TÉCNICO LISBOA

Model predictive control for optimal supply chain management

Eduardo Guilherme Ferreira Morais de Araújo

Thesis to obtain the Master of Science Degree in

Mechanical Engineering

Supervisors: Prof. Miguel Afonso Dias de Ayala Botto
Prof. João Miguel Lemos Chasqueira Nabais

Examination Committee

Chairperson: Prof. Paulo Jorge Coelho Ramalho Oliveira
Supervisor: Prof. João Miguel Lemos Chasqueira Nabais
Member of the Committee: Prof. Maria do Rosário Mauricio Ribeiro Macário

June 2019

To my parents.

Acknowledgments

I want to thank my supervisors Prof. Miguel Ayala Botto of Instituto Superior Técnico (IST), Universidade de Lisboa, and Prof. João Lemos Nabais of Instituto Politécnico de Setúbal, for their close supervision over the course of this work. Both Prof. Botto and Prof. Nabais were always available whenever I had questions about my research or writing. They always believed in my ability to overcome the difficulties faced during this process, and consistently steered me into the right direction whenever I needed it. Please let my profound appreciation be recognized for all the encouragement and opportunities given. I could not have asked for better advisors.

I would also like to thank Prof. Beatriz Silva, and Prof. Ana Carvalho of IST for guiding and helping me during my troubled first years at this institute. In particular, Prof. Beatriz Silva's mentorship was of paramount importance for me to adequately balance my academic and personal lives.

Nevertheless, I have to specially thank my dear friends Ricardo Maia and Maria Sousa for the endless, stimulating discussions, for the sleepless nights we spent working together before deadlines, and for all the fun we have had during these past years.

Finally, I must express my profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study. This accomplishment would not have been possible without them. Thank you.

Resumo

O crescente consumo e globalização das últimas décadas aumentou a necessidade de meios de trasporte, intensificando a concorrência interna deste sector. Por outro lado, para responder à crescente procura em termos de personalização de produto, custo e qualidade de serviço por parte do cliente, as empresas são obrigadas a reduzir custos mantendo os mesmos padrões de qualidade, o que para além do aquecimento global e outras preocupações ambientais, tem pressionado distribuidores, e empresas provedoras de transporte em geral, para reduzir custos de movimento. Esta forte concorrência entre donos de mercadoria e provedores de transporte tem criado um permanente estado de tensão entre os dois sectores, aumentando a procura por níveis mais elevados de eficiência, qualidade, pontualidade, e capacidade de resposta, ao longo das cadeias de abastecimento.

Para resolver o dilema entre (i) entregas pontuais, e (ii) gestão eficiente de transporte, este trabalho apresenta uma nova abordagem dinâmica, baseada em técnicas de controlo predictivo (*MPC*), que integra operações de transporte à gestão em tempo-real de cadeias de abastecimento. Concentrando-se no caso de tempo discreto, o método desenvolvido resulta em representações lineares e invariantes no tempo em espaço-de-estados de cadeias de abastecimento que são controláveis e observáveis. O desempenho da metodologia proposta é ilustrado recorrendo a um caso-de-estudo baseado num cenário real. Os resultados demonstram que o controlador desenvolvido é capaz de (1) lidar com multi-produtos e multi-transportes, (2) gerir inventários, (3) monitorizar trabalho-em-progresso (*WIP*), e (4) sequencializar operações de manufactura e transporte, de uma forma autónoma e integrada respeitando janelas-temporais predefinidas.

Palavras-chave: MPC, Multi-Produto, Multi-Transporte, Gestão de Cadeias de Abastecimento, Tempo-real, Reference Tracking

Abstract

The increasing consumption and globalization over the last decades have created the need for more transportation, strengthening the internal competition of this sector. In turn, to respond to the growing demand in terms of product customisation, price and service levels from the customer side, companies are urged to lower their costs, while still maintaining high quality standards, which in addition to global warming and other environmental concerns have pressured distributors, and transportation service providers in general, towards reducing movement costs. This strong competition between goods owners and transport providers have created a permanent state of tension between the two sectors, increasing the demand for higher levels of efficiency, quality of service, timeliness, and responsiveness across supply chains.

To solve the trade-off between (i) on-time delivery, and (ii) efficient transportation management, this work presents a novel dynamic approach for real-time supply chain management integrating transportation operations, based on a model predictive control framework. Focusing on the discrete time case, a method for developing linear, time-invariant, state-space representations for supply chains that are both controllable and observable is outlined. The performance of the proposed methodology is illustrated resorting to a case study based on a real-world scenario. The results demonstrate that the devised controller is able to (1) deal with multi-products and multi-transports, (2) manage stocks, (3) monitor WIP, and (4) schedule manufacturing and transportation operations in an autonomous and integrated way, while respecting predefined time-windows.

Keywords: MPC, Multi-Product, Multi-Transport, Supply Chain Management, Real-Time, Reference Tracking

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Industry and trade were — and continue to be — the driving forces behind economical growth and the improvement of living standards [45, 48, 49, 61]. This chapter opens with the presentation of the historical context of industry and trade, and their social and economical impact. The discussion proceeds to present how and why did the concept of *Industry 4.0* spread across the world. Section 1.2 discusses the implications of the fourth industrial revolution on supply chain management, presenting the fundamental issues and challenges, as well as the most promising paradigm-shifting technologies in supply chain management. A literature review is outlined in Section 1.3. Section 1.4 presents the objectives and contributions of this work.

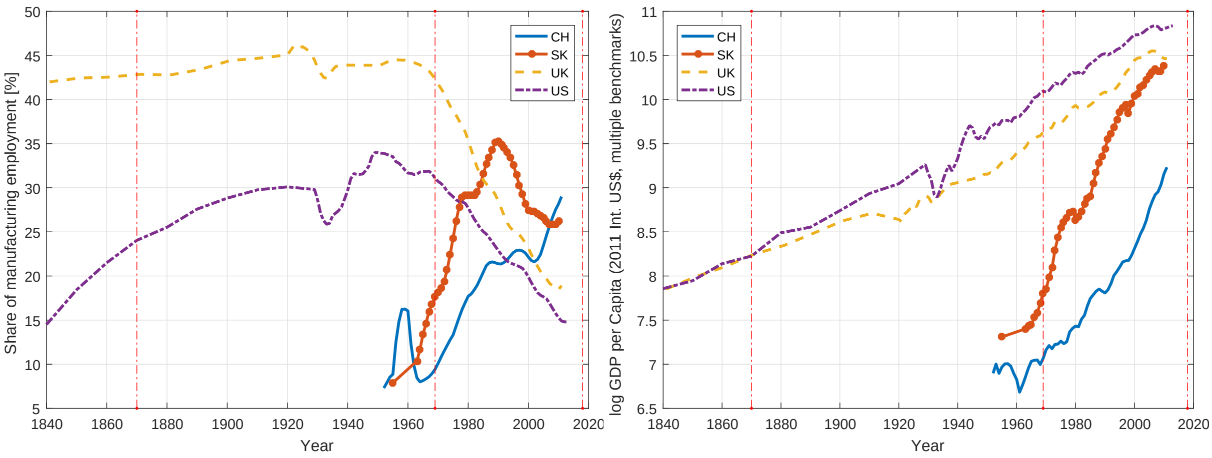
1.1 Historical context — Industry and trade

The movement of people and goods have always been at the heart of all civilizations, and it is not by coincidence that the first highly civilized nations were located at Asia Minor, in the vicinity of the Mediterranean Sea. It was there trade relationships first developed and commerce transportation started to connect different regions and allowing people to take advantage of differences in climate and of special skills which had been developed in previously isolated regions [36, 42, 55].

However, this link between economic growth and the increase of mobility, accessibility and trade was made explicitly apparent only in the wake of the industrial revolution [55], when humanity witnessed the first structural transformation, i.e., when the economic centre of gravity shifted from agriculture to manufacturing. This one-sector growth model became so effective and popular ever since that it was listed by Kuznets [39] in his Nobel prize lecture as one of the six main features of modern economic growth [25].

The shift away from agriculture to non-agricultural pursuits, and more recently from industry to services in developed countries, is such a key aspect of economic development that it is still an active research topic currently. One way to study the process of structural transformation across countries is to track how employment changes across sectors in the economy [51, 59]. To illustrate this phenomenon one can focus on four representative nations: China, South Korea, United Kingdom, and United States. As can be seen in Figure 1.1(a), the evolution of the share of employment in manufacturing over time, reveals the aforementioned shift of the economic activity from agriculture to manufacturing, and later from industry to services in developed countries.

One can still further investigate and compare the GDP per capita with the employment in manufacturing. To that end, Figure 1.1(b) shows that in the case of developed countries, there is a clear hump-shaped relationship between industrialization (measured by employment shares) and incomes [51, 59], which developing countries like China are rapidly catching-up.



(a) (b)

Figure 1.1: Socio-economic Impact of Industrial Revolutions in China, South Korea, United Kingdom, and United States.

There is no doubt the development of manufacturing industry around the world has always been supported by technological and scientific advances, which significantly improve both product offering and operating processes [4]. From a technological evolution perspective, one can identify three stages (or revolutions) which took place during a period of around two centuries [34, 43]. The first industrial revolution started in the second half of the eighteen century with the introduction of water and steam-powered mechanical production facilities and lasted until the end of the nineteenth century. In 1870 the first production line was created (Cincinnati slaughterhouses), marking the advent of the division of labour which in addition to the electrification led to the second industrial revolution. In a response to the car manufacturing industry, Modicon started in 1968 the development of programmable logic controllers (PLCs). One year later, the first PLCs (Modicon 084) were already being implemented in industry, marking the advent of the third industrial revolution, which has been focusing on the automation of manufacturing supported by electronics and information technology until today [34, 43, 52].

One can not overstate the importance these revolutions, in addition to trade liberalization, had in western civilization. Consider once again Figure 1.1 where it is possible to see (1) the impact of each revolution (represented by the vertical red lines) in terms of industrialization (measured by employment shares) per country (Figure 1.1(a)), and (2) the economic growth increasing overtime (Figure 1.1(b)). In turn, Table 1.1 presents a more recent account of the impact of this sector in economic terms.

The key-factor in every revolution has been a relentless strive for efficiency. This is explicit in the case of the third revolution, for only an ever efficient industrial sector would be able of supporting a constant economic-growth overtime while decreasing its need for human resources. However, the rapid de-industrialization of the past few decades across the developed world, may pose difficult questions for the near future. Namely, knowing that (1) although both service and industry sectors are characterised by a concentration of both highly

Table 1.1: World development indicators for the year 2016 (data from The World Bank.)

	Europe & Central Asia	**East Asia & Pacific**	**North America**
Industry value added (% of GDP)[1]	23.0	33.6	18.9
Employment in industry (% of TE)[2]	25.1	25.1	18.9
Manufactures exports (% of ME)[3]	73.0	80.4	61.4

[1] Industry (including construction), value added (% of GDP).
[2] Employment in industry (% of total employment) (modeled ILO estimate).
[3] Manufactures exports (% of merchandise exports)

and low skilled jobs, the ratio of skilled jobs is higher in industry [22], and that (2) a stable middle-class is key to the economic growth of a country [35], one can not but wonder the impact a collapsing (or at least diminishing) middle class may have upon global, interconnected economies. Furthermore, industrial success and innovation have a symbiotic relationship, they both depend on each other, rendering industry a key driver of research, innovation, productivity, job creation and exports [22, 29].

It is not surprising then that after a group of Germany representatives from various fields (such as business, politics, and academia) presented in 2011 at the Hanover Trade Fair their strategy to prepare and strengthen the industrial sector for the future production requirements, most of the other nations followed and "Industry 4.0" became an instant buzzword [29, 52].

As an answer to Germany, that same year the Unite States government promoted the "Advanced Manufacturing Partnership", a series of discussions, actions and recommendations to make sure the US would be prepared to lead the next generation of manufacturing. In 2013, France defined its industrial policy priorities presenting 34 industry-based initiatives called "La Nouvelle France Industrielle". The United Kingdom also presented its long-term strategy "Future of Manufacturing" as a means of supporting the growth and resilience of its industry sector over the future decades. In 2014, the European Commission launched its "Factories of the Future", a public-private partnership to ensure funding until 2020. In turn, South Korea government lead the way in Asia with its "Innovation in Manufacturing 3.0", followed by the Chinese government's "Made in China 2025" one year later. In 2015, the Japanese government launched the “Industrial Value Chain Initiative”. Singapore committed $19 billion to its "Research, Innovation and Enterprise" in 2016, followed by Taiwan's "Smart Machinery Industry" one year later [43, 52].

Concomitantly with this efforts at the nation level, there were also private-sector plans to catalyse and coordinate priorities. These private efforts were triggered by some American companies like AT&T, Cisco, General Electric, IBM and Intel, and rapidly spread world-wide, first to Europe and then to Asia, making the fourth industrial revolution become one of the most frequently discussed topics of many manufacturing conferences, forums and exhibitions in the past few years [43, 52]. For the first time in the history of the industrial revolution the future started to be shaped in a coordinated manner by the joint-efforts of researchers, firms and governments as a way to avoid loosing ground to competitors, both at the corporation as well as at the nation level [22, 34].

Despite the plethora of names used by nations and firms ("Industry 4.0", "Factories of the Future", "Made in China 2025", and so on), they all refer to the same thing: the need to shift away from a machine dominant manufacturing to a digital one [29, 52]. However, it is not yet clear what Industry 4.0 will consist of, or even what it really means. Some researchers are focusing their attention on digitization, defending this revolution will be dominated by intelligent algorithms capable of processing the enormous amounts of data that are being now produced, and assisting (or even substitute) human decision-making. Others defend it will be communication the dominant aspect, envisaging a more human-free manufacturing environment where machines communicate directly with one another. Others, in turn, list flexibility and decentralization as the main factors of the fourth industrial revolution [29, 34, 52]. According to the literature, perhaps the best definition was proposed by Oztemel & Gursev [52]: "Industry 4.0 is a manufacturing philosophy that includes modern automation systems with a certain level of autonomy, flexible and effective data exchanges anchored in the implementation of next generation production technologies, innovation in design, and more personal and more agile in production as well as customized products". Such a definition is able to encompass all the key dominant aspects the next generation of manufacturing needs to have: consistent connectivity and computerization, self-adapting production systems based on transparency and predictive power, and autonomous and decentralised decision making [29].

1.2 Towards Industry 4.0 — A supply chain management perspective

A supply chain (SC) is the set of value-adding activities (i.e. flows and processes) that connects the suppliers and customers of a given company:

$$\text{Receive input from supplier} \rightarrow \text{Process} \rightarrow \text{Deliver to customer,}$$

where "supplier" may refer to one or more external vendors, or one or more upstream processes within the company itself. Similarly, a customer may refer to the actual final customer of the finished product or service, or rather one or more downstream operations which receive as an input the output of one or various other processes [24]. Thus, a SC is a network of various entities (suppliers, manufacturers, distributors, and retailers) which cooperate and coordinate along the entire value chain to provide customers the right product, in the right amount, at the right time, for the right price, and at the right place [13, 32, 41].

As supply chains extend over wide geographical areas they become ever more vulnerable to various global risks. In order to respond to the growing demands in terms of product customization, price and service levels from the customer side, while managing a highly dynamic environment due to economic, social and natural factors, companies must build resilient and flexible supply chains, capable of adapting to changes in the environment. In this sense, information sharing is of paramount importance, since the lack or distortion of information increases uncertainty [15, 31].

Uncertainty is a fundamental property of complex systems and it is at the heart of all supply chain problems. To better understand this, consider the following definitions. *Uncertainty* measures the knowledge one has about the system and the conditions in which it develops. *Risk*, in turn, refers to the deviations (positive or negative) from an expected behaviour of the system. It is clear then that the less uncertainty there is, the less risk exists. The main problem when dealing with complex systems however is related with *disturbances*. That is, the impact perturbations may have on the system, which is linked with its degree of uncertainty, and risk. In the case of supply chains, these perturbations may produce *disruptions*, which can in the limit cease production translating into great revenue losses [27, 31]. Figure 1.2 summarizes these concepts and the relationship between them.

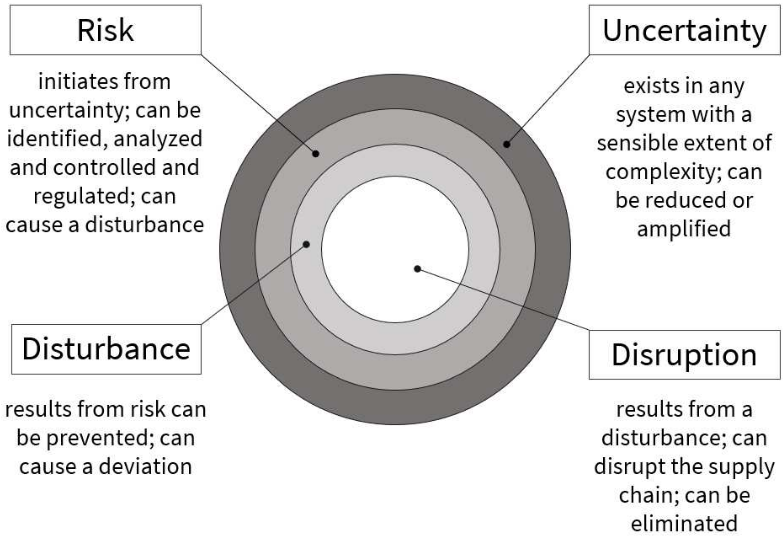


Figure 1.2: Interrelations of uncertainty, risk, disturbance and disruption (based on Ivanov (2018) [31]).

Risks arise from uncertainty and can be classified as *operational (or recurrent) risks*, related to random uncertainty, such as demand fluctuations; or *disruption risks*, related with the impact of sporadic (or exceptional) perturbations, such as natural phenomena, may have upon the SC [31]. These risks are intimately related with the two fundamental problems when dealing with supply chains: the Bullwhip and the Ripple effects, respectively.

