability. The former was verified through a 2-variance Levene test, which showed that differences are not significant. Regarding the latter, we verified there is not lack-of-fit problem. A time horizon of 250 periods was used for each treatment. *Table 2a* displays the layout and *Table 2b* shows the results. These tables highlight the collaborative treatments.

Table 2a. DoE: Inner array (orthogonal matrix).

Run	*X1*	*X2*	*X3*	*X4*	*Z5*	*Z6*	*Z7*	*Z8*	*Z9*	*Z10*	*Z11*
1	OUT	Low	Mid	Low	Low	High	High	Low	High	Low	Low
2	DBR	High	High	Low	Low	High	Low	High	High	Low	Low
3	DBR	Low	High	High	Low	High	High	Low	Low	High	High
4	DBR	High	Low	Low	Low	Low	High	Low	Low	Low	High
5	OUT	High	Low	High	Low	Low	Low	High	High	High	High
6	OUT	Low	Low	High	Low	High	Low	High	Mid	Low	Low
7	DBR	Low	High	High	High	Low	High	Low	High	High	Low
8	DBR	Low	Low	Low	Mid	Low	High	High	High	High	Low
9	OUT	High	High	High	High	High	High	High	High	Low	High
10	DBR	Low	Low	High	High	High	Mid	High	Low	Low	High
11	OUT	Low	Low	Low	High	High	Low	Low	Low	High	Low
12	OUT	High	High	Low	High	High	High	High	Low	High	Low
13	DBR	Low	High	Low	High	Low	Low	High	Low	Low	Low
14	OUT	High	Low	High	High	Low	High	Low	Low	Low	Low
15	OUT	Low	High	Low	High	Low	Low	Low	High	Low	High
16	DBR	High	Low	Low	High	High	Low	Low	High	High	High
17	OUT	Low	High	Low	Low	Low	High	High	Low	High	High
18	DBR	High	High	High	Low	Low	Low	Low	Low	High	Low
19	OUT	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid
20	DBR	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid
21	OUT	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid
22	DBR	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid
23	OUT	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid
24	DBR	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid	Mid

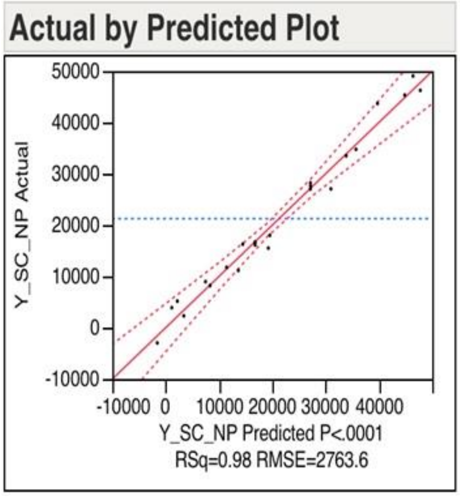
Table 2b. DoE: Results of the different treatments.

Run	*X1*	*NP SC*	*NP Fact*	*NP Dist*	*NP Whol*	*NP Ret*
1	OUT	$27,044.56	$5,718.33	$4,945.86	$7,264.44	$9,115.93
2	DBR	$43,686.43	$16,154.12	$8,339.36	$10,318.59	$8,874.36
3	DBR	$8,205.18	$3,536.14	$947.53	$2,390.29	$1,331.22
4	DBR	$11,757.21	$3,002.22	$2,859.05	$2,887.18	$3,008.76
5	OUT	$34,762.44	$8,875.50	$7,079.23	$9,812.19	$8,995.52
6	OUT	$16,271.97	$4,440.30	$1,683.10	$4,602.04	$5,546.53
7	DBR	$49,044.95	$13,823.97	$11,912.50	$11,968.64	$11,339.84
8	DBR	$46,234.37	$15,300.15	$11,221.03	$10,699.88	$9,013.31
9	OUT	$18,000.34	$287.99	$2,652.27	$7,475.88	$7,584.20
10	DBR	$2,299.29	$4,418.99	$-1,799.86	$1,037.15	$-1,356.99
11	OUT	$8,971.63	$1,816.87	$101.50	$3,565.77	$3,487.49
12	OUT	$-2,972.76	$-4,309.84	$-2,871.24	$1,921.81	$2,286.51
13	DBR	$11,144.96	$5,062.66	$1,577.08	$2,260.19	$2,245.03
14	OUT	$3,896.98	$-4,974.51	$1,606.11	$3,587.80	$3,677.58
15	OUT	$33,491.43	$7,596.18	$6,918.58	$9,335.48	$9,641.19
16	DBR	$45,336.74	$14,061.00	$9,395.73	$11,464.13	$10,415.88
17	OUT	$5,176.96	$-331.77	$-23.12	$2,230.43	$3,301.42
18	DBR	$15,503.55	$4,411.51	$3,628.99	$3,857.10	$3,605.95
19	OUT	$16,697.15	$2,800.37	$2,890.40	$5,280.46	$5,725.92
20	DBR	$28,205.50	$8,521.28	$6,279.26	$7,047.98	$6,357.07
21	OUT	$16,547.36	$2,215.26	$3,078.03	$5,515.68	$5,738.39
22	DBR	$27,655.38	$8,431.51	$5,798.61	$7,047.75	$6,377.51
23	OUT	$16,094.82	$1,977.17	$3,007.92	$5,249.42	$5,860.31
24	DBR	$27,032.78	$8,601.08	$5,571.46	$6,613.59	$6,246.65

Note: "NP SC" represents the overall net profit of the supply chain, while the last four columns show the net profit of the four supply chain members. A more detailed version of the results is available upon request.

5.3. Overall analysis of the results

First, we focus on the results of the entire supply chain. Broadly speaking, this experimentation provides evidence about the sound impact of DBR application to improve supply chain profitability in comparison with the OUT inventory policy. While the average net profit is $26,342.20 when DBR manages the supply chain, it is $16,165.24 when the OUT inventory policy is applied in each participant. This means an improvement of 63%. Nonetheless these impressions must be verified statistically. Due to this reason, Yates' algorithm was applied to compute the estimates of main effects in this factorial experiment. JMP [35] statistical software has been used.



Summary of Fit

RSquare	0.982118
RSquare Adj	0.965727
Root Mean Square Error	2763.646
Mean of Response	21253.72
Observations (or Sum Wgts)	24

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	11	5033827168	457620652	59.9157
Error	12	91652842.8	7637736.9	**Prob > F**
C. Total	23	5125480010		<.0001*

Figure 7. Summary of the Yates' results that confirm the validity of the linear model obtained.