The Bullwhip effect is not a new phenomenon and can be explained as the magnification of demand variability that propagates backwards in a SC. That is, the information of irregular orders taking place at the down-stream part of the supply chain is transmitted up-stream and grows ever more distinct as one moves up the SC. As a result, each supply chain member will over- or under-estimate product demand, producing exaggerated fluctuations, which requires them to hold excessive inventory levels [15, 31].

The Ripple effect, occurs when an unforeseen disruption (e.g. explosion, tsunami) propagates down-stream affecting the supply chain performance, such as sales, service level, or annual revenue. These disruptions are usually of a low-frequency-high-impact type, with the original disruption causing disruption propagation through the SC [31]. Table 1.2 summarizes and compares these two fundamental problems.

Table 1.2: Ripple effect and bullwhip effect (based on Ivanov (2018) [31]).

	Ripple effect	**Bullwhip effect**
Risks	Disruptions (e.g. explosion)	Operative (e.g. demand fluctuation)
Affected areas	Structures and critical parameters (such as service level and total costs)	Operative parameters such as lead-time and inventory
Recovery	Middle- and long-term; significant coordination efforts and investment	Short-term coordination to balance demand and supply
Decreased performance	Output performance such as annual revenues	Mainly current performance such as daily/weekly stock-out/overage costs

Practitioners and researchers have long been interested in dealing with these problems, and numerous approaches can be found in the literature. However, fundamentally all of these consist on (1) better integration of decision-making processes between different SC actors, and (2) more efficient information sharing across the value chain, rendering information technology one of the corner-stones of the next generation of supply chain management. With the goal of Industry 4.0 in mind, SCs are rapidly changing and they will be evermore characterised by consistent connectivity, self-adaptation, and autonomous and decentralised decision-making.

Supply chain management (SCM) is characterised by different levels of decision-making, each one with their particular problems and respective time-frames. Decision-making in a supply chain can occur in three different levels: strategic, tactical, and operational. *Strategic issues* include, for example, determination of the size and location of manufacturing plants or distribution centres, the selection of the right modes of transportation given various supply, production, and distribution constraints, therefore they are long-term decisions with the time-frame ranging from months to years. In turn, *tactical decisions* include production and transportation planning, as well as inventory management and optimisation. The time-frame of such decisions usually ranges form weeks to months. At an *operational level*, decisions include issues as varied as production scheduling, inventory control, vehicle routing, traffic and materials handling, and more. These decisions must be taken usually on an hourly to weekly basis [6, 31]. Therefore, the uncertainty decreases as one moves from the strategic to the operational levels. Table 1.3 presents the decision-making levels in supply chain management.

Table 1.3: Decision-making levels in supply chain management (based on Chandra et al. (2007) [15]).

Decision-making level	Timeline	Type of decision made
Strategic	5–10 years	Investment on plants and capacities. Introduction of new products. Creation of a logistics network.
Tactical	3 months–2 years	Inventory policies to use. Procurement policies to be implemented. Transportation strategies to be adopted.
Operational	Day-to-day	Scheduling of resources. Routing of raw materials and finished products. Solicitation of bids and quotations.

From a short-, mid-term view, there is only one technology development with the potential of shifting the SCM paradigm. Namely, the Internet of Things (IoT), which enables human-thing and thing-thing communications [5]. It is expected this technology to have an impact at the operational level, by providing new levels of visibility, agility and adaptability [15]. One can imagine a network of devices collecting and emitting data, which after processed can be turned into useful information about internal and external aspects of the supply chain. This network of smart-objects will have a great impact on the two fundamental SC problems previously discussed: the Bullwhip and the Ripple effects. In particular, it is expected IoT to greatly impact food supply chains, in which a multitude of actors have to coordinate and deal with perishable goods [5].

Notwithstanding, the main challenge in supply chain management continues to be the coordination of efforts and the fulfilment of goals of the various actors composing the chain. On the one hand, suppliers, manufacturers, and retailers are interested in providing customers the right product, in the right amount, at the right time, for the right price, and at the right place [13, 32, 41]. On the other hand, distributors, and transport providers in general, want to efficiently allocate goods to their resources and minimise the number of movements [28]. This strong competition between shippers (i.e. suppliers, manufacturers, and retailers) and transport providers have created a permanent state of tension between the two sectors, increasing the demand for higher levels of efficiency, quality of service, timeliness, and responsiveness across supply chains [28].

Various scientific communities have devoted attention to the management and optimisation of operations in supply chains. Undoubtedly, operations research methods are the most widely used when modelling those systems. Nevertheless, over the past years, control theory has been attracting the attention from the scientific community as a powerful method to analyse and design supply chains from a dynamical system point of view [32, 56]. However, even though transportation management is an integral part of supply chain management, these topics have either been studied independently from each other, or integrated but assuming transportation resources are readily available whenever and wherever needed. Thus, excluding by design the possibility of dealing with real-world operational issues (e.g. delivery delays due to a lack of transportation resources, transportation monitoring and reallocation, and transportation planning and re-planing), since the symbiotic relationship between the movement of goods and the movement of transportation is disregarded [46, 62]. Furthermore, research on supply chain management has focused on cost, which hampers network planning integration, for it sharpens differences between the actors, pushing them to take individualistic and local views of the value chain [62]. To fulfil the needs of all supply chain actors and achieve a higher quality of service based on the consistency of (1) on-time delivery, and (2) delivery speed, companies must manage their transportation strategies from a holistic supply chain view [21, 62]. Consequently, such an approach entails the need for new supply chain models in which transportation decisions are integrated with dynamic inventory management and service level controls.

1.3 Literature review

Supply chains were initially investigated in a fragmental fashion, with researchers and practitioners focusing on individual processes. However, modern supply chain management (SCM) requires a holistic view, focusing on performance, design and analysis of supply chains [3]. Beamon [3] distinguishes various alternative models for supply chain (SC) design and analysis based on their inputs and objective of the study, and groups them into four categories: (1) deterministic analytical models — all variables are known; (2) stochastic analytical models — at least one variable is unknown and it is assumed to follow a probabilistic distribution; (3) economic game-theoretic models; and (4) simulation models [3].

The most frequently used methods when modelling SCs in a deterministic perspective are operations research (OR) methods [30], and according to Chandra and Grabis [14], and Dolgui and Proth [20], several of those methods have been used to address various supply chain management issues, from inventory management to tactical planning decisions.

Out of the various available quantitative decision-making techniques, mathematical programming stands out as the most significant one. Its main purpose consists on finding an optimal solution for the allocation of scarce resources to different (usually competing) activities [14, 26, 58]. The most common classes of mathematical programming models are linear programming, and integer programming, whose main advantages consist on (1) providing a fairly simple and tractable approximation of complex decision-making problems, (2) finding the optimal set of decisions among a large number of alternatives, and (3) supporting analysis of decisions made. Particularly in the context of supply chains, these models are specially useful in capturing the spacial aspects of its configuration [14, 20].

However, many real-life problems cannot be represented as linear functions, and non-linear models are often required, which highlights a crucial aspect of modelling complex systems: the validity of representing real-world problems is only improved at the expense of tractability, and computational efficiency. In fact, mathematical programming presents important limitations not only regarding the aforementioned issues, but more importantly when it comes to representing the dynamic and stochastic aspects of supply chains [14, 58].

In addition to such methods and techniques, heuristics are also often used specially on scheduling and planning operations [38, 58], and although effective these methods offer no information on how good a solution is achieved, or how much the system's performance can be improved [30].

All in all, although OR methods have been subject to substantial improvements over the past decades, they present some limitations when applied to supply chain management. First, these techniques tend to oversimplify high-dimensional problems, either by reducing them to a simple dimensionality or by employing heuristics. Second, these techniques do not explicitly take into account the uncertainty arising from the interconnection of the different supply chain members [32]. Table 1.4 summarizes the advantages and limitations of OR techniques in application to the SCM.

Table 1.4: Advantages and limitations of OR techniques in application to the SCM (based on Ivanov [32]).

Advantages	**Limitations**
Guarantees optimal or admissible solution	Real high dimensionality problems either reduced to simple dimensionality, or heuristics are employed
Clarity and accessibility	Presents limited flexibility, often resorts to linear models

In light of such limitations, control theory has been attracting more attention from researchers and practitioners in the context of supply chain management over the past years, since it provides the mathematical tools to analyse, design and simulate systems based on dynamic models [32, 56].

The resemblance between SC and dynamical systems soon became apparent, and by the early 1950s classical control techniques were already being employed to address SCM problems [50, 56]. To have a comprehensive literature review goes beyond the scope of this document. Instead of a detailed and chronological account a more taxonomical approach is taken, in which the major control-theoretic contributions are classified by problem areas. For a more detailed overview of the subject, however, the interested reader is referred to Ortega and Lin [50], Sarimveis et al. [56], and Ivanov et al. [33].

According to Ivanov et al. [33], one can distinguish two major problem areas: (1) the application of dynamic feedback control to production-inventory systems, and (2) the application of optimal control to production and scheduling. To solve these problems various alternative methods were devised over the past decades, and various issues have been subject to investigation.

Dynamic feedback production-inventory control focuses on controlling and analysing operational systems making use of closed-loop models. The methodology employed consists on either transfer-function or state-space models, where it is assumed the system either deals with aggregate product levels or a single product [33, 56]. This generic representation is referred to as the inventory and order-based production control system (IOBPCS) family, after Towill [60], and was developed as an alternative framework to the "system dynamics" methodology introduced by Forrester [23] to provide sufficient analytical support and guidance to practitioners on how to improve performance [56]. Each member of the IOBPCS family integrates some or all of the following five components: a forecasting mechanism, a lead-time, an inventory and a work in progress (WIP) feedback loops, and a target stock setting [33, 56]. The IOBPCS family has been subject to both continuous and discrete analyses, using Laplace and z-transformation, respectively [19]. The study of this family of systems seeks to understand the impact the production-inventory control system has upon the overall dynamical operational performance. Various practical implementations and solutions have been developed over the years, however the properties, problems and topics subject to analysis have been always the same: stability, robustness, and resilience. A presentation of the most relevant contributions follows.

In a linear models context, Disney and Towill [18] and Lalwani et al. [40] focused on investigating and mitigating the stability issues arising from poor model designs. In particular, Lalwani et al. [40] have proposed a state-space representation methodology that assures both controllability and observability. In both cases, authors have focused on the discrete time domain, contrary to the majority of contributions. Dejonckheere et al. [17] have dealt explicitly with the variability phenomenons of order-up-to system, proving that those policies always generate a bullwhip effect. Their analyses was based on both statistical and engineering methods. In the end they advocate for the latter, proposing a general inventory replenishment rule capable of dampening the bullwhip effect.

The use of linear models have however been frequently exposed by some authors as a limitation on the study of dynamical systems such as supply chains, and advocate a dynamical systems approach to deal with the non-linear aspects of SC and operations dynamics [33]. In this respect, model predictive control (MPC) has become a popular planning technique of IOBPCS. Perea-Lopez et al. [53] presented a novel approach to supply chains based on a material and information flow perspective. They proposed an innovative modelling approach capable of reproducing realistic SC behaviours making use of a centralized control scheme consisting of heuristic rules frequently used in industry. This novel framework was further developed to accommodate an MPC structure. Perea-Lopez et al. [54] used the MPC framework to model a multi-product, multi-echelon production and distribution network. They assumed demand to be deterministic, and instead of the typical quadratic programming formulation, they used a mixed integer linear programming approach to formulate the optimal control problem. The supply chain under consideration was of considerable complexity and their study focused on a comparison between a centralized and two decentralized approaches. The first decentralized control scheme focused on optimizing distribution, while the production-inventory planning made use of heuristics. Conversely, their second decentralized approach used heuristic rules to manage distribution, while manufacturing was optimized. The authors have concluded that the centralized approach had a better performance. Seferlis and Gianellos [57] proposed an unconventional approach: a two-layered hierarchical control scheme, where a decentralized inventory control based on PID controllers was embedded within an MPC framework. The MPC problem was formulated as a linear, instead of a quadratic programming one. They considered both deterministic and stochastic demand, concluding the bullwhip effect occurred for the deterministic case, whereas for the stochastic demand case a centralized control scheme would need to have a larger prediction horizon to achieve an equivalent performance to the decentralized scheme. In turn, Braun et al. [13] also inspired by the flow perspective presented by Perea-Lopez et al. [53], developed a decentralized MPC scheme for a multi-product multi-echelon SC, capable of dealing with uncertainty resulting from model mismatch (lead times) and demand forecasting errors.

More broadly, control theory has been used to study and analyse specific SC properties, such as the impact of stochastic events, disturbances and fluctuations outside the IOBPCS family domain. Contrary to the OR methods, in which uncertainties, failures and lead times are mostly described by probability distributions and stochastic processes, robust control theory addresses these problems by considering uncertainties as unknown-but-bounded quantities. Moreover, in this framework performance specifications and physical limitations are modelled as hard constraints, meaning they must be satisfied for all realizations of the uncertain quantities [56]. In robust control theory, two types of uncertainty can be considered: exogenous disturbances and plant-model mismatch, i.e. uncertainties due to modelling errors. Regardless of the uncertainty type, however, the control problem is then formulated as a min-max problem, in which the goal is to minimise a worst-case objective function, subject to the appropriate constraints [37].

In this respect, the works of Blanchini et al. [7–9] and Boukas et al. [10–12] stand out. Blanchini et al. [7] addressed the problem of dynamic production / distribution, in which inventory levels must be kept inside prescribed bounds for all possible demands. The first step of this investigation started with the examination of necessary and sufficient conditions for stability. Taking a flow approach, it was proven it was possible to solve the problem provided the following two conditions were satisfied: (1) the controlled admissible flow must dominate the uncontrolled flow, and (2) the inventory capacities must be large enough to be able to fulfil current customer demands. Making use of such conditions, Blanchini et al. [8] proceed to devise a way of achieving the least worst-case inventory level. Several important points were noted. Namely, the devised feed-back controller was equivalent to a periodic-review, order-up-to level policy and was equal to the least inventory level. More, this least inventory level turned out to be the actual steady-state one would desire the dynamic production / inventory system to achieve. Employing a method based on linear programming an on-line method for robust optimal control was devised. This investigation was further expanded to take into account the lead times [9].

In turn, Boukas et al. [10–12] have focused on the robustness and resilience analysis of different systems, considering faults, failures, and uncertainties. According to Sarimveis et al. [56], by addressing a manufacturing system problem in which machines were subject to failure and demand was unknown-but-bounded, Boukas et al. [10] were able to provide a verification theorem that gave the sufficient conditions a feedback controller must satisfy to guarantee optimality. Even though an example was presented in which a controller was derived in closed form, it was noted that for complex schemes of control the same derivation would be practically impossible. Boukas et al. [11] proceeded to investigate yet another problem, this time a continuous-time production–inventory problem with deteriorating items which aimed to minimise a quadratic, finite-horizon production and inventory/shortage cost. A controller design method which guarantees stochastic quadratic stability for the closed-loop system, as well as a suitable performance level were presented. Making use of an example of a tracking problem where a machine prone to failure produced one deteriorating item, they have shown the tracking error converged to values close to zero. In turn, Boukas et al. [12] tackled an inventory–production system with uncertain processing time and delay in control. Their goal was to make the closed-loop system asymptotically stable while satisfying an infinite-horizon cost function, which was achieved deriving sufficient linear matrix inequality conditions for the satisfaction of input constraints for the feedback controller.

1.4 Objectives and contributions

To solve the trade-off between (i) on-time delivery, and (ii) efficient transportation management, this document sets forth a new approach for real-time supply chain management based on the model predictive control framework. The proposed methodology is based on a flow perspective and focuses on the discrete time case. It integrates ideas from operations research and control theory, resulting in interpretable, tractable and flexible dynamic models. The outlined modelling framework produces linear, time-invariant, state-space supply chain representations that are both controllable and observable. The presented approach was initially based on the work of Nabais et al. [47], and evolved into an extension of the works of Perea-Lopez et al. [53] and Braun et al. [13]. While their work focused on coordinating production and inventory activities across the network, this work generalizes and integrates both the perspectives of the shippers, and the transportation actors, enabling the proposed managing tool to be employed by any supply chain member, regardless of its role. That is, on the one hand suppliers, manufacturers and retailers are able to (1) deal with multi-products, (2) monitor and manage stocks, (3) schedule production activities, (4) monitor work in progress (WIP), and (5) define reception and dispatch time-windows. On the other hand, transport providers can (1) monitor different transportation types, (2) deal with costs associated with the different resources' capacities, and (3) monitor the location and state (i.e. with or without cargo) of the transportation resources composing the fleet.

The thesis contributions can be divided into two main areas:

1. **Supply chain modelling**

   A systematic and flexible modelling framework based on a flow perspective that encapsulates the different views in supply chains (i.e. shippers *vs* transport providers) is proposed.

2. **Operations management**

   A constrained MPC scheme capable of integrating transportation operations (e.g. monitoring and managing multiple transportation types, dealing with costs associated with resources' capacities) with manufacturing activities (e.g. monitoring and managing multi-product inventories, scheduling production activities, monitoring WIP) is presented.

Some results have already been submitted. There is also on-the-work publications, as follows.

- **Distributed Model Predictive Control For Optimal Warehouse Management**

  Submitted to Jornadas de Distribuição e Logística – JDL 2018, Setúbal. In this paper, operations management in a warehouse is modelled as a network of internal storing units and stated as a tracking control problem, in which inventory levels at the network should follow desired values over time. A distributed model predictive control (DMPC) scheme is proposed as an optimisation framework for optimal warehouse management [1].

- **Real-time, integrated transportation planning for optimal supply chain management**

  To be submitted to a journal paper. This article presents the mathematical formulation and the centralized control scheme of the model developed in this work.

Chapter 2

Methodology

The present chapter opens with the presentation of the modelling approach in Section 2.1, where the two basic flows — information, and material flows — are thoroughly discussed and analysed to understand the dynamics of supply chains. The knowledge leveraged from this initial discussion is then used to define the building-blocks of the proposed modelling approach. Then the presentation of the established notation, which is used throughout the document, will follow. The section ends with he definition of the performance indexes to be used to evaluate the system's performance. In Section 2.2, the control structure based on the MPC framework is outlined, making use of the modelling approach presented in the previous section. Both the variables' mapping from the MPC into the supply chain domain, as well as the deduction of the imposed constraints are presented. The section ends with the presentation of the optimisation problem written in the quadratic programming form.

2.1 Modelling supply chains

To provide a prompt response to variable demand, supply chains (SCs) usually work in a pull system, meaning that a given node reacts to a replenishing order placed by its succeeding node by either producing and/or replenishing it, or by transferring that order to upstream nodes if it is not possible to fulfil that order [53], producing a cascade effect. Figure 2.1 illustrates the basic flows and mechanisms present in such systems.

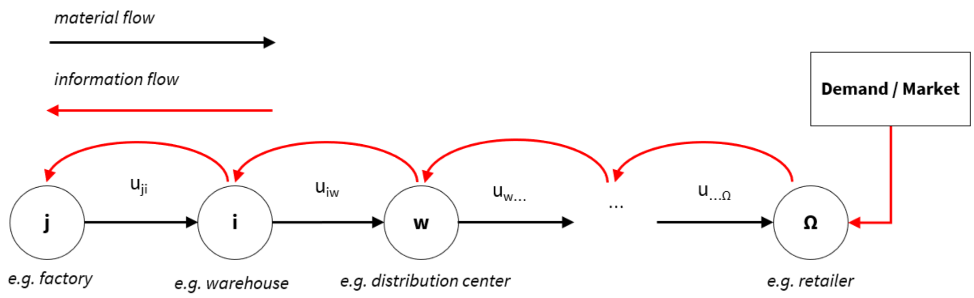


Figure 2.1: Supply chain dynamics — definition of nodes, arcs, and flows of material and information.

In SCs, actors (e.g. suppliers, manufacturers, retailers, transporters) must cooperate to move commodities from the point of origin to the point of consumption in order to meet customers' requirements. Logistics management is the part of supply chain management (SCM) which plans, implements, and controls the flow and storage of goods, services and related information [16]. Traditionally, *outbound logistics* dealt with the movement of goods from a manufacturer to its distribution partners, comprising activities such as distribution, shipping, and customer delivery and service. In turn, *inbound logistics* is related with the purchasing and transportation of raw materials to a factory or storage facility, dealing with activities such as purchasing, procurement, and raw materials inventory management [2].

Nevertheless, different SC members may have different goals. For instance, suppliers usually do not work exclusively for one particular manufacturer, and retailers often do not depend on one single manufacturer. This distinction is particularly accentuated when considering transportation as a service. In that case, the primordial goal of transport providers (TP) is not the delivery of goods to customers on time, but rather to efficiently allocate goods to their transportation resources. This differences in goals create tensions, and uncertainty arises from the interconnection between shippers and TP.

From a modelling perspective, a supply chain is described as a network of independent entities or nodes, connected by links (or arcs) representing the flows of material and information. In turn, these flows affect the contents of each node. Therefore, the dynamics of a SC can be modelled resorting to mass balances at each node. Figure 2.2 presents a three-echelon sample supply chain network made publicly available by Willems [63].

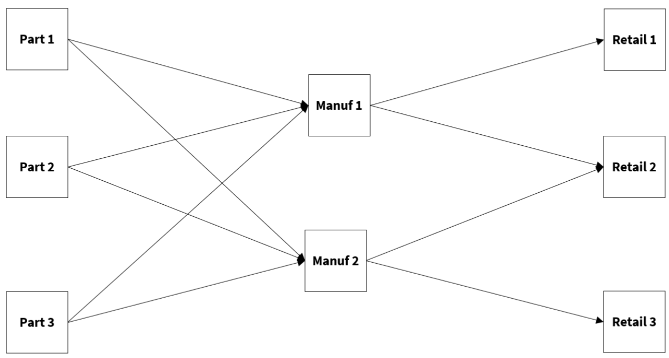


Figure 2.2: Sample supply chain network. (Based on Willems [63].)

At each node there may exist different types of material resources: (1) goods, i.e. raw materials, WIP, and finished products, and (2) transportation vehicles. Throughout the document, raw materials and/or finished products will be referred to as *goods* or *commodities* interchangeably, and transportation vehicles as *transport agents* (or simply *agents*) or *vehicles* interchangeably.

The movement of material resources across the network creates flows of material. The general assumption in literature so far, is twofold: (1) it takes the flow of goods and the flow of transportation to be one and the same thing, and (2) the material flow is assumed to be in steady-state subject to punctual disturbances (or variations) that must be mitigated. Both assumptions present some limitations. First, assuming transportation and goods are part of the same flow, one excludes the possibility of vehicles moving without cargo, which is a recurrent problem in SC. More, it implies there is only one goal in the transportation of goods. That is, it is assumed TP have the same goals as the companies they provide the service to, which is not usually the case in SCs. Secondly, the steady-state assumption is unrealistic, since it requires transportation resources to be readily available at every moment. Several points must be noted here: (i) such an assumption puts the onus on the TP, showing the parcelled approach to supply chain management problems, in which the management of goods and transportation resources are two distinct problems, which perpetuates the existing tension between shippers and transport providers; (ii) models based on the assumption that transportation is always available exclude by design both the uncertainty that arises from not having the necessary resources available and the impact it may have on the delivery of goods. For instance, a vehicle may actually be available but displaced across the network. Hence, a model which is not capable of simultaneously monitoring the location and the status (i.e. with or without cargo) of transportation resources disregards the symbiotic relationship between movement of goods and movement of transport agents. In a nutshell, the assumptions found in the literature so far are critical and result in models with limited capacity to tackle real-world operational issues in an integrated way.

Therefore, one can distinguish between two different and independent material flows: (1) the flow of commodities, and (2) the flow of transport agents. Furthermore, two different groups of node contents (i.e. inventories) can be discriminated: (1) inventory of commodities (e.g. stocked raw materials and/or products awaiting to be shipped), and (2) inventory of transport agents (e.g. free, parked vehicles awaiting for shipment assignment). Both flows and inventories are subject to particular levels of availability and capacity. Namely, flows depend on the availability of resources and the capacity of links, whereas inventories are bounded by a minimum and maximum node capacity. Moreover, flows may be *controllable* by a given SC member (e.g. reception/shipment of commodities from/to suppliers/customers within the standard SC, movement of transportation agents across the network) or *uncontrollable* (e.g. acquisition of new vehicles, unavailability of a given agent due to maintenance, reception/shipment of commodities from/to suppliers/customers outside the standard supply chain, end-customer acquisition of finished goods). Hence, one can discriminate two fundamental layers: (1) the commodity layer, and (2) the transportation layer. Each layer consists of (possibly different) independent networks across which matter is allowed to flow. Thus, the SC dynamics can be considered as a superposition of these two fundamental layers, in which the movement of a commodity is interpreted as a synchronised, superimposed flow of material in both the commodity and the transportation layers, as shown in Figure 2.3.

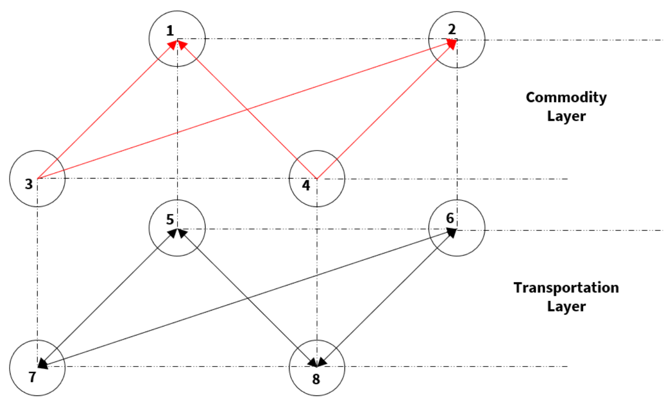


Figure 2.3: Fundamental layers in a SC. The movement of a commodity from node 3 to node 1, implies the synchronised flow of material from nodes 3 and 7 to nodes 1 and 5, respectively.

Generally, however, each commodity poses particular transportation requirements with respect to its weight, physical state, heating, packaging, etc. The combination of these characteristics specifies a certain commodity type. Analogously, a transportation type may be characterised by different features as well, such as speed, loading capacity, transportation cost, authorized areas of operation, etc. Consequently, a generic supply chain is given as a collection of stacked networks (or layers) of different commodity and transportation types. It should be stressed that there can be multiple commodities and multiple transportation types, and their number can differ, i.e. it is not required to have the same amount of commodity and transportation types. Figure 2.4 illustrates a generic supply chain consisting on stacked commodity and transportation networks.

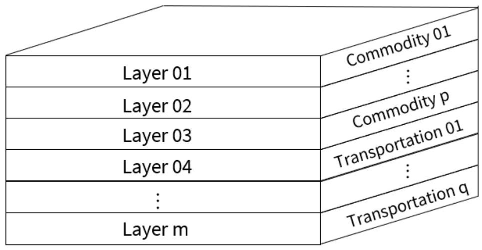


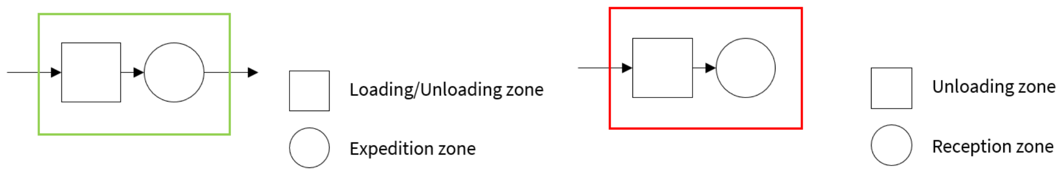
Figure 2.4: Generic supply chain stack of p commodities and q transportation networks.

From an operations management point of view, there are a variety of operations one can find in any supply chain. Namely, one or more SC members may focus on the procurement of raw material, others in processing and transforming those materials into finished goods, others, still, may focus only on storing and/or distributing commodities. It is therefore expected that nodes with different functions within a supply chain should present different mechanisms of mapping inputs onto outputs.

Taking a holistic view of the SC, two fundamental types of node can be discerned. Namely, *source nodes*, and *sink nodes*. Sink nodes are defined as the most down-stream members of the network, whereas the rest are referred to as source nodes. The difference in nomenclature is important because their internal mechanisms are different. Source nodes receive (and/or transform) commodities that will be dispatched and flow through the network until they arrive at the sink nodes, which form the last echelon in a supply chain, usually representing the retailing level. Therefore, sink nodes receive and stock commodities. Since these two types of node are fundamentally different, the names given to their internal zones should also differ.

Source nodes: are composed of a *loading / unloading zone*, followed by an *expedition zone*. The loading / unloading zone refers to the areas where commodities and transportation agents are stored awaiting for assignment, either because cargo has arrived or because it is waiting to be shipped. Once an order is placed, if the required resources are available, assignment takes place and cargo is loaded. Then, the loaded transport goes into the expedition zone, from where it can take different paths to the down-stream nodes. Figure 2.5(a) presents a schematic representation of such nodes. Source nodes may be further divided into transformation nodes.

Sink nodes: are composed of an *unloading zone*, followed by a *reception zone*. The unloading zone is the area where commodities and transportation agents are decoupled. Transport agents must stop and wait for further instructions, while commodities proceed to the reception zone, a restricted area where commodities are stocked for the final customer to pick them up. Figure 2.5(b) presents a schematic representation of such nodes.



(a) Source node. (b) Sink node.

Figure 2.5: Schematic representation of source and sink nodes.

Transformation nodes: are a special type of source nodes that represent a production or assembly process, where the output is given as a combination of the inputs. Figure 2.6 presents a schematic representation of such nodes.

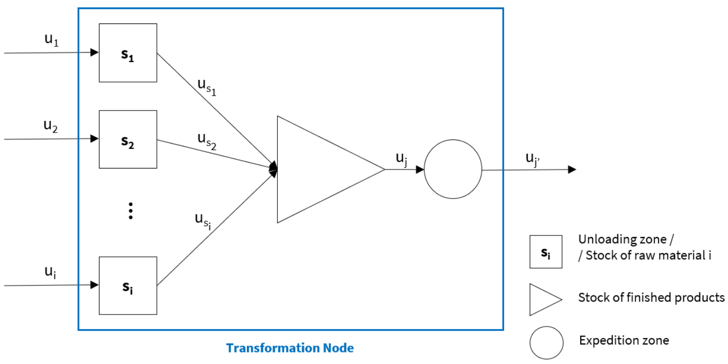


Figure 2.6: Transformation nodes.

Transshipment nodes: may not represent precise geographical locations, but rather virtual ones, in order to give information about: (1) the amount and whereabouts of in-transit commodities, and (2) the amount, whereabouts and status (i.e. if agents are *occupied* moving cargo, or *free* going towards a new location) of moving transport agents. An illustration of such nodes can be found in Figure 2.8.

Table 2.1 summarizes the previous definitions, presenting the expected contents of each node area at any given time, when applied.

Table 2.1: Nodes topology and their expected contents by area (when applicable).

Node Types	Loading/Unloading	Expedition/Reception	Stock of finished prod.
Source	commodities waiting to be loaded free transport agents	cargo loaded and prepared to ship occupied transport agents	— —
Sink	unloaded commodities free transport agents	stock of commodities —	— —
Transformation	commodities waiting to be loaded/processed free transport agents	cargo loaded and prepared to ship occupied transport agents	finished commodities —
Transshipment	in-transit commodities in-transit occupied transport agents		

In turn, one can describe the transportation process by a sequence of five actions: (1) wait assignment; (2) reallocate, if needed; (3) load cargo; (4) move cargo, (5) unload goods. Figure 2.7 presents a summary of the transportation process.

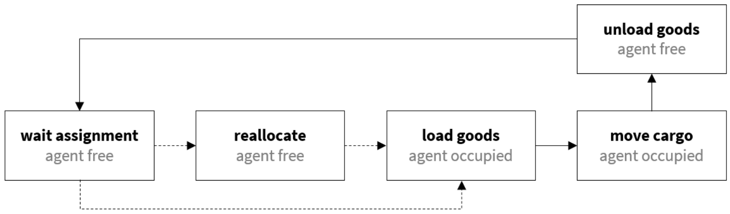
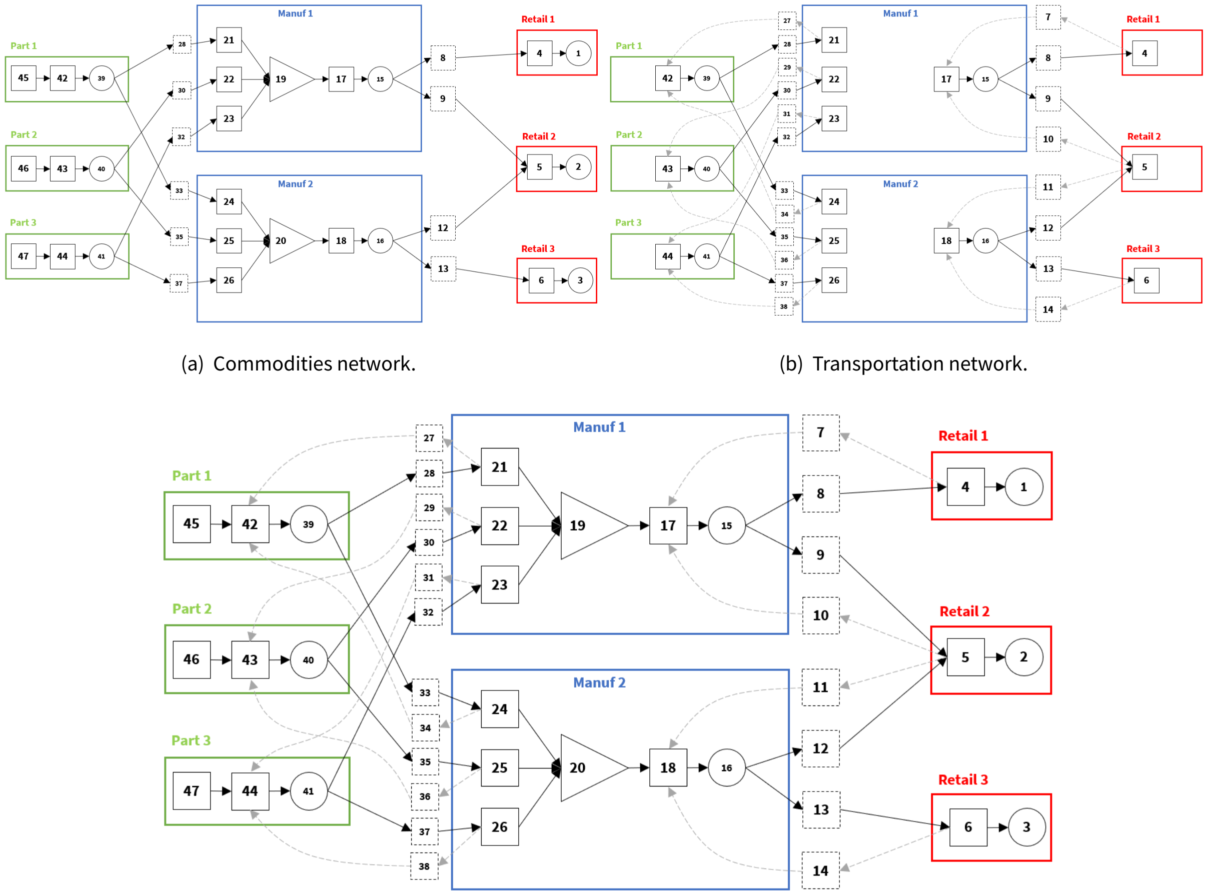


Figure 2.7: Transportation process.

Finally, Figure 2.8 presents the model of the SC of Figure 2.2, employing the presented modelling approach.



(a) Commodities network.

(b) Transportation network.

(c) Supply chain model, superimposing the commodities and the transportation networks.

Figure 2.8: Model. Dashed arrows, flow of transport agents only. Solid arrows, commodities and/or transport agents. Dashed squares, transshipment nodes.

2.1.1 Notation

As previously stated, the dynamics of supply chain may be modelled resorting to mass balances at each node. Since the flows consist of commodities and/or transport agents, for an easy and clear exposition commodities and transport agents will be referred to as *resources*, distinguishing between the two whenever necessary. Consider Figure 2.9 as a starting point.