Figure 7 shows that a linear model is enough to explain the results obtained. As the coefficient of determination (R^2) is considerably high, the variability is more absorbed through the model than by the residual. Hence, there is a large capacity to explain system response between controllable and noise factors. The ANOVA study concludes that the model is relevant (p-value significantly lower than 5%).

Table 3 displays the effect diagram, with the parameter estimates (and the standard error), the t-ratio (which tells about the relative importance of each factor), and the p-value. This shows six significant factors at the confidence level 95%. As expected, the management policy is one of them. This one and the gross margin are the more relevant factors. That is, the main hypothesis is confirmed: the holistic DBR method significantly outperforms the reductionist OUT. On the other hand, although both are relevant, transport cost has shown to be more important than storage cost, while the range of the order lead time has a higher impact than the one of the product lead time.

Figure 8 exhibits the main effects plot. It graphically shows the influence of the various controllable and noise factors on supply chain net profit. It should be remarked that, according to the main goal of the paper, it represents the screening. Hence some of the effects that can be seen are negligible. Note that neither the production nor the transportation capacities have proven to be significant. Surprisingly, the Z safety parameter does not have a considerable impact on the net profit, while this metric does not have a significant relationship with the standard deviation of the demand. In addition, it is not possible to verify, at the confidence level 95%, a great negative effect caused by the defective product rate on the net profit.

Table 3. DoE screening: Effect analysis for the whole supply chain net profit.

Factor	*Estimate*	*Std Error*	*t-ratio*	**p-value**
Intercept	4,864.30	3,535.65	1.38	0.1940
Management policy (X1) (*)	-5,190.72	574.69	-9.03	0.0000
Production capacity (X2)	-0.09	0.13	-0.70	0.4991
Transport capacity (X3)	-0.18	0.13	-1.35	0.2019
Z safety parameter (X4)	-1,911.71	1,273.53	-1.50	0.1592
St. Dev. of the demand (Z5)	-50.23	67.86	-0.74	0.4734
Transport cost (Z6)	-624,120.30	166,365.90	-3.75	0.0028
Storage cost (Z7)	-462,745.70	171,117.30	-2.70	0.0192
Defective products rate (Z8)	-0.50	0.24	-2.08	0.0596
Gross margin (Z9)	77,435.22	3,413.62	22.68	0.0000
Order lead time: Range (Z10)	4,743.19	1,241.29	3.82	0.0024
Product lead time: Range (Z11)	-1,035.85	442.83	-2.34	0.0374

Note: (*): In this categorical variable, the results refer to OUT in comparison with DBR; (1): This table highlights the significant factors at the confidence level 95%.

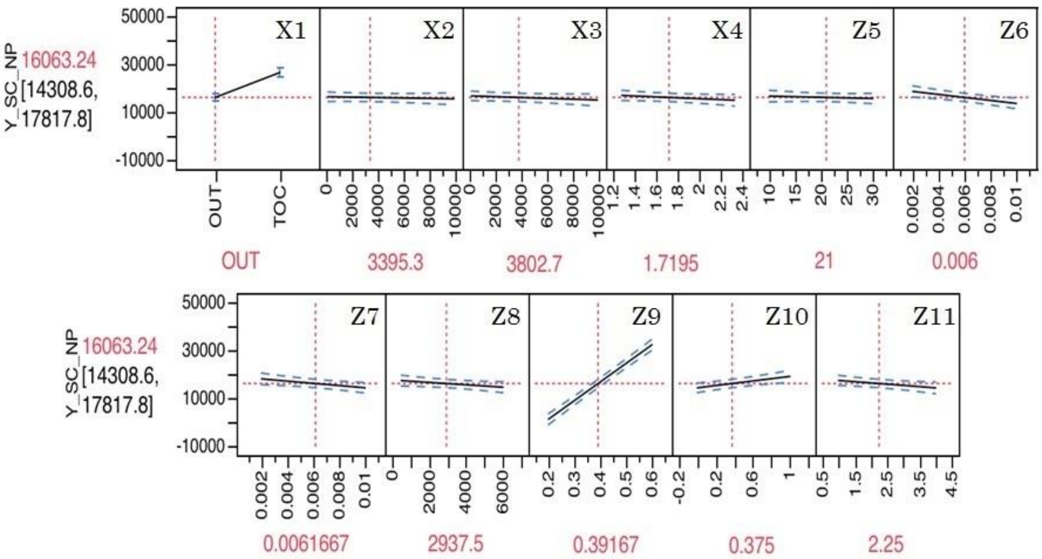


Figure 8. DoE screening: Main effects of the different factors on supply chain performance.

5.4 Understanding the improvement

The previous results demonstrate the improvement induced by the holistic management on the supply chain in economic terms, but how is this achieved? To answer this question, we focus on block 2 of the DoE: the central points. *Table 4* shows the average and the standard deviation of the operational indicators in these tests when the system is managed through the OUT replenishment policy (runs 19, 21, and 23) and the DBR methodology (runs 20, 22, 24).

Table 4. Operational indicators in block 2: mean and standard deviation (in brackets).

Management policy (X1)	*Throughput*	*Operating Expense*	*Inventory*
OUT inventory policy	$24,249.30 ($328.35)	$5,319.94 ($312.79)	$2,830.00 ($428.00)
DBR methodology	$35,193.20 ($187.48)	$4,579.09 ($71.29)	$2,882.53 ($50.97)

These results outline that the net profit grows due to two reasons: the throughput tends to increase, and the operating expense tends to decrease. Nonetheless, the contribution of the throughput has a greater significance. It is not a surprise: the TOC proposes an innovative management focused on increasing the throughput, whereas traditional practices are aimed at cutting costs. However, and paradoxically, the DBR method also leads to a cost reduction. *Table 5* helps to interpret these results, displaying some key indicators that underline the differences between both approaches.

The increase in the throughput comes mainly from the rise in the total sales in the system, *i.e.* the reduction in lost sales achieved by the DBR method. Note that the increase by 25% in total sales translates into a higher increase (45%) in the throughput due to the operating leverage. The large amount of lost sales, even when working with high service levels, within the reductionist system is a direct consequence of the problems caused by the Bullwhip Effect (variability along the supply chain is significantly higher). Notice lost sales increase dramatically even though the inventory level is similar (see the inventory in *Table 4*, or the average time in the system in *Table 5*), since it is not appropriately distributed in the supply chain to protect the bottleneck.