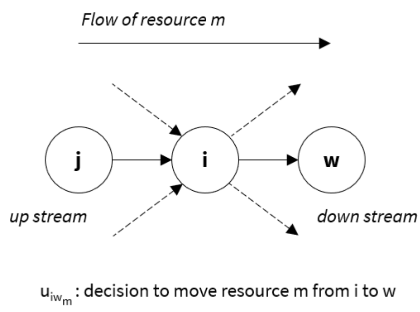


Figure 2.9: Transformation nodes.

Generically, the inventory of a given resource m at node i can be computed as follows:

$$x_{i_m}(k+1) = x_{i_m}(k) + \Delta x_{i_m}(k) \tag{2.1}$$
$$= x_{i_m}(k) + \sum_j u_{ji_m}(k) - \sum_w u_{iw_m}(k) + d_{i_m}(k), \tag{2.2}$$

where x_{i_m} is the inventory level of resource m at node i; Δx_{i_m} is the variation of inventory level of resource m at node i; u_{ji_m} is the incoming stream of a given resource m to be processed at node i coming from node j; u_{iw_m} is the outgoing stream of a given resource m from node i to be processed at node w; d_{i_m} represents an exogenous input into node i regarding resource m (e.g. acquisition of new transport agents, unavailability of a given agent due to maintenance, reception/shipment of commodities from/to suppliers/customers outside the standard supply chain, end-customer acquisition of finished goods); and k is the discrete-time base period, also referred to as "sampling time instant", which depends on the dynamic characteristics of the network, i.e. is dependent on the application.

Transformation nodes belong to a special type of node which involve more manipulations. As shown in Figure 2.6, transformation nodes are composed of various sub-nodes, or zones: unloading zone/stock of raw materials, stock of finished products, and expedition zone. Transformation nodes define their outputs as a combination of the inputs. That is, whenever the necessary raw materials required to produce one unit of finished product are in place, these are simultaneously processed and transformed into the final product. This transformation takes place whenever material flows from the stock of raw materials nodes into the stock of finished products one. Thus, the stream of raw materials works as a "transformation switch".

From a mathematical point of view, the stock of transformed (or finished) products can be modelled as follows:

$$x(k+1) = x(k) + Q_B\, C_G \sum_i u_{s_i} - u_j, \tag{2.3}$$

where Q_B is the batch quantity, and C_G is the compatibility gain, defined as $C_G = \frac{1}{N}$, where N is the total amount of different raw materials required to produce one unit of finished product.

In turn, the inventory of raw materials can be defined as follows:

$$s_i(k+1) = s_i(k) - Q_B\, M_i\, u_{s_i} + u_i, \tag{2.4}$$

where Q_B is the same batch quantity, and M_i is the necessary quantity of raw material i to produce a transformed product.

Denoting by n the number of nodes, n_m the number of resources, n_u the number of links between adjacent nodes, and n_z the number of output nodes, one can represent the SC model by making use of a state-space dynamic model representation as follows,

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) \ + \ \mathbf{B}\mathbf{u}(k) \ + \ \mathbf{B}_d\mathbf{d}(k), \\ \mathbf{y}(k) &= \mathbf{C}_y\mathbf{x}(k), \\ \mathbf{z}(k) &= \mathbf{C}_z\mathbf{x}(k), \end{aligned} \tag{2.5}$$

where

$$\mathbf{x}(k) = \ [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{n_m}]^T \in \mathbb{R}^{[n \times n_m] \times 1} \text{ is the state vector,} \tag{2.6}$$

$$\mathbf{u}(k) = \ [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{n_m}]^T \in \mathbb{R}^{[n_u \times n_m] \times 1} \text{ is the input vector,} \tag{2.7}$$

$$\mathbf{d}(k) = \ [\mathbf{d}_1, \mathbf{d}_2, \cdots, \mathbf{d}_{n_m}]^T \in \mathbb{R}^{[n \times n_m] \times 1} \text{ is the exogenous input vector,} \tag{2.8}$$

$$\mathbf{y}(k) = \ [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_{n_m}]^T \in \mathbb{R}^{[n \times n_m] \times 1} \text{ is the vector of measured outputs,} \tag{2.9}$$

$$\mathbf{z}(k) = \ [\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_{n_m}]^T \in \mathbb{R}^{n_z \times 1} \text{ is the vector of outputs which are to be controlled.} \tag{2.10}$$

$\mathbf{A}, \mathbf{B}, \mathbf{B}_d, \mathbf{C}_y$, and $\mathbf{C}_z$ are matrices of appropriate size given as follows,

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_{n_m} \end{bmatrix}, \mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B}_{n_m} \end{bmatrix}, \mathbf{B}_d = \begin{bmatrix} \mathbf{B}_{d_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{d_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B}_{d_{n_m}} \end{bmatrix}, \tag{2.11}$$

$\mathbf{C}_y = \mathbf{I}$, and $\mathbf{C}_z$ is defined by the user.

Table 2.2 presents a brief summary of the developed notation.

Table 2.2: Glossary.

x_{i_m}	inventory level of node i, of resource m
$\mathbf{x}$	collection of inventory levels per node, per resource
u_{ij_m}	link from node i to node j, of resource m
$\mathbf{u}$	collection of nodes' connections per node, per resource
d_{i_m}	exogenous input into node i regarding resource m
$\mathbf{d}$	collection of exogenous inputs per node, per resource

The interested reader is referred to Appendix A, where the proofs that the proposed approach results in linear, time-invariant, controllable and observable models take place.

2.1.2 Performance indexes

Since one may be interested in gaining a panoramic and integrated overview of the supply chain as whole, two performance indexes are proposed. Namely, *Service rate*, and *Fleet usage rate*.

Service rate: service (or fill) rate measures the number of units filled as a percentage of the total ordered, and can be defined as follows,

$$SR = \left[1 - \frac{max(D) - N}{max(D)}\right] \times 100, \quad \text{for } D > 0, \tag{2.12}$$

where, D is the demand of a given product type, and N is the number filled units of a given product type.

Fleet usage rate: fleet usage rate measures the number of transportation resources that are in use as a percentage of the total transportation resources, and can be defined as follows,

$$FUR = \left[1 - \frac{A - O(k)}{A}\right] \times 100, \quad \text{for } A > 0, \tag{2.13}$$

where, A is the existing transportation resources (or agents), and O the number of occupied agents at time k.

2.2 Controlling supply chains

Model Predictive Control (MPC) has become an important framework for controlling complex, dynamic systems. Over the last decades, MPC has proven to be successful in the process industry [13, 33], and its growing popularity on supply chain management applications is rooted in the relative ease with which it can be understood, and its ability to handle constraints [13]. Figure 2.10 presents the basic features and mechanisms of a model predictive control.

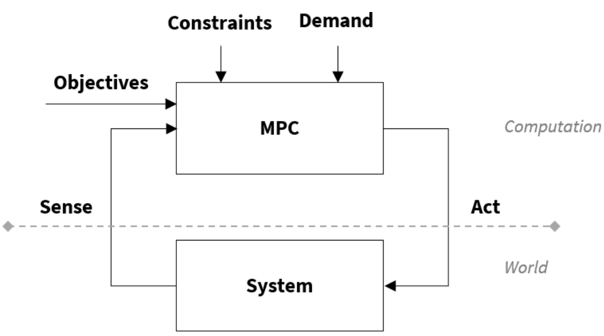


Figure 2.10: Basic features and mechanisms of MPC. (Based on Nabais et al. [47].)

MPC is a control strategy that produces a sequence of control actions based on the predicted behaviour of the system. The control actions are chosen by repeatedly minimising (or maximising) a performance index at each time-step. Since the controller at each sampling time predicts a control sequence into the future over some horizon (prediction horizon, H_p) but implements only the first one, it results in what is called the *receding horizon window*: the perception that the predicted horizon is always receding away from the present. Thus, any MPC problem considers three main components: a prediction model, a performance index (or cost function) and a set of constraints.

Regarding the cost function, the problem at hand is formulated as a modified standard reference tracking problem, in which the controller must minimise the tracking error between a given reference, $\mathbf{r}$, and the system output, $\tilde{\mathbf{z}}$, over a given prediction horizon, H_p. The cost function to be minimised is defined as follows,

$$J = \sum_{i=1}^{Hp} [\mathbf{r}(k+i) \; - \; \tilde{\mathbf{z}}(k+i)]^2_{\mathbf{Q}} \; + \; [\mathbf{u}(k+i-1)]^2_{\mathbf{R}} \, , \tag{2.14}$$

where $\mathbf{u}$ is the collection of predicted decision variables that minimise J, and $\mathbf{Q}$ and $\mathbf{R}$ are weighing parameters of appropriate dimensions.

Table 2.3 presents the variables' mapping from the MPC into the supply chain domain.

Table 2.3: Variables mapping from MPC to supply chain domains.

MPC	**Supply Chain**
references $\mathbf{r}$	inventory level targets per node
predicted outputs $\tilde{\mathbf{z}}$	future inventory levels per node (individual and/or aggregate levels)
inputs $\mathbf{u}$	"nodes' connections, or links", decision to move/allocate resource
weighing matrix $\mathbf{Q}$	cost of not achieving a prescribed stock level
weighing matrix $\mathbf{R}$	transportation costs

In the previous section the concept of stacked resource layers was presented, in which each layer consisted on the network a given resource was able to move across. It was yet mentioned that the movement of a particular commodity is interpreted as a synchronised, superimposed flow of material in both the commodity's layer, and its respective transport agent's one. From the control perspective, the flow of material is the decision variable to be optimized. Therefore, for the controller to be able to produce a synchronised, superimposed flow of material the following constraint should be imposed. Let $u_{ij_{m_P}}$ and $u_{ij_{m_T}}$ represent the decisions to move a product, P, and a transportation type, T, respectively. Then, the superimposed flow of material (i.e. products, plus transportation) can be written as follows,

$$\sum_{m_P} u_{ij_{m_P}}(k) \leq u_{ij_{m_T}}(k). \tag{2.15}$$

Notwithstanding, an important feature to consider when distinguishing between transportation types is their load capacity, as previously noted. Therefore, to effectively model transportation, the previous constraint must be further extended, yielding,

$$\sum_{m_P} u_{ij_{m_P}}(k) \leq \lambda u_{ij_{m_T}}(k), \tag{2.16}$$

where $\lambda > 0$, and represents the maximum load capacity.

In turn, transformation nodes also require some constraints. Recall that in the previous section it was stated that whenever the necessary raw materials required to produce one unit of finished product are in place, they are then *simultaneously* processed and transformed into the final product. The transformation takes place when the material flows from the stock of raw materials nodes into the stock of finished products node, making the stream of raw materials work as a "transformation switch". To do so, control actions, u_{s_i} (see Figure 2.6) must assume only binary values. Further, to accomplish the "switch-like" mechanism all control actions must assume the same value (either 0 or 1) at the same time. To translate this constraint into a formal mathematical way, consider the following.

Let $\mathcal{S}$ be the set of all inputs to a given stock of finished products node. Let $u_{s_i}(k)$ denote the i^{th} element in $\mathcal{S}$, and $u_{s'_i}(k)$ the i^{th} element in $\mathcal{S} \setminus \{u_{s_i}(k)\}$, at any given time-instant k. Thus, two elements, $u_{s_i}(k)$ and $u_{s'_i}(k)$, are said to be equal *iff* the following condition holds,

$$|u_{s_i}(k) - u_{s'_i}(k)| \leq 0. \tag{2.17}$$

Forcing all elements in $\mathcal{S}$ to simultaneously assume the same value, would imply a set of combinatorial conditions of the form of equation (2.17). Since the number of necessary combinations are mathematically expressed by C_2^N, where N denotes the total number of elements in $\mathcal{S}$, the following expression will be used as a shorthand to represent them,

$$|u_{s_i}(k) - u_{s'_i}(k)|_{C_2^N} \leq 0. \tag{2.18}$$

Conversely, to prevent transformation nodes of producing two commodities, m and n, in parallel, one can define the following constraint,

$$u_{s_{i_m}}(k) + u_{s_{i_n}}(k) \leq 0. \tag{2.19}$$

where $u_{s_{i_m}}$ and $u_{s_{i_n}}$ are the flows of raw material i of commodities m and n, respectively. Generalising such a constraint would imply a set of conditions of the form of equation (2.19). Namely, the number of necessary conditions is equal to the number of commodities that must be produced in sequence minus one. For brevity's sake, the following expression will be used as a shorthand to represent them,

$$\left[u_{s_{i_m}}(k) + u_{s_{i_n}}(k)\right]\Big|_{\forall\, m,n} \leq 0. \tag{2.20}$$

To guarantee that a given resource is at a given node at the time of pulling, one could define the following constraint:

$$\sum_i u_{ij_m}(k) \leq x_{i_m}(k). \tag{2.21}$$

However, since different nodes may have different processing mechanisms, their processing-times may also vary. From a modelling perspective, the processing-time can be considered as a pure time-delay. Thus, one can rewrite equation (2.21) as follows,

$$\sum_i u_{ij_m}(k) \leq x_{i_m}(k - \tau_i), \tag{2.22}$$

where τ_i is the time-delay produced by the processing-time of node i.

Equation (2.22) makes explicit use of past information, namely the past inventories $x_{i_m}(k - \tau_i)$, which means the controller must have sufficient internal memory to save the necessary information. The most efficient way of encapsulating the required information is to make a dynamical model that updates constraints each time the controller is updated with the system's state. Figure 2.11 illustrates the control scheme.

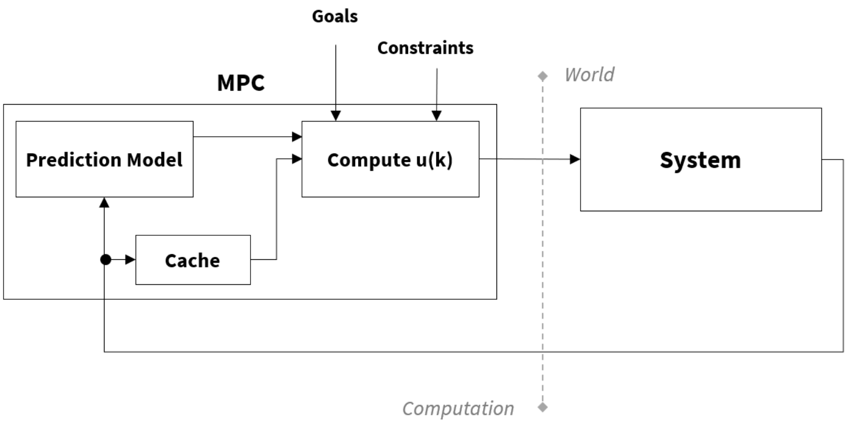


Figure 2.11: MPC scheme.

Expanding equation (2.22) yields,

$$
\begin{aligned}
\sum_i u_{ij_m}(k) &\leq \; x_{i_m}(k-\tau_i), \\
\sum_i u_{ij_m}(k+1) &\leq x_{i_m}(k-\tau_i+1), \\
&\vdots \\
\sum_i u_{ij_m}(k+\tau_i) &\leq x_{i_m}(k).
\end{aligned}
$$

Noticing that as k increases, $x_{i_m}(k)$ "moves upwards", it is possible to build a state-space model for a system of n nodes and n_m commodities, as follows,

$$\mathbf{p}(k+1) = \mathbf{A}_\tau \mathbf{p}(k) + \mathbf{B}_\tau \Delta \mathbf{y}(k) + \boldsymbol{\Omega} \mathbf{B}_{u_\tau} \mathbf{u}(k-1), \tag{2.23}$$

$$\mathbf{P}(k) = \boldsymbol{\Omega}_\tau^T \mathbf{p}(k) \tag{2.24}$$

where

$$\mathbf{A}_\tau = diag\left(\mathbf{A}_{\tau_1},\, \mathbf{A}_{\tau_2},\, \cdots,\, \mathbf{A}_{\tau_{n_m}}\right) \in \mathbb{R}^{[n \times n_m(\tau_i+1)] \times [n \times n_m(\tau_i+1)]}, \tag{2.25}$$

$$\mathbf{B}_\tau = diag\left(\mathbf{B}_{\tau_1},\, \mathbf{B}_{\tau_2},\, \cdots,\, \mathbf{B}_{n_m}\right) \in \mathbb{R}^{[n \times n_m(\tau_i+1)] \times n_m}, \tag{2.26}$$

$$\boldsymbol{\Omega} = diag\left(\boldsymbol{\Omega}_1,\, \boldsymbol{\Omega}_2,\, \cdots,\, \boldsymbol{\Omega}_{n_m}\right) \in \mathbb{R}^{[n \times n_m(\tau_i+1)] \times n_m}, \tag{2.27}$$

$$\mathbf{B}_{u_\tau} = f\left(\mathbf{B}_u\right) \in \mathbb{R}^{n_m \times n_m}, \tag{2.28}$$

where, $\mathbf{B}_{u_\tau}$ is given as a function of the control matrix, $\mathbf{B}_u$, of the system to be controlled. More accurately, $f(\cdot)$ was meant to represent two operations: 1) substitute every positive element by zero, and 2) delete null rows.

$$\mathbf{A}_{\tau_{\mathbf{n_m}}} = \begin{bmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{n(\tau_i+1)\times n(\tau_i+1)}, \tag{2.29}$$

$$\mathbf{B}_{\tau_{\mathbf{n_m}}} = [0,\,\cdots,\,0,\,1]^T \in \mathbb{R}^{n(\tau_i+1)\times 1}, \tag{2.30}$$

$$\mathbf{\Omega}_{\mathbf{n_m}} = [1,\,0,\,\cdots,\,0]^T \in \mathbb{R}^{n(\tau_i+1)\times 1}, \tag{2.31}$$

$$\mathbf{p}(k) = [\mathbf{x}_1(k-\tau),\cdots,\,\mathbf{x}_{n_m}(k-\tau)]^T \in \mathbb{R}^{[n\times n_m(\tau_i+1)]\times 1}, \tag{2.32}$$

$$\mathbf{x}_{n_m}(k-\tau) = [x_{1_m}(k-1),\cdots,\,x_{i_m}(k-\tau_i)]^T \in \mathbb{R}^{n(\tau_i+1)\times 1}, \tag{2.33}$$

$$\Delta\mathbf{y}(k) = \mathbf{y}(k) - \mathbf{y}(k-1) \in \mathbb{R}^{[n\times n_m(\tau_i+1)]\times 1}, \tag{2.34}$$