Table 5. Total sales, average time in the system (per unit), rolled throughput yield (RTY, or percentage of defect free units), and Bullwhip Effect (ratio between the variance of the overall inventory and the variance of the demand): mean and standard deviation (in brackets).

Management policy (X1)	*Total sales*	*Average time in the system*	*RTY*	*Bullwhip Effect*
OUT inventory policy	22,027.0 (536.6)	58.07 (10.12)	89.84% (0.82%)	392.21 (48.35)
DBR methodology	27,552.7 (90.0)	34.16 (0.89)	93.37% (0.23%)	127.48 (49.36)

A growth in sales usually leads to an increase in operating expense. However, it does not occur in this case. The reason is the total time of the product in the supply chain. With the OUT policy, the product tends to be unnecessarily (far from the customer) stored, and consequently the percentage of defective products increases (*i.e.* the rolled throughput yield reduces). As a result of both effects, storage costs are dramatically higher in the reductionist approach.

It can also be noticed from inspection of *Tables 4 and 5* that variations in results are smaller with the DBR method. That is, the reductionist approach is more sensitive to the repetitiveness of the experiment. The holistic approach seems to be more robust.

5.5. Analysis by nodes of the results

Once TOC economic superiority in this noise-extended environment has been verified when compared to the classic OUT policy for the overall supply chain, this leads to an unavoidable key question: Do all supply chain members benefit in the same way from collaboration? Therefore, the research is moved towards the node-by-node analysis. We have carried out the same study for the net profit of each member.

When results are observed in detail, it can be noticed that the wealth generated by the holistic approach is not equitably distributed along the various supply chain echelons. *Table 6* exhibits the average net profit of the four nodes both when the DBR methodology and the OUT policy manage the supply chain. *Figure 9* displays the main effects of the management policy factor in order to graphically show the significant difference in how members benefit from collaboration.

When analyzing the reductionist approach, the dramatic economic consequences of the Bullwhip Effect within the supply chain arise. This phenomenon creates large differences in profits along the distribution system, although the gross margin is the same and the throughput only undergoes slight changes (due to defective products and storage). These variations lead to an increase in operating expense as it moves away from the customer. Nonetheless, there is not a great difference in factory and distributor. The reason could be the production limitation. This constraint is a good solution to tame the Bullwhip Effect at the factory [36], as prevents the factory from generating large variations, smoothing its behavior. This limitation tends to increase lost sales at the retailer, but in certain scenarios the cost reduction compensates it.

Table 6. Local results of the same treatment when both alternatives are used to manage the supply chain.

Factor	*OUT _ Net profit*	*DBR _ Net profit*	*Percentage increase*
Supply Chain	$16,165.24	$26,342.20	+62.96%
Factory	$2,175.99	$8,777.05	+303.36%
Distributor	$2,589.05	$5,477.56	+111.57%
Wholesaler	$5,486.78	$6,466.04	+17.85%
Retailer	$5,913.42	$5,621.55	-4.94%

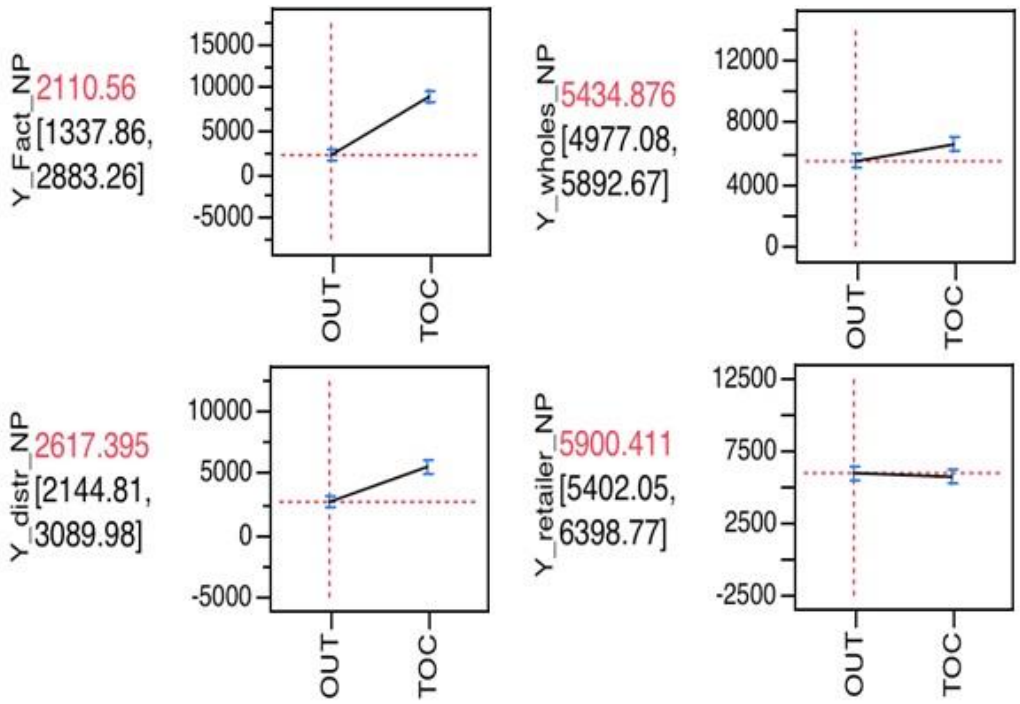


Figure 9. DoE screening: Main effects of the management policy factor on the various nodes.

As seen, the great increase in overall profits induced by the holistic management comes both from the growth in throughput (Goldratt's practices are aimed at protecting the bottleneck) and the decrease in operating expense (as the Bullwhip Effect is significantly reduced). Nevertheless, as TOC solution is based on keeping the inventory near to customer in order to minimize lost sales, the retailer will assume higher inventory costs. Thus, as shown in *Table 6*, cost distribution varies considerably. This causes that, when the DBR method is used, those members distant to customer obtain better results (in case of equality in margins), especially if unit costs are high. In the factory, storage costs are more relevant since the lead time is greater, while in the rest of the nodes, transport costs are more relevant.

In summary, although there is a dramatic improvement generated by the collaborative approach in the supply chain (overall profit is increased by 63%), this increase is much higher as the members move away from the customer. In fact, global optimization might lead to economic losses in some nodes –this could happen in the retailer in the case analyzed. In this context, the interest in adopting a collaborative policy will be very different at the various nodes. Therefore, if collaboration does not generate fair benefits in all echelons, some barriers to holism emerge in the supply chain, and the system would run into a non-optimal solution.