Note that at any given time the controller must be updated with $\mathbf{t}(k)$, $\Delta\,\mathbf{y}\,(k)$, and $\mathbf{u}(k-1)$. The last two are trivial, since they depend only on already computed variables. On the other hand, $\mathbf{t}(k)$ must be updated at every time-step using the following relationship,

$$\mathbf{p}(k) = \mathbf{A}_\tau \mathbf{p}(k-1) + \mathbf{B}_\tau \Delta\mathbf{y}(k). \tag{2.35}$$

The MPC optimisation problem can then be written as follows:

$$\min_{\tilde{u}} J = \sum_{i=1}^{Hp} [\mathbf{r}(k+i)\; - \;\tilde{\mathbf{z}}(k+i)]^2_{\mathbf{Q}} \; + \; [\mathbf{u}(k+i-1)]^2_{\mathbf{R}}\,, \tag{2.36}$$

$$s.t. \quad \mathbf{x}(k) \geq 0, \tag{2.37}$$

$$\underline{\mathbf{z}} \; \leq \mathbf{z}(k) \; \leq \; \overline{\mathbf{z}}, \tag{2.38}$$

$$\underline{\mathbf{u}} \; \leq \mathbf{u}(k) \; \leq \; \overline{\mathbf{u}}, \tag{2.39}$$

$$\sum_i u_{ij_m}(k) \leq x_{i_m}(k-\tau_i), \tag{2.40}$$

$$\sum_{m_P} u_{ij_{m_P}}(k) \leq \lambda\, u_{ij_{m_T}}(k), \tag{2.41}$$

$$|u_{s_i}(k) - u_{s'_i}(k)|_{C_2^N} \leq 0 \tag{2.42}$$

$$\left[u_{s_{i_m}}(k) + u_{s_{i_n}}(k)\right]\Big|_{\forall\; m,n} \leq 0. \tag{2.43}$$

Constraints (2.37) – (2.43) impose the network's structural features. Equation (2.37) assures all states are positive at every time-step. Equation (2.38) imposes the minimum and maximum nodes' capacities, denoted by $\underline{\mathbf{z}}$ and $\overline{\mathbf{z}}$, respectively. Equation (2.39) limits the material flow between nodes, imposing minimum and maximum admissible values as well, denoted by $\underline{\mathbf{u}}$ and $\overline{\mathbf{u}}$, respectively. Equation (2.40) guarantees the pulled resource is available at the node at the time of pulling, where τ represents the time node i requires to process resource m. Equation (2.41) assures the maximum transport loading capacity, λ, is respected. Equation (2.42) assures the raw materials flow in the right quantity, and at the right time into the transformation nodes. Equation (2.43) assures the raw materials of different commodities flow in the right quantity, and at the right time into the transformation nodes, forcing the controller to schedule manufacturing activities. It should however be stressed that only constraints (2.40) – (2.43) depend on the network's configuration.

To solve the MPC minimisation problem resorting to quadratic programming algorithms, it suffices to rewrite equations (2.36) –(2.42) in the form:

$$J = \frac{1}{2}\,\mathbf{u}^T\mathbf{H}\mathbf{u} \,+\, \mathbf{c}^T\mathbf{u}, \tag{2.44}$$

$$\mathbf{M}\,\mathbf{u} \,\leq\, \mathbf{\Lambda}, \tag{2.45}$$

where

$$\mathbf{H} = 2\left[\mathbf{R}_u^T \mathcal{Q}_Q\,\mathbf{R}_u \,+\, \mathcal{Q}_R\right], \tag{2.46}$$

$$\mathbf{c}^T = 2\left[\mathbf{R}_u^T \mathcal{Q}_Q\left(\mathbf{R}_x\mathbf{x} \,+\, \mathbf{R}_d\mathcal{D} \,-\, \mathcal{R}\right)\right]^T, \tag{2.47}$$

$$\mathbf{M} = \begin{bmatrix} \mathbf{S}\,\mathbf{R}_{u_y} \\ \mathbf{G}\,\mathbf{R}_{\mathbf{u}} \\ \mathbf{F} \\ \mathbf{F}_{\mathbf{p}} \\ \mathbf{F}_{\mathbf{a}} \\ \mathbf{F}_{\mathbf{t}} \end{bmatrix}, \tag{2.48}$$

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{s} - \mathbf{S}\,\left(\mathbf{R}_{x_y}\mathbf{x} + \mathbf{R}_{d_y}\mathcal{D}\right) \\ \mathbf{g} \,-\, \mathbf{G}\,\left(\mathbf{R}_x\mathbf{x} + \mathbf{R}_d\mathcal{D}\right) \\ \mathbf{f} \\ \mathcal{P}(k) \\ \mathbf{f}_{\mathbf{a}} \\ \mathbf{f}_{\mathbf{t}} \end{bmatrix}. \tag{2.49}$$

Deduction of equations (2.44) — (2.49) are presented in Appendix B.

Chapter 3

Simulation experiment

To validate the proposed approach a case-study based on a real-world supply chain is devised. In Section 3.1 a comprehensive description of the problem is outlined, followed by the presentation of the proposed model and its computational implementation in Section 3.2. The chapter concludes with the exhibition of the achieved results in Section 3.3.

3.1 Problem description

The supply chain under study is presented in Figure 3.1 and is based on a data set made publicly available by Willems [63] comprising 38 real-world supply chains which have been implemented in practice by either company analysts or consultants. The chosen network is a three-echelon vertical integrated chain dedicated to the production of three types of product (P1, P2, and P3). To produce each product-type, manufacturing sites (Manuf 1, and Manuf 2) require three different raw materials (RM1, RM2, and RM3) which are then mixed in various proportions to produce the finished goods. Further, these must be delivered to three different retailers (Retail 1, Retail 2, and Retail 3), each one requiring specific daily amounts of finished products. Table 3.1 presents the total lead time (in hours) for each commodity. Tables 3.2 and 3.3 present the bill of materials and the average daily demand of each retailer, respectively.

Table 3.1: Total lead time, in hours.

	Retail 1	**Retail 2**		**Retail 3**
		via Manuf 1	**via Manuf 2**	
RM1	2 + 2	2 + 3	3 + 4	3 + 2
RM2	2 + 2	2 + 3	2 + 4	2 + 2
RM3	3 + 2	3 + 3	2 + 4	2 + 2
P1	5	—	—	—
P2	—	6	6	—
P3	—	—	—	5

Differences in hardware, available space, and product-mix require different production schemes for each

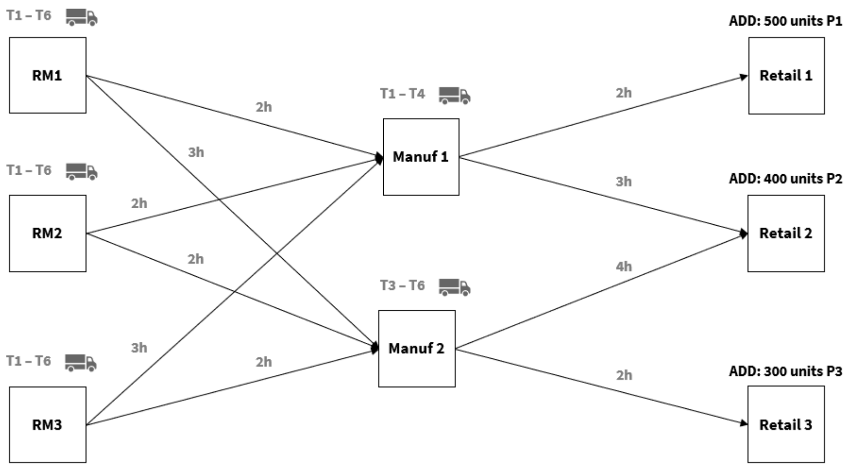


Figure 3.1: Supply chain configuration. (Based on Willems [63].)

Table 3.2: Bill of materials — units of raw material per unit of finished product.

	RM1	**RM2**	**RM3**
P1	2	1	1
P2	1	1	1
P3	1	1	2

Table 3.3: Average daily demand.

	P1 (units)	**P2 (units)**	**P3 (units)**
Retail 1	500	—	—
Retail 2	—	400	—
Retail 3	—	—	300

manufacturing site. While Manuf 1 produces P1 and P2, Manuf 2 focuses on producing P2 and P3. Each product type requires different processing times, resulting in heterogeneous production rates and throughput times. Namely, a batch of 500 units of P1 and a batch of 300 units of P3 take 1 hour to be produced, and after that time commodities of such types can readily be stored. However, compared to P1 and P3, products of type P2 require one extra hour to be produced before storage. A batch of 200 units of P2 is accomplished in 1 hour, thus the processing time of each batch equals 2 hours. Moreover, both manufacturing sites work in a flow shop scheme, meaning that it is not possible to produce two product types in parallel. Table 3.4 presents a summary of manufacturing information.

Table 3.4: Manufacturing information.

	Production rate			**Processing time**		
	P1 (units/h)	**P2 (units/h)**	**P3 (units/h)**	**P1 (h/batch[1])**	**P2 (h/batch[1])**	**P3 (h/batch[1])**
Manuf 1	500	200	—	1	2	—
Manuf 2	—	200	300	—	2	1

[1] One batch equals the amount of produced products in one hour.

In terms of raw material supply, it is assumed that they are always and immediately available to be shipped whenever necessary.

Regarding transportation, each commodity can be moved by two different modes, where different transportation types differ only on load capacity. Table 3.5 presents the specifications of each transportation type, as well as the type of commodity each mode can be assigned to. In turn, the average trip duration is shown in Figure 3.1 and Table 3.1.

Table 3.5: Maximum load capacity by transportation mode and commodity type.

	RM1 (units)	**RM2 (units)**	**RM3 (units)**	**P1 (units)**	**P2 (units)**	**P3 (units)**
T1	250	250	250	250	—	—
T2	500	500	500	500	—	—
T3	100	100	100	—	100	—
T4	200	200	200	—	200	—
T5	100	100	100	—	—	100
T6	200	200	200	—	—	200

Finally, each supply chain member defines a working-day as a 12 hour period (from 8h to 20h), meaning that there is no processing, nor transportation, of commodities outside this time window. Moreover, the demand nodes (Retail 1, Retail 2, Retail 3) have specific time windows of cargo reception. Namely, these members are able to receive goods only at two particular times per day: from 8h to 10h, or from 16h to 19h.

The problem to be solved consists in monitoring and control (1) the transportation of goods from source to demand nodes, and (2) the position and status of the available transport agents over time, while assuring that the delivery is made on time.

3.2 Implementation, computational properties, and initial set up

To solve the case described in the previous section, the model presented in Figure 2.8 (reproduced again in Figure 3.2 for convenience) was implemented using MATLAB R2016a on a machine with the following specifications: Intel Core i7-4790 3.60GHz, 8GB DDR4. For optimisation purposes, it was used the OPTI Toolbox's SCIP solver.

Two simulation tests are performed (*Simulation 1* and *Simulation 2*). The difference between simulations consists in the weights attributed to the different transportation modes, as will be discussed below. For clear illustrations, each simulation assumes the SC to be empty of commodities at starting time. Transportation wise, Table 3.6 presents the fleet disposition at the starting time.

Both simulations represent a 48h period, in which each time-instant represents 1 hour. The main objective is the same in both tests: to deliver the right amount of finished goods (500 units of P1, 400 units of P2, and 300 units of P3), at the right place, and at the right time, while minimising transportation and inventory costs. Namely, the desired delivery time is at 18h (of the first day). However, due to processing times in upstream sites, it is still acceptable to receive the goods at the next window of opportunity, that is the following pre-established reception time. It is further assumed that the maximum capacity of each node is much larger than the amounts

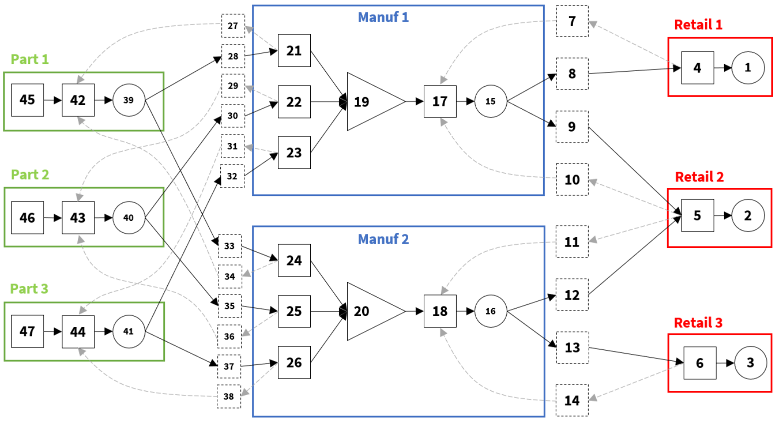


Figure 3.2: Supply Chain of the case study.

Table 3.6: Fleet disposition at the starting time.

	Node 42 (units)	Node 43 (units)	Node 44 (units)	Node 17 (units)	Node 18 (units)
T1	2	2	3	3	—
T2	1	1	1	1	—
T3	2	2	3	4	4
T4	1	1	1	1	1
T5	2	2	3	—	3
T6	1	1	1	—	1

of commodities being transported. Besides, at each time-step it can be transported as many commodities as necessary, the only limitation being the availability of transportation.

To solve this problem, the MPC controller is set to have a prediction horizon, H_p, equal to 20 time-instants. That is, at each time-step the controller plans the current control action considering the next 20 hours. Moreover, matrix $\mathbf{Q}$ is set to be equal in both simulations and is defined as follows

$$\mathbf{Q} = diag(1000, 100, \cdots, 100, 1, \cdots, 1).$$

This means that (1) the tracking error is more costly in the commodity layers than the in transportation ones ($\mathbf{Q}_{ii} = 100 > \mathbf{Q}_{ii} = 1$), (2) the inventory level of finished products in nodes representing the retail level are the most critical ones ($\mathbf{Q}_{ii} = 1000$).

As previously stated, the difference between simulations consists in the weights attributed to different transportation modes. Consequently, matrix $\mathbf{R} = diag(\rho_1, \cdots, \rho_i)$ is set to be different in each case. In Simulation 1 the movement of transport agents with greater loading capacity is considered to be more *costly* than the rest. Namely, T1, T3, and T5 are set to have a weight, ρ_i, equal to 1.5, whereas a $\rho_i = 3.5 \times 10^4$ was set for T2, T4, and T6. In turn, in Simulation 2 all transportation types were set to have the same weight, yielding $\mathbf{R}_{ii} = 1.5$. Table 3.7 presents a summary of the MPC parameter specifications for both simulations.

Table 3.7: MPC parameter specifications.

	H_p	$\mathbf{Q}$	ρ	
			T1, T3, T5	**T2, T4, T6**
Simulation 1	20	$diag(1000, 100, \cdots, 100, 1, \cdots, 1)$	1.5	3.5×10^4
Simulation 2	20	$diag(1000, 100, \cdots, 100, 1, \cdots, 1)$	1.5	1.5

3.3 Results

Figure 3.3 presents the inflow of commodities into the different demand nodes (Retail 1, Retail 2, and Retail 3), whereas Figure 3.4 presents the transformation of raw materials into finished products. As can be seen, the demand is met, while respecting the pre-established reception time-windows. It is interesting to take note of the delay on the shipment of P3. To better understand this, recall that a transport agent takes 2h to move from Manuf 2 to Retail 3 (see Figure 3.1). Inspecting Figure 3.4(b) it is clear that P3 is stored at 17h (i.e. it enters node 17, see Figure 3.2) . Considering that goods would yet need to be loaded (i.e. pass through node 15) — which would require 1h more — and the travelling between sites takes 2h, the order would arrive at Retail 2 at 20h, which would violate the reception time-windows previously defined. Therefore, the controller postponed its shipment to the next available time-slot.

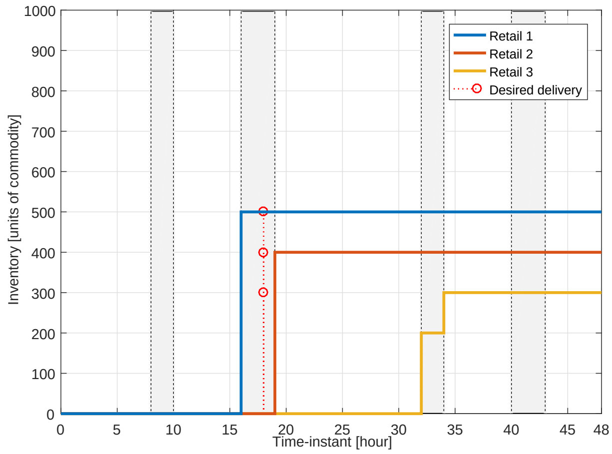
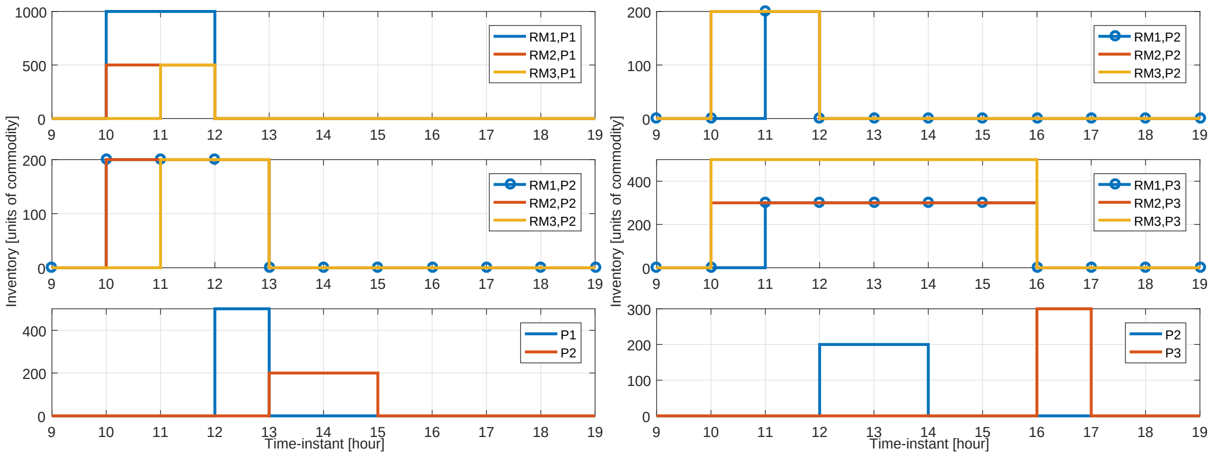


Figure 3.3: Flow of commodities into demand nodes (Retail 1, Retail 2, and Retail 3). Reception periods are represented as grey area.

Figure 3.4 clearly shows the production scheduling. Two points should be noted. First, the constraint that manufacturing sites work in a flow shop, i.e. it is not possible to have parallel production, is satisfied. Second, the choice of which product to produce first is a decision left entirely to the controller, which has to decide what would be the best sequence of decisions in order to meet the demand.

Transportation wise, one is interested in monitoring the position and status of the available transport agents over time. To illustrate how the proposed approach is able to provide the required information, focus on Simulation 1. Figure 3.5 presents a holistic view of the transportation resources over time. At any given moment, it is possible to see how many agents of each type are free, i.e. awaiting task assignment. To have a more detailed information about how and where are these resources being used, one has to resort to a representation such as the one shown in Figure 3.6, where one can have a much deeper insight into the flow of transportation



(a) Manuf 1.

(b) Manuf 2.

Figure 3.4: Transformation of raw materials into finished products. (Nodes 19 — 26, see Figure 3.2.)

occurring across a specific area (in this case the manufacturing sites). Note as Figures 3.4 and 3.6 complement each other, giving a detailed account of what happened in manufacturing sites, and showing the impact of transportation on all supply chain activities. In this respect, notice that the production of P3 (Figure 3.4(b)) was postponed because a delivery delay of 100 units of raw material RM3 that only arrived at 16h at Manuf 2 (see T6 in Figure 3.6(b)). Additionally, one can make use of other ways of monitoring the fleet, as shown in Figure 3.7. Namely, one can monitor what is happening in a single node, regarding a particular transportation type as presented in Figure 3.7(a), or take a more holistic view and monitor the occupied agents across the network, to know how many are waiting assignment, how many are loading, and how many are already in-transit, as show in Figure 3.7(b).

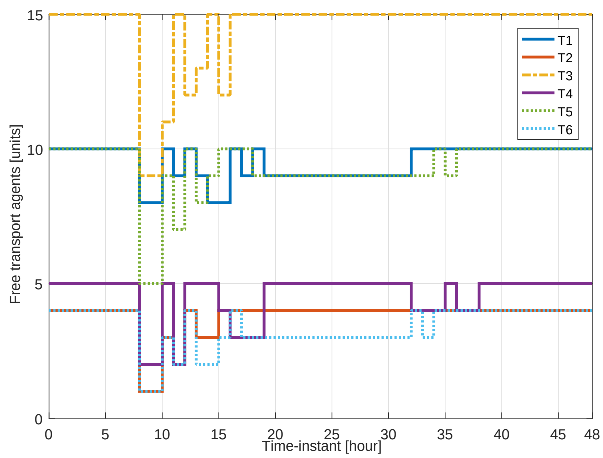
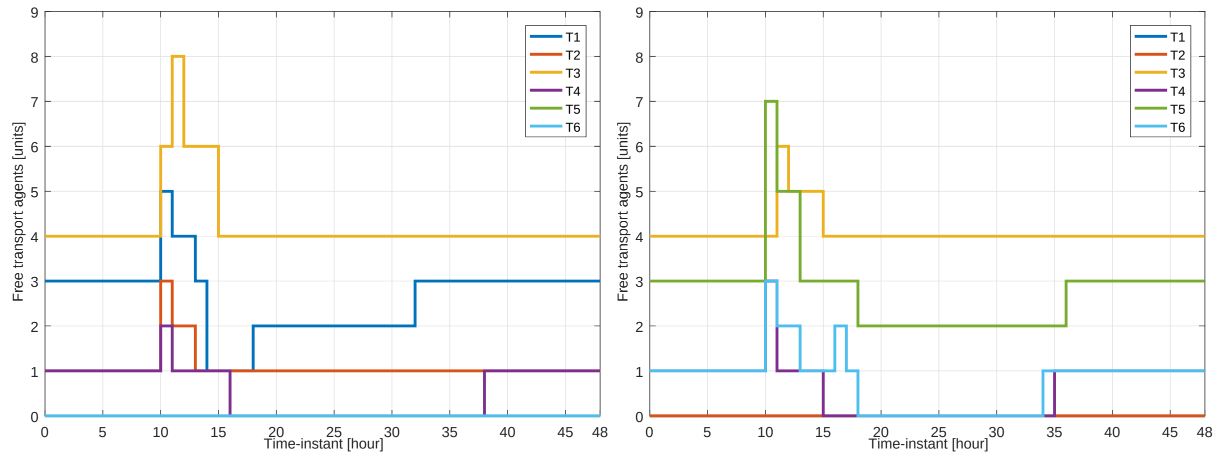


Figure 3.5: Aggregate number of transport agents awaiting task assignment, over time.

Notwithstanding, it is yet possible to monitor how many trips were made, and if they consist on either shipments, or repositioning trips. As can be seen in Table 3.8, all transport agents got back to their initial position, which extends the insight of Figure 3.5.

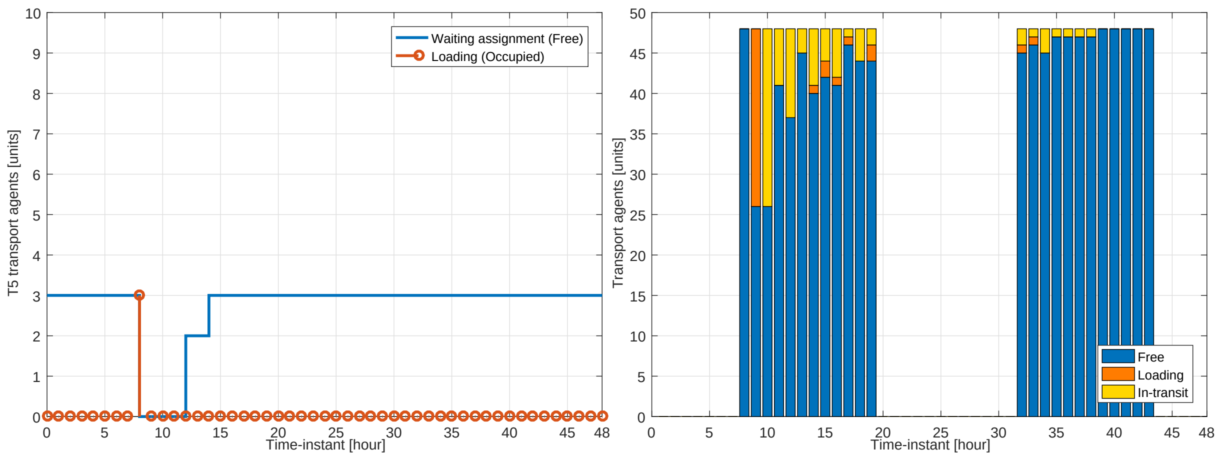
In turn, Figure 3.8 presents a comparison of the transportation modes assigned in each simulation, stressing an important information presented in Table 3.8. As can be seen, the controller tended to assign more transport



(a) Manuf 1.

(b) Manuf 2.

Figure 3.6: Flow of transport agents across manufacturing sites. (Nodes 17 — 18, and 21 — 26, see Figure 3.2.)



(a) Agents of type T5 in RM3. (Nodes 47 – 41, see Figure 3.2).

(b) Fleet status overview.

Figure 3.7: Additional fleet monitoring tools.

Table 3.8: Total performed trips in both simulations.

	Simulation 1						**Simulation 2**					
	T1	**T2**	**T3**	**T4**	**T5**	**T6**	**T1**	**T2**	**T3**	**T4**	**T5**	**T6**
With cargo	8	7	14	10	12	10	4	9	12	11	14	8
Without cargo	8	7	14	10	12	10	4	9	12	11	14	8

agents with higher loading capacity in Simulation 2. Moreover, it is clear that whenever there is a possibility to choose between two modes, the controller selected the cheapest one more often. This is specially visible when comparing T1 and T2. In Simulation 1, the assignment of T1 occurred twice more often than in Simulation 2, which enabled to save two of the more expensive T2 shipments.

Furthermore, one may be interested on gaining a panoramic and integrated overview of the supply chain as whole. Figure 3.9 presents a SC overview, integrating both the perspective of shippers and transport providers, where it is clear fleet-usage decreases over time as commodities are moved from up- to down-stream nodes.

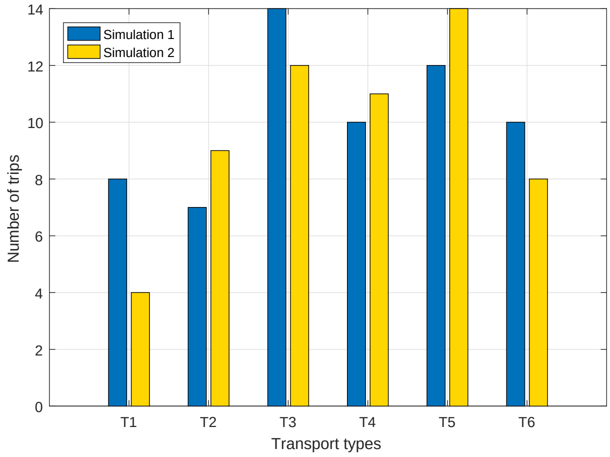


Figure 3.8: Transportation mode assignment.

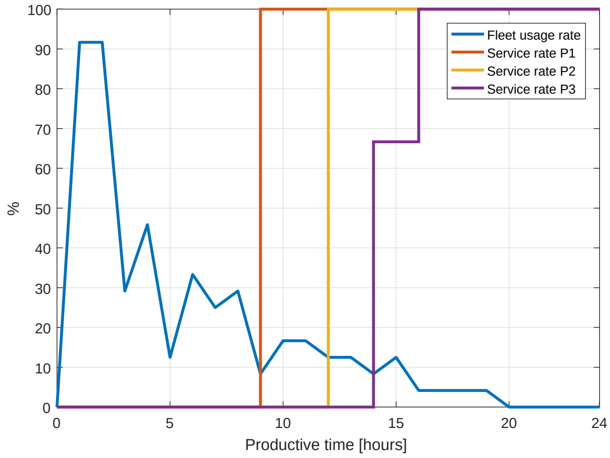


Figure 3.9: SC overview, integrating both the perspective of shippers and transport providers.

Finally, Table 3.9 presents information regarding computational performance. As can be seen, the maximum iteration time is bellow 3 minutes (i.e. each hour is simulated in less than 3 minutes), which in a supply chain management context is fast enough to be considered *real-time*. It should be noted that, new levels of flexibility could be achieved by employing a distributed control approach. In particular, an extension of the work presented in Araújo et. al [1] would reduce the computational burden, and create the possibility of addressing situations such as geographical extensions of the supply chain without the need for re-modelling the whole chain.

Table 3.9: Computational performance.

	min It. Time [min]	**max It. Time [min]**	**avg It. Time [min]**	**Total Sim. Time [h]**
Simulation 1	1.65	2.10	1.72	1.37
Simulation 2	1.59	2.32	1.69	1.35

Chapter 4

Conclusion and future work

To answer to the increasing demand for higher levels of efficiency, quality of service, timeliness, and responsiveness across supply chains (SCs), a new dynamic approach for real-time supply chain management integrating transportation operations, based on a model predictive control (MPC) framework was proposed. On the one hand, shippers are interested in providing customers the right product, in the right amount, at the right time, and at the right place. On the other hand, transport providers want to efficiently allocate goods to their resources and minimise the number of movements. However, these topics have either been studied independently from each other, or integrated for tactical or strategic purposes only. The assumptions found in the literature result in models with limited capacity to tackle real-world operational issues in an integrated way. Namely, the assumption that transportation resources are always available is a critical one, since it disregards the symbiotic relationship between movement of goods and movement of transports.

The outlined modelling framework is based on a flow perspective and builds on the fact that there are two fundamental flows in SCs. Namely, *information* and *material* flows, where the material flow may be further divided into flow of goods and flow of transports. In turn, a SC can be represented as a network of independent actors. In the proposed modelling approach, nodes represent different SC members, and links represent a connection between two SC members through which material may flow. Each node is characterised by a minimum and maximum inventory level, where inventory refers to either stocked goods, or parked transportation vehicles. Links are also limited by a minimum and maximum flow capacities. Thus, the SC dynamics can be considered as a superposition of two fundamental layers: (1) the layer of goods, and (2) the layer of transportation, where each layer consists of (possibly different) networks across which material is allowed to flow. Extending this fundamental notion of flow of goods *vs* flow of transportation, a generic SC is given as a collection of stacked networks (or layers) of different goods and transports. In the proposed modelling framework, the movement of goods is represented by a synchronised, superimposed flow of goods and its respective means of transportation.

The devised notation was then used to develop a centralised, constrained MPC scheme, where the variables' mapping from the MPC framework to the SC domain was accomplished by representing inventories as states, and flows of material as control actions. To achieve the desired system behaviour, a set of constraints was defined. The MPC problem was formulated as a quadratic programming problem, in which desired inventory levels must be achieved.

The proposed modelling framework was shown to result in linear, time-invariant, state-space representations of supply chains that are both controllable and observable.

Results have validated the proposed approach and shown the autonomous planning capabilities of the devised controller. Namely, it was shown that the proposed approach is able to integrate manufacturing and transportation operations so that: (1) pre-defined reception time-windows are respected; (2) moving costs are minimised; (3) transportation reallocation policies are respected; and (4) path-planning for on-time delivery is accomplished. Regarding inventory and manufacturing activities alone, the devised approach is able to effectively and efficiently manage production (i.e. to determine the amount of goods to produce, where to produce them, and in which sequence), resulting in optimal inventory levels. That is, the right amount of commodities was stored at the right time, and at the right place only to assure manufacturing activities could take place, enabling supply chain members to reduce their inventory levels to zero. Furthermore, the proposed approach was able not only to effectively control the supply chain at an operational level, but also to provide deep insights into each activity and their symbiotic interactions.

Interpretability, tractability and flexibility came, however, at the expense of some limitations. First, the proposed approach requires a pre-defined transportation fleet where transportation types must be already defined and assigned to specific commodities. That is, it is not possible for a given transportation type to morph into another. Consider, for instance, the case in which a given transport could vary its maximum load capacity. Such a case, would require a model in which transport agents could belong to different types over time, depending on their specific features. Thus, a valuable follow-up implementation would be able to address these situations. Second, new levels of flexibility could be achieved by employing a distributed control approach for horizontal integration. In such a potential situation, one would be able to explicitly define both the shippers' and the transport providers' goals independently, providing a more realistic and flexible way of tackling supply chain management problems. Third, the proposed framework results in dynamical models that make explicit use of sparse matrices, which hampers scalability. To cope with such a limitation, the implementation of a decentralized/distributed control scheme seems to be the a possible solution. Fourth, although the proposed modelling approach takes into consideration different types of goods, they are assumed to not deteriorate over time. Such an assumption excludes by design supply chains of perishable goods. A valuable future implementation would be to address such challenging problems. Fifth, the transportation costs were defined resorting to a parameter, ρ, which has not a direct real-world meaning. A future implementation would benefit from having a more meaningful cost-function.

Nevertheless, the developed methodology is very relevant for the supply chain management domain. First, it provides a framework for better interoperability between independent SC members. Second, the presented modelling approach can be easily coupled with other state-of-the art technologies. Namely, it is not difficult to picture a situation in which a forecasting system based on artificial intelligence (e.g. machine, deep learning techniques) is used to dynamically adapt to the ever-changing customer demand, while the control system optimises decisions based on those adaptive predictions, rendering the optimisation itself to be adaptive as well. All in all, the inherent flexibility of the proposed approach, as well as its interoperability and information-sharing features make it a suitable operations management tool for the current Industry 4.0 paradigm.

Bibliography

[1] Eduardo Araújo, João Lemos Nabais, and Miguel Ayala Botto. Distributed Model Predictive Control For Optimal Warehouse Management. In *Jornadas de Distribuição e Logística – JDL 2018*, Setúbal, 2018.

[2] Michael R. Bartolacci, Larry J. LeBlanc, Yasanur Kayikci, and Thomas A. Grossman. Optimization Modeling for Logistics: Options and Implementations. *Journal of Business Logistics*, 33(2):118–127, jun 2012. ISSN 07353766. doi: 10.1111/j.0000-0000.2012.01044.x. URL http://doi.wiley.com/10.1111/j.0000-0000.2012.01044.x.

[3] Benita M Beamon. Supply chain design and analysis:: Models and methods. *International Journal of Production Economics*, 55(3): 281–294, aug 1998. ISSN 0925-5273. doi: 10.1016/S0925-5273(98)00079-6. URL https://www.sciencedirect.com/science/article/pii/S0925527398000796.

[4] Valeria Belvedere, Alberto Grando, and Paola Bielli. A quantitative investigation of the role of information and communication technologies in the implementation of a product-service system. *International Journal of Production Research*, 51(2):410–426, jan 2013. ISSN 0020-7543. doi: 10.1080/00207543.2011.648278. URL http://www.tandfonline.com/doi/abs/10.1080/00207543.2011.648278.

[5] Mohamed Ben-Daya, Elkafi Hassini, and Zied Bahroun. Internet of things and supply chain management: a literature review. *International Journal of Production Research*, pages 1–24, nov 2017. ISSN 0020-7543. doi: 10.1080/00207543.2017.1402140. URL https://www.tandfonline.com/doi/full/10.1080/00207543.2017.1402140.