This node-by-node study brings evidence of the need for the fifth feature according to the Simatupang and Sridharan's framework [10]. Some kind of incentive alignment is required in order to achieve the system optimal solution through a Nash equilibrium, *i.e.* without incentives to deviate. That is, sharing costs, risks, and benefits among the various members is essential for taking the system from reductionism to holism. Thus, experimentation through simulation allows manager to anticipate to this problem by defining an appropriate cost structure within the supply chain.

6. Main Conclusions and Future Research

Although holistic supply chain solutions are considered to outperform traditional reductionist alternatives, they are not yet widespread since the adoption of an efficient collaborative solution requires a complex decision-making process to implement an appropriate framework [37]. This research focuses on this transition from local to global optimization understanding experimentation as a powerful engine for gaining knowledge. We employ an agent-based approach as a model-driven decision support system, where practitioners can explore a complex network of interdependences that would be unmanageable through other methodologies.

There are some aspects that the authors (with practical experience in supply chains and change management) consider essential in this transition. One of them is the educational phase that is required to move supply chain participants away from their natural individualistic behavior. Simulation can lead them to gain confidence in collaborative practices [38], since motivation is crucial in this decision-making process. These studies can reproduce the known environment (which would be inconceivable through an analytical approach), and allow managers to explore complex cause-effect relationships within an inexpensive and risk-free context.

This work provides evidence of TOC economic robustness in comparison with the OUT inventory policy when managing a four-echelon supply chain with several noise sources. The overall improvement came mainly from the increase of the throughput, which is a strong argument against traditional cutting costs-based management. However, paradoxically, the operating expense is also reduced due to the taming in the Bullwhip Effect and hence the reduction in storage costs. Several experiments have been carried out to statistically confirm this hypothesis, and the average increase in the net profit of the whole supply chain has been 63%.

Once the improvement is perceived, supply chain actors focus on the implementation. In the required integrative framework, information sharing acts as an (indispensable) enabler. This creates a visibility environment, which facilitates decision making to be carried out by a headquarter office that accounts for the interest of the whole system.

To find an appropriate collaborative solution for the supply chain, it is essential to integrate processes, synchronize decisions, and define a systemic performance scorecard [10]. To solve these issues, a solution based on Goldratt's TOC is proposed, in which the DBR method defines the collaborative behavior and the TA is used to determine the impact of the decisions on supply chain performance.

We provide evidence of the fact that the net profit distribution significantly varies when adopting collaborative solutions. While the OUT tends to damage upstream echelons due to the Bullwhip Effect, the TOC approach usually favors these members. Under these circumstances, aligning incentives within the supply chain is required. In this sense, trust is essential, and risks and benefits must be shared in order to avoid opportunistic behaviors [39]. Computer-based prototypes can be used by managers as business laboratories to define an appropriate cost structure within the supply chain.

Once studied the widely used Beer Game (serial) supply chain, future work is aimed at confirming the robustness of the holistic approach in divergent networks topologies.

In addition, we intend to further explore the reductionism-to-holism transition in terms of incentive alignment. One simple way companies can define a robust adaptive mechanism (it must be able to function over time) is by altering contracts with the aim of fairly distributing the benefit induced by collaboration. It means establishing linear contracts so that each node is rewarded according to its contribution.

A third avenue for future work is to incorporate Lean Management mechanisms in the agent-based system. We aim to use simulation for contrasting the TOC with the most known holistic paradigm. Our preliminary research suggests that while there is not a significant difference between both methodologies in low-noise scenarios, the TOC makes a difference when the supply chain faces harmful noise conditions.

References

1. T.M. Simatupang, A.C. Wright, R. Sridharan, Applying the Theory of Constraints to supply chain collaboration, Supply Chain Manag.: An Int. J., 9 (1) (2004), 57-70.
2. J. Costas, B. Ponte, D. De la Fuente, R Pino, R., J. Puche, J. Applying Goldratt's Theory of Constraints to reduce the Bullwhip Effect in a supply chain through agent-based modeling, Expert Syst. with Appl. 42 (4) (2015) 2049–2060.
3. G.P. Cachon, P.H. Zipkin, Competitive and cooperative inventory policies in a two-stage supply chain, Manag. Sci., 45 (7) (1999) 936-953.
4. T.M. Simatupang, R. Sridharan, The collaborative supply chain, Int. J. of Logist. Manag. 13 (1) (2002) 15-30.
5. J. Dejonckheere, S.M. Disney, M.R. Lambrecht, D.R. Towill, Measuring and avoiding the Bullwhip Effect: A control theoretic approach, Eur. J. of Oper. Res. 147 (3) (2003) 567-590.
6. N. Gilbert, Agent-based models, Sage Publications, London, 2008.
7. D.J. Power, R. Sharda, Model-driven decision support systems: Concepts and research directions, Decis. Support Syst. 43 (3) (2007) 1044-1061.
8. E.M. Goldratt, Theory of Constraints, North River Press, Croton-on-Hudson, 1990.
9. F. Chen, Z. Drezner, J.K. Ryan, D. Simchi-Levi, Quantifying the Bullwhip Effect in a simple supply chain: The impact of forecasting, lead times, and information, Manag. Sci. 46 (3) (2000) 436-443.
10. T.M. Simatupang, R. Sridharan, An integrative framework for supply chain collaboration, Int. J. of Logist. Manag. 16 (2) (2005) 257-274.
11. D.W. Smith, The measurement nightmare, Lucie Press, Boca Raton, 2000.
12. H.L. Lee, V. Padmanabhan, S. Whang, The Bullwhip Effect in supply chains, Sloan Manag. Rev. 38 (3) (1997) 93-102.

13. M. Holweg, S. Disney, J. Holmström, J. Småros. Supply Chain Collaboration: Making Sense of the Strategy Continuum. Eur. Manag. J. 23 (2) (2005) 170-181.

14. M. Moschandreas, The role of opportunism in transaction cost economics, J. of Econ. Issues 31 (1) (1997) 39-57.

15. E.M. Goldratt, E. Schragenheim, C.A. Ptak, Necessary but not sufficient, North River Press, Croton-on-Hudson, 2000.

16. V.J. Mabin, S.J. Balderstone, The performance of the theory of constraints methodology: analysis and discussion of successful TOC applications, Int. J. of Oper. & Prod. Manag. 23 (6) (2003) 568-595.