[6] Robert Blackburn, Josef Kallrath, and Steffen T. Klosterhalfen. Operations research in BASF's supply chain operations. *International Transactions in Operational Research*, 22(3):385–405, may 2015. ISSN 09696016. doi: 10.1111/itor.12118. URL http://doi.wiley.com/10.1111/itor.12118.

[7] F. Blanchini, F. Rinaldi, and W. Ukovich. Least inventory control of multistorage systems with non-stochastic unknown inputs. *IEEE Transactions on Robotics and Automation*, 13(5):633–645, 1997. ISSN 1042296X. doi: 10.1109/70.631225. URL http://ieeexplore.ieee.org/document/631225/.

[8] Franco Blanchini, Franca Rinaldi, and Walter Ukovich. A Network Design Problem for a Distribution System with Uncertain Demands. *SIAM Journal on Optimization*, 7(2):560–578, may 1997. ISSN 1052-6234. doi: 10.1137/S1052623494266262. URL http://epubs.siam.org/doi/10.1137/S1052623494266262.

[9] Franco Blanchini, Raffaele Pesenti, Franca Rinaldi, and Walter Ukovich. Feedback control of production-distribution systems with unknown demand and delays. *IEEE Transactions on Robotics and Automation*, 16(3):313–317, jun 2000. doi: 10.1109/70.850649. URL http://ieeexplore.ieee.org/document/850649/.

[10] E. K. Boukas, H. Yang, and Q. Zhang. Minimax production planning in failure-prone manufacturing systems. *Journal of Optimization Theory and Applications*, 87(2):269–286, nov 1995. ISSN 0022-3239. doi: 10.1007/BF02192564. URL http://link.springer.com/10.1007/BF02192564.

[11] E. K. Boukas, P. Shi, and A. Andijani. Robust inventory-production control problem with stochastic demand. *Optimal Control Applications and Methods*, 20(1):1–20, jan 1999. doi: 10.1002/(SICI)1099-1514(199901/02)20:1<1::AID-OCA642>3.0.CO;2-L. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/{%}28SICI{%}291099-1514{%}28199901/02{%}2920{%}3A1{%}3C1{%}3A{%}3AAID-OCA642{%}3E3.0.CO{%}3B2-L.

[12] E. K. Boukas, P. Shi, and R. K. Agarwal. An application of robust control technique to manufacturing systems with uncertain processing time. *Optimal Control Applications and Methods*, 21(6):257–268, nov 2000. doi: 10.1002/oca.677. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/oca.677.

[13] Martin W. Braun, Daniel E. Rivera, M.E. Flores, W. Matthew Carlyle, and Karl G. Kempf. A Model Predictive Control framework for robust management of multi-product, multi-echelon demand networks. *Annual Reviews in Control*, 27(2):229–245, jan 2003. doi: 10.1016/J.ARCONTROL.2003.09.006. URL https://www.sciencedirect.com/science/article/pii/S1367578803000312.

[14] Charu Chandra and Jānis Grabis. *Supply Chain Configuration*. Springer US, 2007. ISBN 978-0-387-68155-9. doi: 10.1007/978-0-387-68155-9. URL https://www.springer.com/us/book/9781441937780.

[15] Charu Chandra, Jānis Grabis, and Armen Tumanyan. Problem taxonomy: a step towards effective information sharing in supply chain management. *International Journal of Production Research*, 45(11):2507–2544, jun 2007. ISSN 0020-7543. doi: 10.1080/00207540601020486. URL http://www.tandfonline.com/doi/abs/10.1080/00207540601020486.

[16] Council of Supply Chain Management Professionals. SCM Definitions and Glossary of Terms. URL https://cscmp.org/CSCMP/Educate/SCM{_}Definitions{_}and{_}Glossary{_}of{_}Terms/CSCMP/Educate/SCM{_}Definitions{_}and{_}Glossary{_}of{_}Terms.aspx?hkey=60879588-f65f-4ab5-8c4b-6878815ef921.

[17] Jeroen Dejonckheere, Stephen M. Disney, Marc R. Lambrecht, and Dennis R. Towill. Measuring and avoiding the bullwhip effect: A control theoretic approach. *European Journal of Operational Research*, 147(3):567–590, jun 2003. doi: 10.1016/S0377-2217(02)00369-7. URL https://www.sciencedirect.com/science/article/pii/S0377221702003697.

[18] Stephen M. Disney and Denis R. Towill. A discrete transfer function model to determine the dynamic stability of a vendor managed inventory supply chain. *International Journal of Production Research*, 40(1):179–204, jan 2002. doi: 10.1080/00207540110072975. URL http://www.tandfonline.com/doi/abs/10.1080/00207540110072975.

[19] Stephen M. Disney and Dennis R. Towill. Eliminating drift in inventory and order based production control systems. *International Journal of Production Economics*, 93-94:331–344, jan 2005. doi: 10.1016/J.IJPE.2004.06.031. URL https://www.sciencedirect.com/science/article/pii/S0925527304002531.

[20] Alexandre Dolgui and Jean-Marie Proth. *Supply Chain Engineering*. Springer-Verlag London, London, 2010. ISBN 978-1-4471-6200-1. doi: 10.1007/978-1-84996-017-5. URL https://www.springer.com/gp/book/9781849960168.

[21] Chuanwen Dong, Robert Boute, Alan McKinnon, and Marc Verelst. Investigating synchromodality from a supply chain perspective. *Transportation Research Part D: Transport and Environment*, 61:42–57, jun 2018. ISSN 1361-9209. doi: 10.1016/J.TRD.2017.05.011. URL https://www.sciencedirect.com/science/article/pii/S136192091630387X.

[22] Anne Dujin, Cornelia Geissler, and Dirk Horstkötter. Industry 4.0. The new industrial revolution How Europe will succeed. *Think-Act, Roland Berger SC*, 2014.

[23] J W Forrester. Industrial Dynamics. *Journal of the Operational Research Society*, 48(10):1037–1041, oct 1997. ISSN 0160-5682. doi: 10.1057/palgrave.jors.2600946. URL https://www.tandfonline.com/doi/full/10.1057/palgrave.jors.2600946.

[24] Terry P. Harrison. Principles for the Strategic Design of Supply Chains. In *The Practice of Supply Chain Management: Where Theory and Application Converge*, pages 3–12. Springer-Verlag, New York, 2004. doi: 10.1007/0-387-27275-5_1. URL http://link.springer.com/10.1007/0-387-27275-5{_}1.

[25] Berthold Herrendorf, Richard Rogerson, and Ákos Valentinyi. Growth and Structural Transformation. *Handbook of Economic Growth*, 2: 855–941, jan 2014. ISSN 1574-0684. doi: 10.1016/B978-0-444-53540-5.00006-9. URL https://www.sciencedirect.com/science/article/pii/B9780444535405000069.

[26] Frederick S Hillier. *Introduction to operations research*. McGraw-Hill, 2001. ISBN 9780071181631.

[27] William Ho, Tian Zheng, Hakan Yildiz, and Srinivas Talluri. Supply chain risk management: a literature review. *International Journal of Production Research*, 53(16):5031–5069, aug 2015. ISSN 0020-7543. doi: 10.1080/00207543.2015.1030467. URL http://www.tandfonline.com/doi/full/10.1080/00207543.2015.1030467.

[28] Arild Hoff, Henrik Andersson, Marielle Christiansen, Geir Hasle, and Arne Løkketangen. Industrial aspects and literature survey: Fleet composition and routing. *Computers & Operations Research*, 37(12):2041–2061, dec 2010. ISSN 0305-0548. doi: 10.1016/J.COR.2010.03.015. URL https://www.sciencedirect.com/science/article/pii/S0305054810000699.

[29] Erik Hofmann and Marco Rüsch. Industry 4.0 and the current status as well as future prospects on logistics. *Computers in Industry*, 89:23–34, aug 2017. ISSN 0166-3615. doi: 10.1016/J.COMPIND.2017.04.002. URL https://www.sciencedirect.com/science/article/pii/S0166361517301902.

[30] Dmitry Ivanov. An adaptive framework for aligning (re)planning decisions on supply chain strategy, design, tactics, and operations. *International Journal of Production Research*, 48(13):3999–4017, jul 2010. ISSN 0020-7543. doi: 10.1080/00207540902893417. URL https://www.tandfonline.com/doi/full/10.1080/00207540902893417.

[31] Dmitry Ivanov. *Structural Dynamics and Resilience in Supply Chain Risk Management*, volume 265 of *International Series in Operations Research & Management Science*. Springer International Publishing, Cham, 2018. ISBN 978-3-319-69304-0. doi: 10.1007/978-3-319-69305-7. URL http://link.springer.com/10.1007/978-3-319-69305-7.

[32] Dmitry Ivanov, Boris Sokolov, and Joachim Kaeschel. Integrated supply chain planning based on a combined application of operations research and optimal control. *Central European Journal of Operations Research*, 19(3):299–317, sep 2011. ISSN 1435-246X. doi: 10.1007/s10100-010-0185-0. URL http://link.springer.com/10.1007/s10100-010-0185-0.

[33] Dmitry Ivanov, Suresh Sethi, Alexandre Dolgui, and Boris Sokolov. A survey on control theory applications to operational systems, supply chain management, and Industry 4.0. *Annual Reviews in Control*, oct 2018. ISSN 1367-5788. doi: 10.1016/J.ARCONTROL.2018.10.014. URL https://www.sciencedirect.com/science/article/pii/S1367578818301238.

[34] H Kagermann, J Helbig, A Hellinger, and W Wahlster. Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Securing the future of German manufacturing industry; final report of the. 2013.

[35] Yener Kandogan and Scott D. Johnson. Role of economic and political freedom in the emergence of global middle class. *International Business Review*, 25(3):711–725, jun 2016. ISSN 0969-5931. doi: 10.1016/J.IBUSREV.2015.02.005. URL https://www.sciencedirect.com/science/article/abs/pii/S0969593115000311.

[36] Paul M. Kennedy. *The rise and fall of the great powers : economic change and military conflict from 1500 to 2000*. William Collins, 10 edition, 2017. ISBN 0006860524.

[37] Mayuresh V. Kothare, Venkataramanan Balakrishnan, and Manfred Morari. Robust constrained model predictive control using linear matrix inequalities. *Automatica*, 32(10):1361–1379, oct 1996. ISSN 0005-1098. doi: 10.1016/0005-1098(96)00063-5. URL https://www.sciencedirect.com/science/article/pii/0005109896000635.

[38] Stephan Kreipl and Michael Pinedo. Planning and Scheduling in Supply Chains: An Overview of Issues in Practice. *Production and Operations Management*, 13(1):77–92, jan 2009. ISSN 10591478. doi: 10.1111/j.1937-5956.2004.tb00146.x. URL http://doi.wiley.com/10.1111/j.1937-5956.2004.tb00146.x.

[39] Simon Kuznets. Modern Economic Growth: Findings and Reflections, 1973. URL https://www.jstor.org/stable/1914358.

[40] Chandra S. Lalwani, Stephen M. Disney, and Denis R. Towill. Controllable, observable and stable state space representations of a generalized order-up-to policy. *International Journal of Production Economics*, 101(1):172–184, may 2006. ISSN 0925-5273. doi: 10.1016/J.IJPE.2005.05.014. URL https://www.sciencedirect.com/science/article/pii/S0925527305001428.

[41] Douglas M Lambert and Martha C Cooper. Issues in Supply Chain Management. *Industrial Marketing Management*, 29(1):65–83, jan 2000. ISSN 0019-8501. doi: 10.1016/S0019-8501(99)00113-3. URL https://www.sciencedirect.com/science/article/abs/pii/S0019850199001133.

[42] William Mather Lewis. The Significance of Transportation to Civilization. *The ANNALS of the American Academy of Political and Social Science*, 187(1):1–6, sep 1936. ISSN 0002-7162. doi: 10.1177/000271623618700102. URL http://journals.sagepub.com/doi/10.1177/000271623618700102.

[43] Yongxin Liao, Fernando Deschamps, Eduardo de Freitas Rocha Loures, and Luiz Felipe Pierin Ramos. Past, present and future of Industry 4.0 - a systematic literature review and research agenda proposal. *International Journal of Production Research*, 55(12): 3609–3629, jun 2017. ISSN 0020-7543. doi: 10.1080/00207543.2017.1308576. URL https://www.tandfonline.com/doi/full/10.1080/00207543.2017.1308576.

[44] Jan Marian Maciejowski. *Predictive control: with constraints*. Pearson education, 2002.

[45] Brad McDonald. International Trade: Commerce among Nations, 2018. URL https://www.imf.org/external/pubs/ft/fandd/basics/trade.htm.

[46] Alyssa Mosca, Navneet Vidyarthi, and Ahmet Satir. Integrated Transportation - Inventory Models: A Review. *Operations Research Perspectives*, 6:100101, feb 2019. ISSN 2214-7160. doi: 10.1016/J.ORP.2019.100101. URL https://www.sciencedirect.com/science/article/pii/S2214716018302562.

[47] João Lemos Nabais, Rudy R. Negenborn, and Miguel Ayala Botto. A Novel Predictive Control Based Framework for Optimizing Intermodal Container Terminal Operations. In Voß S. Hu H., Shi X., Stahlbock R., editor, *Computational Logistics. ICCL 2012. Lecture Notes in Computer Science, vol 7555.*, pages 53–71. Springer, Berlin, Heidelberg, 2012. ISBN 978-3-642-33587-7. doi: 10.1007/978-3-642-33587-7_4. URL http://link.springer.com/10.1007/978-3-642-33587-7{_}4.

[48] Risto Nieminen, Laura Puccio, and Marine Charlotte André. The added value of international trade and impact of trade barriers. Technical report, European Parliament, 2017. URL https://epthinktank.eu/2017/10/27/the-added-value-of-international-trade-and-impact-of-trade-barriers/.

[49] OECD. *Policy Priorities for International Trade and Jobs*. 2012. URL http://www.oecd.org/site/tadicite/policyprioritiesforinternationaltradeandjobs.htm.

[50] M. Ortega and L. Lin. Control theory applications to the production–inventory problem: a review. *International Journal of Production Research*, 42(11):2303–2322, jun 2004. ISSN 0020-7543. doi: 10.1080/00207540410001666260. URL https://www.tandfonline.com/doi/full/10.1080/00207540410001666260.

[51] Esteban Ortiz-Ospina and Nicolas Lippolis. Our World in Data - Are emerging economies deindustrializing too quickly?, 2018. URL https://ourworldindata.org/growth-and-structural-transformation-are-emerging-economies-industrializing-too-quickly.

[52] Ercan Oztemel and Samet Gursev. Literature review of Industry 4.0 and related technologies. *Journal of Intelligent Manufacturing*, pages 1–56, jul 2018. ISSN 0956-5515. doi: 10.1007/s10845-018-1433-8. URL http://link.springer.com/10.1007/s10845-018-1433-8.

[53] Edgar Perea, Ignacio Grossmann, Erik Ydstie, and Turaj Tahmassebi. Dynamic modeling and classical control theory for supply chain management. *Computers & Chemical Engineering*, 24(2-7):1143–1149, jul 2000. ISSN 0098-1354. doi: 10.1016/S0098-1354(00)00495-6. URL https://www.sciencedirect.com/science/article/pii/S0098135400004956.

[54] Edgar Perea-López, B.Erik Ydstie, and Ignacio E. Grossmann. A model predictive control strategy for supply chain optimization. *Computers & Chemical Engineering*, 27(8-9):1201–1218, sep 2003. ISSN 0098-1354. doi: 10.1016/S0098-1354(03)00047-4. URL https://www.sciencedirect.com/science/article/pii/S0098135403000474.

[55] Jean-Paul Rodrigue, C Comtois, and Brian Slack. *The geography of transport systems*. 2016. doi: 10.4324/9781315618159. URL https://www.researchgate.net/publication/318837359{_}The{_}geography{_}of{_}transport{_}systems.

[56] Haralambos Sarimveis, Panagiotis Patrinos, Chris D. Tarantilis, and Chris T. Kiranoudis. Dynamic modeling and control of supply chain systems: A review. *Computers & Operations Research*, 35(11):3530–3561, nov 2008. ISSN 0305-0548. doi: 10.1016/J.COR.2007.01.017. URL https://www.sciencedirect.com/science/article/pii/S0305054807000366.

[57] Panos Seferlis and Nikolaos F. Giannelos. A two-layered optimisation-based control strategy for multi-echelon supply chain networks. *Computer Aided Chemical Engineering*, 14:509–514, jan 2003. doi: 10.1016/S1570-7946(03)80166-9. URL https://www.sciencedirect.com/science/article/pii/S1570794603801669.

[58] Hartmut Stadtler. *Supply Chain Management — An Overview*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. doi: 10.1007/978-3-540-74512-9_2. URL http://link.springer.com/10.1007/978-3-540-74512-9{_}2.

[59] Marc Teignier. The role of trade in structural transformation. *Journal of Development Economics*, 130:45–65, jan 2018. ISSN 0304-3878. doi: 10.1016/J.JDEVECO.2017.09.002. URL https://www.sciencedirect.com/science/article/pii/S0304387817300652.

[60] Denis R. Towill. Dynamic analysis of an inventory and order based production control system. *International Journal of Production Research*, 20(6):671–687, nov 1982. doi: 10.1080/00207548208947797. URL http://www.tandfonline.com/doi/abs/10.1080/00207548208947797.

[61] United Nations Industrial Development Organization. Industrial Development Report 2018. Demand for Manufacturing: Driving Inclusive and Sustainable Industrial Development. Overview. Technical report, Vienna, 2017. URL https://unido.org/resources-publications-flagship-publications-industrial-development-report-series/industrial-development-report-2018.

[62] Bart van Riessen, Rudy R. Negenborn, and Rommert Dekker. Synchromodal Container Transportation: An Overview of Current Topics and Research Opportunities. pages 386–397. 2015. doi: 10.1007/978-3-319-24264-4_27. URL http://link.springer.com/10.1007/978-3-319-24264-4{_}27.