17. T. Corbett, Throughput accounting: TOC's management accounting system, North River Press, Croton-on-Hudson, 1998.

18. S. Cannella, J. Ashayeri, P.A. Miranda, M. Bruccoleri, Current economic downturn and supply chain: the significance of demand and inventory smoothing, Int. J. of Comput. Integr. Manuf. 27 (3) (2014) 201-212.

19. J.D. Sterman, Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment, Manage. Sci. 35 (3) (1989) 321-339.

20. S. Karlin, Dynamic inventory policy with varying stochastic demands, Manage. Sci. 6 (3) (1960) 231-258.

21. K.L. Croxon, S.J. García-Dastugue, D.M. Lambert, D.S. Rogers, The supply chain management processes, Int. J. of Logist. Manag. 12 (2) (2001) 13-36.

22. K. Youngman, A Guide to Implementing the Theory of Constraints, 2009, available at http://www.dbrmfg.co.nz/ (accessed 26 March 2015).

23. H.H. Wu, C.P. Chen, C.H. Tsai, T.P. Tsai, A study of an enhanced simulation model for TOC supply chain replenishment system under capacity constraint, Expert Syst. with Appl. 37 (9) (2010), 6435-6440.

24. H.H. Wu, A.H.I. Lee, T.P. Tsai, A two-level replenishment frequency model for TOC supply chain replenishment systems under capacity constraint, Comput. & Ind. Eng. 72 (1) (2014) 152–159.

25. F. Castiglione, Diffusion and aggregation in an agent based model of stock market fluctuations, Int. J. of Mod. Phys. C 11 (5) (2000) 865-879.

26. D.C. Chatfield, T.P. Harrison, J.C Hayya, SISCO: An object-oriented supply chain simulation system, Decis. Support Syst. 42 (1) (2006) 422-434.

27. C.M. Macal, M.J. North, Tutorial on agent-based modeling and simulation, in WSC'05 Proceedings of the 37th Conference on Winter Simulation, Orlando (2005) 2-15.

28. U. Wilensky, NetLogo, 1999, available at http://ccl.northwestern.edu/netlogo/ (accessed 28 April 2015).

29. A. Getchell, Agent-based modeling, Phys. 22 (6) (2008) 757-767.

30. E. Bonabeau, Agent-based modeling: Methods and techniques for simulating human systems, Proceedings of the National Academy of Sciences of the United States of America 99 (3) (2002) 7280-7287.

31. W. Zinn, H. Marmorstein, Comparing two alternative methods of determining safety stock levels: The demand and the forecast systems, J. of Bus. Logist. 11 (1) (1990) 95-110.

32. S.M. Disney, A. Maltz, X. Wang, R.D. Warburton, Inventory management for stochastic lead times with order crossovers, Eur. J. of Oper. Res. 248(2) (2016) 473-486.

33. J. Ravichandran, Six-sigma milestone: an overall sigma level of an organization, Total Qual. Manage. 17(8) (2006) 973-980.

34. R.A. Fisher, Statistical Methods, Experimental Design and Scientific Inference, Oxford University Press, New York, 1990.

35. J. Sall, A. Lehman, A., M.L. Stephens, L. Creighton, L. JMP start statistics: a guide to statistics and data analysis using JMP, SAS Institute, Cary, 2012.

36. S. Cannella, E. Ciancimino, A.C. Márquez, Capacity constrained supply chains: A simulation study. Inter. J. of Simul. and Process Model. 4 (2) (2008) 139-147.

37. J. Puche, B. Ponte, J. Costas, R. Pino, D. de la Fuente, Systemic approach to supply chain management through the viable system model and the theory of constraints, Prod. Plan. & Control in press.

38. M. Holweg, J. Bicheno, Supply chain simulation – a tool for education, enhancement and endeavour, Int. J. of Prod. Econ. 78 (2) (2002) 163-175.

39. X. Fu, M. Dong, S. Liu, G. Han, Trust based decisions in supply chains with an agent, Decis. Support Syst. 82 (1) (2016) 35-46.

CONCLUSIONS

The specific conclusions of each one of the six studies that conform the present Doctoral Thesis have been discussed separately in the previous chapters. Therefore, we devote this last section to extract some overall conclusions by seeing the whole picture. We aim to synthetize the main ideas obtained from this research as well as reflect on them to derive managerial implications. Moreover, we delineate some possible future avenues of research stemming from the works included in this Doctoral Thesis.

In general terms, this Doctoral Thesis highlights the imperative need to analyze the supply chain in its entirety. Through a simulation study carried out in a wide range of noise conditions, we have first observed that the interaction of individual strategies in supply chains (*reductionism*) creates a largely inefficient scenario, mainly as it fosters the Bullwhip phenomenon. This fact unquestionably damages the profit margin of the various nodes. Secondly, we have brought evidence of how managing globally the supply chain (*holism*) results in breakthrough improvements. In this sense, collaboration generates win-win solutions: the inventory (operating costs) may dramatically decrease while paradoxically the customer fill rate (sales revenue) surges. That is, supply chain managers can benefit from an enhanced financial statement by analyzing the interdependencies among processes and decisions across the supply chain.

Under these circumstances, a key question lies in why, even when practitioners concede that supply chains are strongly built on interdependencies and the huge benefits derived from collaboration, these global optimization approaches are far from being widespread (Schneider, 2013) —in other words, why supply chain collaboration often fails in practice (Fawcett et al., 2015). To tackle this matter, we strongly concur with Simatupang and Sridharan's (2005) view, who underscore five indispensable fields of collaboration: (1) information transparency; (2) overall performance system; (3) process integration; (4) decision synchronization; and (5) incentive alignment. All of them must be taken into consideration to ensure the viability and take full advantage of the collaborative solution.

From this point on, we have designed a collaborative framework for supply chains built on the integration of the *Viable System Model* (VSM) (Beer, 1984) and the *Theory of Constraints* (TOC) (Goldratt, 1990). Both methodological notions perfectly fit together: the VSM defines the systemic structure of the supply chain (orchestrates the framework), while the TOC implements the systemic behavior of the supply chain (integrate processes around the main goal of the system).

The materialization of this framework has been investigated in detail by means of agent-based techniques. In particular, we adopt a bottleneck orientation to design a *Drum-Buffer-Rope* (DBR) mechanism for governing the material flow of the supply chain. We also define a suitable scorecard based on the *Throughput Accounting* (TA) to guide supply chains towards achieving their main financial goals. This approach has shown to increase enormously both the efficiency and the agility of the system in comparison with traditional alternatives based on the mass production paradigm.

We have also explored the role of artificial intelligence-based techniques in forecasting customer demand. *Artificial neural networks* (ANNs) —both under multi-layer perceptron (MLP) and radial basis function (RBF) architectures— present a powerful model to deal with complex time series, as they are able to accurately capture its trend and its seasonality. Agent-based tools enabled us to create a forecasting mechanism that hourly selects the best forecasting method in function of recent demand, as well as to integrate the forecasting model in a system with a wider scope. The designed forecasting system provides great performance in terms of alleviating the generation of the Bullwhip phenomenon, which positively impacts the management of the supply chain.

From this perspective, we forcefully stand up for artificial intelligence-based modeling techniques, such as the aforementioned agent-based systems, as powerful laboratories for business exploration and transformation. These prototypes —that can reproduce the known scenario and enable managers to study complex relationships within a cost-free and risk-free environment— may act as a catalyst to move supply chain participants away from their natural individualistic behavior. In addition, they can also be highly useful in the transition process from reductionism to holism in the supply chain. They allow managers to develop and implement an integrative framework for collaboration, being also useful in the essential phase of aligning incentives; i.e. the nodes' motivations to deviate from the collaborative scheme must be completely removed.

This approach to problem-solving and decision-making support, which is summarized in figure 1, consists of a three-step procedure: (1) modeling and implementation; (2) simulation and analysis; and (3) real-world development. It should be highlighted that a major advantage of the use of agents lie in the modular nature of this prototyping methodology. Agent-based modeling and simulation systems are highly flexibles: they can be easily integrated in a system with a wider scope and/or they can be simply scaled to additional restrictions and properties.

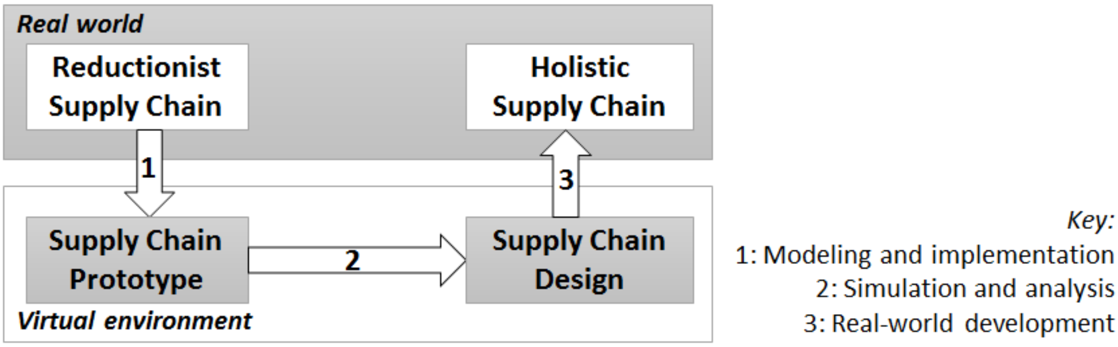


Figure 1. Prototyping for business decision-making.

There are plenty of future research works that we wish to carry out as next steps on this intriguing topic.

Incorporating Lean Management mechanisms in the collaborative framework will be a first avenue for future work. Lean is a brilliant systemic philosophy focused on minimizing sources of waste in the system, and we will use simulation tools to compare it versus the TOC-based framework developed. A preliminary analysis suggests that Lean, based on simple mechanisms, offers a great performance in moderately complex scenarios, but TOC can make a difference when facing harmful conditions.

We also aim to explore the development a robust adaptive mechanism (it must function over time) for aligning incentives throughout the supply chain. We believe that game theory is a powerful approach to this issue. It would entail taking into account aspects such as the contribution made by each node to the collaborative solution and its bargaining power. Moreover, rules and controls must be placed to prevent opportunistic behaviors of the nodes against the supply chain major interests.

The lack of considering the impact of the topology of the supply chain, as well as morphogenetic aspects, in the simulation study can be understood as a noticeable limitation. In this sense, we wish to investigate the effect of the horizontal structure of the supply chain on the results obtained. Furthermore, considering the dynamic behavior of the supply chain not only in steady state but also against ramp-up (an increase in firm production ahead of anticipated increases in product demand) and phase-out (the end of the life cycle of a product) scenarios represents an interesting field to explore.

References

1. Beer, S. (1984). The viable system model: Its provenance, development, methodology and pathology. *Journal of the Operational Research Society*, 35(1), 7-25.

2. Fawcett, S. E., McCarter, M. W., Fawcett, A. M., Webb, G. S., & Magnan, G. M. (2015). Why supply chain collaboration fails: the socio-structural view of resistance to relational strategies. *Supply Chain Management: An International Journal*, 20(6), 648-663.

3. Goldratt, E. M. (1990). *Theory of Constraints*. Croton-on-Hudson: North River.

4. Schneider, D. 2013. Systems thinking and the supply chain. Supply Chain Digest [online], available from: http://www.scdigest.com/assets/FirstThoughts/13-08-09.php?cid=7297 [last access 20.02.2016]

5. Simatupang, T. M., & Sridharan, R. (2005). An integrative framework for supply chain collaboration. *International Journal of Logistics Management*, 16(2), 257-274.

CONCLUSIONES

Las conclusiones específicas de cada uno de los seis trabajos que componen esta Tesis Doctoral han sido presentadas por separado en los capítulos anteriores. Esta última sección observa el problema en su totalidad. Nuestra intención es sintetizar las principales ideas derivadas de este trabajo, así como discutir sus implicaciones prácticas en el campo de la gestión. Además, trazaremos posibles líneas de futura investigación en este campo.

En términos generales, esta Tesis Doctoral subraya la imprescindibilidad de analizar la cadena de suministro desde una perspectiva sistémica. A través de un estudio de simulación llevado a cabo en un amplio rango de escenarios, hemos observado que la interacción de estrategias individuales en la cadena de suministro (*reduccionismo*) crea un escenario altamente ineficiente, ya que potencia la generación del Efecto Bullwhip. Este hecho empeora el resultado financiero de los distintos nodos que conforman la cadena de suministro. A partir de ahí, hemos demostrado cómo una gestión global del sistema (*holismo*) guía la cadena de suministro hacia una situación financiera mucho más favorable. La colaboración genera soluciones del tipo *win-win*: el inventario (y por lo tanto los costes operacionales) decrece enormemente a la vez que, paradójicamente, el nivel de servicio al consumidor (y por lo tanto los ingresos) aumenta. De esta forma, los gestores de cadenas de suministro pueden obtener grandes mejoras, tanto operacionales como financieras, si analizan las interdependencias entre los distintos procesos y decisiones a lo largo de toda la cadena de suministro.

Bajo estas circunstancias, una cuestión clave reside en por qué —incluso cuando los profesionales reconocen las cadenas de suministro como un sistema construido sobre interdependencias y las ventajas derivadas de buscar la optimización global— las soluciones colaborativas no son habituales en las cadenas de suministro reales (Schneider, 2013). En otras palabras, por qué la colaboración a veces falla en la práctica (Fawcett et al., 2015). Para estudiarlo, nos basamos en la visión de Simatupang y Sridharan (2005), quienes señalan cinco factores clave en la colaboración en cadenas de suministro: (1) transparencia de información; (2) uso de indicadores sistémicos; (3) integración de procesos; (4) sincronización en la toma de decisiones; y (5) alineamiento de incentivos. Todas ellas han de tenerse en cuenta con el objetivo de asegurar la viabilidad de la solución colaborativa y obtener el máximo rendimiento de ella.

Desde esa base, hemos diseñado un marco colaborativo para cadenas de suministro basado en la integración del Modelo de los Sistemas Viables (*Viable System Model*, VSM) (Beer, 1984) y la Teoría de las Restricciones (*Theory of Constraints*, TOC) (Goldratt, 1990). Ambas encajan a la perfección: el Modelo de los Sistemas Viables define la estructura sistémica de la cadena de suministro (orquesta la colaboración), mientras la Teoría de las Restricciones implementa el comportamiento colaborativo del sistema (integra los procesos en torno al objetivo principal del sistema).

La materialización de este marco se ha explorado a través de técnicas basadas en agentes. En concreto, a través de un mecanismo *Drum-Buffer-Rope* (DBR) hemos adoptado una orientación basada en el cuello de botella del sistema para gestionar el flujo de materiales a lo largo del mismo. Además, se ha definido un cuadro de mando de acuerdo a la Contabilidad del Throughput (*Throughput Accounting*, TA) con el objetivo de guiar la cadena de suministro hacia sus objetivos financieros. Esta solución incrementa significativamente tanto la eficiencia como la agilidad del sistema en comparación con alternativas tradicionales basadas en el paradigma de la producción en masa.

Por otro lado, hemos investigado la eficacia de las técnicas basadas en la inteligencia artificial en la previsión de la demanda del consumidor. En este sentido, las redes neuronales artificiales (*artificial neural networks*, ANNs) —tanto en arquitecturas del tipo perceptrón multicapa (*multi-layer perceptron*, MLP) como del tipo función de base radial (*radial basis function*, RBF)— permiten capturar tanto la tendencia como la estacionalidad de series temporales complejas. A través de herramientas basadas en agentes, hemos creado un mecanismo de previsión que selecciona en cada momento el mejor método de previsión, y hemos incorporado el sistema de previsión dentro de un sistema multi-agente con un objetivo más amplio. El mecanismo desarrollado ofrece un gran rendimiento en la reducción del Efecto Bullwhip, lo cual ha demostrado tener un impacto muy positivo sobre la gestión de la cadena de suministro.

Así, esta Tesis Doctoral resalta el papel de las técnicas de modelado y simulación, como los mencionados sistemas basados en agentes, como poderosos laboratorios de ensayo para el análisis empresarial. En primer lugar, estos prototipos —que pueden reproducir con precisión un escenario conocido y habilitan a los managers para estudiar complejas relaciones en un escenario libre de riesgos y costes— pueden actuar como un catalizador para alejar a los miembros de la cadena de suministro de comportamientos basados en la optimización local. Además, pueden ser muy útiles en el proceso de transición desde soluciones reduccionistas hasta soluciones holistas en la cadena de suministro, facilitando el desarrollo y la implementación de un marco integrador para la colaboración. También en la esencial fase del alineamiento de incentivos, orientada a eliminar los motivos de los distintos miembros para alejarse del comportamiento colaborativo.

Esta aproximación a la resolución de problemas y al apoyo en la toma de decisiones empresariales, ilustrada en la figura 1, consiste en un procedimiento de tres pasos: (1) modelado e implementación; (2) simulación y análisis; y (3) desarrollo en el mundo real. Una ventaja significativa del uso de agentes en este proceso se encuentra en la naturaleza modular de esta metodología. Los sistemas basados en agentes son muy flexibles: pueden ser fácilmente integrados en sistemas con horizontes más amplios, así como simplemente adaptados a nuevos requisitos y propiedades en los distintos agentes que lo forman.

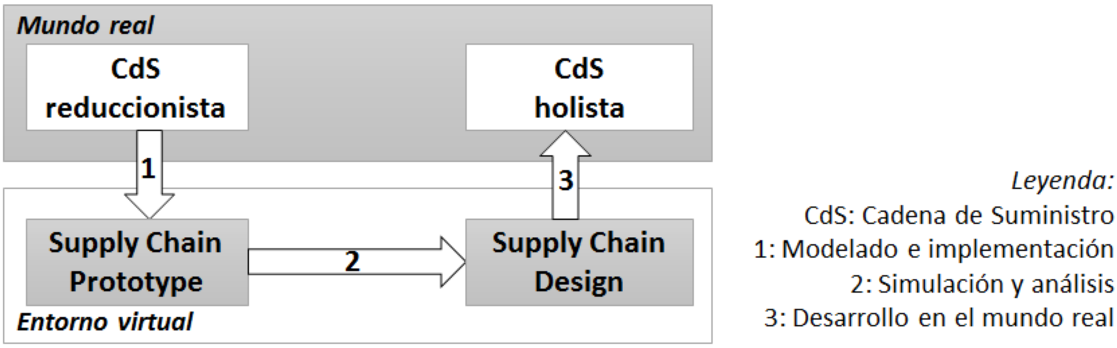


Figura 1. Uso de prototipos en la toma de decisiones empresariales.

Por último, trazamos distintas líneas de futura investigación como próximos pasos en este campo.

Una primera línea para futuros trabajos se basa en la incorporación de mecanismos basados en Lean Management al marco colaborativo. Lean es una metodología sistémica focalizada en la reducción de las fuentes de desperdicios en el sistema. Tenemos la intención de utilizar el sistema basado en agentes para comparar éstos con los diseñados de acuerdo a la Teoría de las Restricciones. Un análisis preliminar nos ha sugerido que Lean, que se basa en mecanismos más simples, ofrece un gran rendimiento en entornos con una complejidad moderada, mientras que la Teoría de las Restricciones marca las diferencias cuando la cadena de suministro opera un escenario con muchas fuentes de incertidumbre.

También pretendemos explorar el desarrollo de un mecanismo adaptativo robusto (ya que ha de ser capaz de funcionar a lo largo del tiempo) para el alineamiento de incentivos en la cadena de suministro. La teoría de juegos ofrece una interesante aproximación a esta materia, ya que podríamos tener en cuenta aspectos como la contribución realizada por cada nodo y su poder de negociación. Además, algún tipo de control es necesario a lo largo de la cadena de suministro para evitar comportamientos oportunistas de los distintos nodos que la forman en contra del interés global del sistema.

Nótese, además, que esta investigación no ha considerado aspectos topológicos ni morfogenéticos en la cadena de suministro, lo cual podría ser entendido como una limitación del mismo. En este sentido, tenemos la intención de investigar el impacto de la estructura horizontal de la cadena de suministro sobre los resultados obtenidos. Igualmente, queremos considerar el comportamiento dinámico de los agentes de la cadena de suministro no solo en régimen permanente, sino también en las distintas fases del ciclo de vida del producto.

Referencias

1. Beer, S. (1984). The viable system model: Its provenance, development, methodology and pathology. *Journal of the Operational Research Society*, 35(1), 7-25.
2. Fawcett, S. E., McCarter, M. W., Fawcett, A. M., Webb, G. S., & Magnan, G. M. (2015). Why supply chain collaboration fails: the socio-structural view of resistance to relational strategies. *Supply Chain Management: An International Journal*, 20(6), 648-663.
3. Goldratt, E. M. (1990). *Theory of Constraints*. Croton-on-Hudson: North River.
4. Schneider, D. 2013. Systems thinking and the supply chain. Supply Chain Digest [online], available from: http://www.scdigest.com/assets/FirstThoughts/13-08-09.php?cid=7297 [last access 20.02.2016]
5. Simatupang, T. M., & Sridharan, R. (2005). An integrative framework for supply chain collaboration. *International Journal of Logistics Management*, 16(2), 257-274.

Informe de calidad de las publicaciones

Tesis Doctoral: **Bullwhip Effect Reduction through Artificial Intelligence-Based Techniques**

Presentada por: **Borja Ponte Blanco**

Dirigda por: **David de la Fuente García y Raúl Pino Díez**

- -

CHAPTER 1

Reference: Ponte, B., Ruano, L., Pino, R., & de la Fuente, D. (2015). The Bullwhip effect in water demand management: taming it through an artificial neural networks-based system. *Journal of Water Supply: Research and Technology-Aqua*, 64(3), 290-301.

DOI: 10.2166/aqua.2015.087

Available online at: http://aqua.iwaponline.com/content/64/3/290

Journal of Water Supply: Research and Technology—AQUA (http://aqua.iwaponline.com/), indexed in the ISI, is a Q3-journal in the category 'Engineering: Civil' (Science Edition). Impact Factor 2015: 0.807.

CHAPTER 2

Reference: Ponte, B., Fuente, D. D. L., Pino, R., & Rosillo, R. (2015). Real-time water demand forecasting system through an agent-based architecture. *International Journal of Bio-Inspired Computation*, 7(3), 147-156.

DOI: 10.1504/IJBIC.2015.069559

Available online at: http://www.inderscienceonline.com/doi/abs/10.1504/IJBIC.2015.069559

International Journal of Bio-Inspired Computation (http://www.inderscienceonline.com/loi/ijbic), indexed in the ISI, is a Q2-journal in the category 'Computer Science: Theory and Methods' (Science Edition). Impact Factor 2015: 1.390.

CHAPTER 3

Reference: Ponte, B., de la Fuente, D., Parreño, J., & Pino, R. (2016). Intelligent Decision Support System for Real-Time Water Demand Management. *International Journal of Computational Intelligence Systems*, 9(1), 168-183.

DOI: 1080/18756891.2016.1146533

Available online at: http://www.tandfonline.com/doi/abs/10.1080/18756891.2016.1146533

International Journal of Computational Intelligence Systems (http://www.tandfonline.com/toc/tcis20/current), indexed in the ISI, is a Q4-journal in the category 'Computer Science: Artificial Intelligence' (Science Edition). Impact Factor 2015: 0.391.

CHAPTER 4

Reference: Puche, J., Ponte, B., Costas, J., Pino, R., & de la Fuente, D. (2016). Systemic approach to supply chain management through the viable system model and the theory of constraints. *Production Planning & Control*, 27(5), 421-430.

DOI: 10.1080/09537287.2015.1132349

Available online at: http://www.tandfonline.com/doi/abs/10.1080/09537287.2015.1132349

Production Planning & Control: The Management of Operations (http://www.tandfonline.com/toc/tppc20/current), indexed in the ISI, is a Q2-journal in the category 'Operations Research and Management Science' (Science Edition). Impact Factor 2015: 1.532.

CHAPTER 5

Reference: Costas, J., Ponte, B., de la Fuente, D., Pino, R., & Puche, J. (2015). Applying Goldratt's Theory of Constraints to reduce the Bullwhip Effect through agent-based modeling. *Expert Systems with Applications*, 42(4), 2049-2060.

DOI: 0.1016/j.eswa.2014.10.022

Available online at: http://www.sciencedirect.com/science/article/pii/S0957417414006460

Expert Systems with Applications (http://www.journals.elsevier.com/expert-systems-with-applications/), indexed in the ISI, is a Q1-journal in the category 'Operations Research and Management Science' (Science Edition). Impact Factor 2015: 2.981.

CHAPTER 6

Reference: Ponte, B., Costas, J., Puche, J., de la Fuente, D., & Pino, R. (2016). Holism versus reductionism in supply chain management: An economic analysis. *Decision Support Systems*, 86, 83-94.

DOI: 0.1016/j.dss.2016.03.010

Available online at: http://www.sciencedirect.com/science/article/pii/S0167923616300446

Decision Support Systems (http://www.journals.elsevier.com/decision-support-systems/), indexed in the ISI, is a Q1-journal in the category 'Operations Research and Management Science' (Science Edition). Impact Factor 2015: 2.604.