[63] Sean P. Willems. Data Set —Real-World Multiechelon Supply Chains Used for Inventory Optimization. *Manufacturing & Service Operations Management*, 10(1):19–23, jan 2008. doi: 10.1287/msom.1070.0176. URL http://pubsonline.informs.org/doi/abs/10.1287/msom.1070.0176.

Appendix A

System characterisation

In this chapter, it is shown that the proposed modelling approach results in linear, time-invariant models that are both reachable and observable.

A.1 Linearity

A system is said to be linear if and only if the principles of superposition (sometimes also referred to as additivity) and homogeneity are verified.

Superposition The principle of superposition states that for a given system G with inputs $x_1(n)$ and $x_2(n)$, the following relationship holds:

$$G\left[x_1(n)\ +\ x_2(n)\right]\ =\ G[x_1(n)]\ +\ G[x_2(n)]. \tag{A.1}$$

Homogeneity The principle of homogeneity, in turn, states that for a given system G, an input $x_1(n)$ and a constant b, the following relationship holds:

$$G\left[b\, x_1(n)\right]\ =\ b\, G\left[x_1(n)\right]. \tag{A.2}$$

Recall that the system at hand is a network of various nodes connected by arcs, in which the inventory of a given resource m at node i is modelled as follows,

$$x_{i_m}(k+1) = x_{i_m}(k) + \Delta x_{i_m}(k) \tag{A.3}$$

$$= x_{i_m}(k) + \sum_j u_{ji_m}(k) - \sum_w u_{iw_m}(k) + d_{i_m}(k). \tag{A.4}$$

Note that, for each node there are three types of input. Namely, u_{ji_m}, u_{iw_m}, and d_{i_m}, representing the in- and out-flows, and the exogenous input, respectively. For an easier exposition, consider the following notation,

$$x_0(k) := \text{ initial inventory, at time } k, \tag{A.5}$$

$$u_i(k) := \text{ inflow of resources, at time } k, \tag{A.6}$$

$$u_o(k) := \text{ outflow of resources, at time } k, \tag{A.7}$$

$$d(k) := \text{ exogenous input, at time } k, \tag{A.8}$$

$$y(k) := \text{ output, at time } k. \tag{A.9}$$

Thus, a system composed of a single node as the one represented in Figure A.1, yields,

$$y[u_i, u_o, d](k) = x_0(k) + u_i(k) - u_o(k) + d(k); \tag{A.10}$$

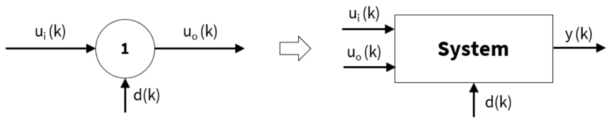


Figure A.1: Change from network to system representation.

Letting the system starts from rest (i.e. $x_0(k) = 0$) yields,

$$y_1(\alpha u_{i_1}, \beta u_{o_1}, \theta d_1)(k) + y_2(u_{i_2}, u_{o_2}, d_2)(k) = \tag{A.11}$$

$$= (\alpha u_{i_1} - \beta u_{o_1} + \theta d_1) + (\alpha u_{i_2} - \beta u_{o_2} + \theta d_2) \tag{A.12}$$

$$= \alpha\left(u_{i_1} + u_{i_2}\right) - \beta\left(u_{o_1} + u_{o_2}\right) + \theta\left(d_1 + d_2\right) \tag{A.13}$$

$$= y[\alpha\left(u_{i_1} + u_{i_2}\right), \beta\left(u_{o_1} + u_{o_2}\right), \theta\left(d_1 + d_2\right)](k), \tag{A.14}$$

for all $\alpha, \beta, \theta, u_{i_1}, u_{o_1}, d_1 \neq 0$. Thus, the system is linear.

Analogously, one can take the case in which the system is given as a network of a generic size, yielding,

$$\mathbf{y}_1(\alpha \mathbf{u}_{i_1}, \beta \mathbf{u}_{o_1}, \theta \mathbf{d}_1)(k) + \mathbf{y}_2(\alpha \mathbf{u}_{i_2}, \beta \mathbf{u}_{o_2}, \theta \mathbf{d}_2)(k) = \tag{A.15}$$

$$= (\alpha \mathbf{u}_{i_1} - \beta \mathbf{u}_{o_1} + \theta \mathbf{d}_1(k)) + (\alpha \mathbf{u}_{i_2} - \beta \mathbf{u}_{o_2} + \theta \mathbf{d}_2) \tag{A.16}$$

$$= \alpha\left(\mathbf{u}_{i_1} + \mathbf{u}_{i_2}\right) - \beta\left(\mathbf{u}_{o_1} + \mathbf{u}_{o_2}\right) + \theta\left(\mathbf{d}_1 + \mathbf{d}_2\right) \tag{A.17}$$

$$= \mathbf{y}[\alpha(\mathbf{u}_{i_1} + \mathbf{u}_{i_2}), \beta(\mathbf{u}_{o_1} + \mathbf{u}_{o_2}), \theta(\mathbf{d}_1 + \mathbf{d}_2)](k), \tag{A.18}$$

for all $\alpha, \beta, \theta, \neq 0$, and $\mathbf{u}_{i_1}, \mathbf{u}_{o_1}, \mathbf{d}_1 \neq \mathbf{0}$.

A.2 Time-invariance

A dynamic system is said to be time-invariant if the mapping of inputs into outputs does not change in time. In other words, if an input signal is delayed τ time-instants, so must be the output, y. Thus, a system is said to be time-invariant if the following relationship holds,

$$y(t-\tau) = F\left[u(t-\tau)\right]. \tag{A.19}$$

Clearly, the system

$$\mathbf{y}(\mathbf{u}_i, \mathbf{u}_o, \mathbf{d})(k) = \mathbf{x}_0(k) + \mathbf{u}_i(k) - \mathbf{u}_o(k) + \mathbf{d}(k)$$

is time-invariant, since

$$\mathbf{y}(\mathbf{u}_i, \mathbf{u}_o, \mathbf{d})(k-\tau) = \mathbf{x}_0(k-\tau) + \mathbf{u}_i(k-\tau) - \mathbf{u}_o(k-\tau) + \mathbf{d}(k-\tau). \tag{A.20}$$

A.3 Controllability and observability

A linear dynamic process is said to be controllable if it is possible to drive the system from any initial state $\mathbf{x}(0) \neq 0$ to any other state $\mathbf{x}(N) = 0$, in a finite time interval N, by means of a sequence of control actions $\mathbf{u}(k)$. The term reachability is used instead when it is possible to drive the system from $\mathbf{x}(0) = 0$ to any other state $\mathbf{x}(N) \neq 0$, in a finite time interval N, by means of a sequence of control actions $\mathbf{u}(k)$.

A necessary and sufficient condition for complete state controllability is that the rank of the controllability matrix $\mathcal{C}$, if non-singular, must equal the rank of matrix A, where $\mathcal{C}$ is given by

$$\mathcal{C} = \left[\, \mathbf{B},\ \mathbf{AB},\ \cdots,\ \mathbf{A}^{N-1}\mathbf{B} \,\right] \tag{A.21}$$

and hence,

$$rank\left(\mathcal{C}\right) = n \tag{A.22}$$

with n as the order of the matrix $\mathbf{A}$.

In turn, if $\mathcal{C}$ is singular, one of two conditions must be verified:

- $\mathbf{A}^N\mathbf{x}(0) \in span(\mathcal{C}), \forall \mathbf{x} \neq \mathbf{0}$,
- $\mathbf{A}^k = 0,\ \ \forall k \in \mathbb{R}$.

Analogously, a necessary and sufficient condition for complete state observability is that the rank of the observability matrix $\mathcal{O}$, if non-singular, must equal the rank of matrix A, where $\mathcal{O}$ is given by

$$\mathcal{O} = \left[\, \mathbf{C},\ \mathbf{AC},\ \cdots,\ \mathbf{A}^{N-1}\mathbf{C} \,\right] \tag{A.23}$$

and hence,

$$rank\left(\mathcal{O}\right) = n \tag{A.24}$$

with n as the order of the matrix $\mathbf{A}$.

In the proposed modelling framework the input matrix, $\mathbf{B}$, depends explicitly on the network's configuration. Thus, to prove it results in controllable and observable systems, an exhaustive approach will follow. However, to maintain the problem tractable, demonstrations will focus on networks composed of four nodes. Then, both properties will be tested for all possible network's configurations. Figure A.2 presents the possible configurations a four-node network can assume.

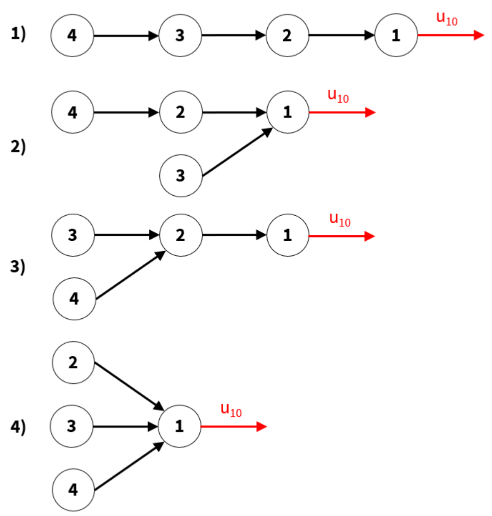


Figure A.2: Possible network's configurations (black), including the control variable that would allow to control the inventory of the most down-stream nodes (red).

One can model the first network as follows,

$$x_1(k+1) = x_1(k) + u_{21} \tag{A.25}$$

$$x_2(k+1) = x_2(k) + u_{32} - u_{21} \tag{A.26}$$

$$x_3(k+1) = x_3(k) + u_{43} - u_{32} \tag{A.27}$$

$$x_4(k+1) = x_4(k) - u_{43}. \tag{A.28}$$

The previous system may yet be written in a more compact form, as follows,

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}_1\mathbf{u}_1(k), \tag{A.29}$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k), \tag{A.30}$$

where $\mathbf{x} = [x_1, x_2, x_3, x_4]^T$, $\mathbf{u}_1 = [u_{21}, u_{32}, u_{43}]^T$, $\mathbf{y} = [y_1, y_2, y_3, y_4]^T$, and

$$\mathbf{A} = \mathbf{C} = \mathbf{I} \in \mathbb{R}^{4\times 4}, \tag{A.31}$$

$$\mathbf{B}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}. \tag{A.32}$$

Analogously, one can define both the input vectors and matrices of the remaining systems as follows,

$$\mathbf{u}_2 = [u_{21}, u_{31}, u_{42}]^T, \ \mathbf{B}_2 = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \tag{A.33}$$

$$\mathbf{u}_3 = [u_{21}, u_{32}, u_{42}]^T,\ \mathbf{B}_3 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \tag{A.34}$$

$$\mathbf{u}_4 = [u_{21}, u_{31}, u_{41}]^T,\ \mathbf{B}_4 = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{A.35}$$

Controllability Making use of the definition of controllability (equation (A.21)), one can compute the controllability matrix for all systems and characterise the system accordingly as follows,

$$\mathcal{C} = \mathcal{C}_1 = \mathcal{C}_2 = \mathcal{C}_3 = \mathcal{C}_4 = \tag{A.36}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 \\ 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 \\ 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 \end{bmatrix}. \tag{A.37}$$

Since $\mathcal{C}$ is rectangular and $\mathbf{A} = \mathbf{I}$, it is necessary to check if the relationship $\mathbf{A}^N\mathbf{x}(0) \in span(\mathcal{C}), \forall \mathbf{x} \neq \mathbf{0}$ holds. In other words, $\mathbf{A}^N\mathbf{x}(0)$ must be written as a linear combination of the vectors that constitute the basis of $\mathcal{C}$, which is defined as follows,

$$\mathbf{V} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}. \tag{A.38}$$

Let $\mathbf{v}$ denote a generic vector defined as $\mathbf{v} = [x, y, z, t]^T \neq 0$, where $x, y, z, t \in \mathbb{R}$. If $\mathbf{v} \in \mathbf{V}$, then there is a vector $\mathbf{w} = [\alpha, \beta, \theta]^T \neq 0$ that satisfies the following relationship,

$$\mathbf{V}\mathbf{w} = \mathbf{v} \tag{A.39}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \theta \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} \Rightarrow y - x = t - z. \tag{A.40}$$

Thus, the system is not completely state controllable.

Notwithstanding, if the uncontrollable state is the most down-stream node, i.e. node 1 (see Figure A.2), then one can still claim the system is controllable, since all the states of interest are controllable. That is, in a supply chain management domain, one can not control when the final customer will buy products, thus the inventory of the nodes representing the retail level are, in fact, uncontrollable.

Since the controllable matrix is rectangular, one can not determine the uncontrollable state by inspection. Therefore, to prove that all states of interest are controllable, it suffices to modify the previous systems, adding a control variable to the most down-stream nodes (see Figure A.2) and repeat the previous analysis. If the rank of the *new* controllability matrix equals the order of the system, then it is proven the uncontrollable states are the nodes representing the retailing level.

Repeating the previous procedure, one can define the new input vectors and matrices as follows,

$$\mathbf{u}_1 = [u_{10}, u_{21}, u_{31}, u_{42}]^T, \; \mathbf{B}_1 = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \tag{A.41}$$

$$\mathbf{u}_2 = [u_{10}, u_{21}, u_{31}, u_{42}]^T, \; \mathbf{B}_2 = \begin{bmatrix} -1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \tag{A.42}$$

$$\mathbf{u}_3 = [u_{10}, u_{21}, u_{32}, u_{42}]^T, \; \mathbf{B}_3 = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \tag{A.43}$$

$$\mathbf{u}_4 = [u_{10}, u_{21}, u_{31}, u_{41}]^T, \; \mathbf{B}_4 = \begin{bmatrix} -1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{A.44}$$

Then, the controllability matrix is given by,

$$\mathcal{C} = \mathcal{C}_1 = \mathcal{C}_2 = \mathcal{C}_3 = \mathcal{C}_4 = \tag{A.45}$$

$$= \begin{bmatrix} -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \end{bmatrix}. \tag{A.46}$$

The the basis of $\mathcal{C}$ is now defined as follows,

$$\mathbf{V} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{A.47}$$

Once again, let $\mathbf{v}$ and $\mathbf{w}$ denote two generic non-null vectors defined as $\mathbf{v} = [x, y, z, t]^T$, and $\mathbf{w} = [\alpha, \beta, \theta, \omega]^T$, where $\mathbf{v}, \mathbf{w} \in \mathbb{R}^{4\times1}$. Then,

$$\mathbf{V}\mathbf{w} = \mathbf{v} \tag{A.48}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \theta \\ \omega \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} \Rightarrow \begin{cases} \alpha & = -(x+y+z+t) \\ \beta & = -(y+z+t) \\ \theta & = -(z+t) \\ \omega & = -t \end{cases}, \tag{A.49}$$

as was intended to be proven. Thus, the system will be said to be controllable, since the only uncontrollable state represents the inventory of the retailing nodes.

Observability Regarding observability, making use of the definition presented in equation (A.23), one can compute the observability matrix and prove the system is observable as follows,

$$\mathcal{O} = \mathcal{O}_1 = \mathcal{O}_2 = \mathcal{O}_3 = \mathcal{O}_4 = \mathbf{I} \in \mathbb{R}^{4\times4} \Rightarrow rank(\mathcal{O}) = 4. \tag{A.50}$$

Appendix B

Solving the MPC problem via quadratic programming

In order for the controller to select the *best* sequence of control actions, a prediction model is required, which can be obtained by continuously expanding equation (2.5). Its compact form can then be written as follows,

$$\mathcal{Z}(k) = \mathbf{R}_x \mathbf{x}(k) \, + \, \mathbf{R}_u \mathcal{U}(k) \, + \, \mathbf{R}_d \mathcal{D}(k), \tag{B.1}$$

where,

$$\mathcal{Z}(k) = \left[\tilde{\mathbf{z}}(1), \tilde{\mathbf{z}}(2), \, \cdots, \, \tilde{\mathbf{z}}(H_P)\right]^T, \tag{B.2}$$

$$\mathcal{U}(k) = \left[\mathbf{u}(1), \mathbf{u}(2), \, \cdots, \, \mathbf{u}(H_P)\right]^T, \tag{B.3}$$

$$\mathcal{D}(k) = \left[\tilde{\mathbf{d}}(1), \tilde{\mathbf{d}}(2), \, \cdots, \, \tilde{\mathbf{d}}(H_P)\right]^T, \tag{B.4}$$

$$\mathbf{R_x} = \left[\mathbf{C}_z\mathbf{A}, \, \mathbf{C}_z\mathbf{A}^2, \, \cdots, \, \mathbf{C}_z\mathbf{A}^{Hp}\right]^T, \tag{B.5}$$

$$\mathbf{R_u} = \begin{bmatrix} \mathbf{C}_z\mathbf{B}_u & 0 & 0 & \cdots & 0 \\ \mathbf{C}_z\mathbf{A}\mathbf{B}_u & \mathbf{C}_z\mathbf{B}_u & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_z\mathbf{A}^{Hp-1}\mathbf{B}_u & \mathbf{C}_z\mathbf{A}^{Hp-2}\mathbf{B}_u & \mathbf{C}_z\mathbf{A}^{Hp-3}\mathbf{B}_u & \cdots & \mathbf{C}_z\mathbf{B}_u \end{bmatrix}, \tag{B.6}$$

$$\mathbf{R_d} = \begin{bmatrix} \mathbf{C}_z\mathbf{B}_d & 0 & 0 & \cdots & 0 \\ \mathbf{C}_z\mathbf{A}\mathbf{B}_d & \mathbf{C}_z\mathbf{B}_d & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_z\mathbf{A}^{Hp-1}\mathbf{B}_d & \mathbf{C}_z\mathbf{A}^{Hp-2}\mathbf{B}_d & \mathbf{C}_z\mathbf{A}^{Hp-3}\mathbf{B}_d & \cdots & \mathbf{C}_z\mathbf{B}_d \end{bmatrix}, \tag{B.7}$$

and $\mathcal{Z}(k), \mathcal{U}(k)$, and $\mathcal{D}(k)$ are the future (predicted) outputs, control vectors, and exogenous inputs respectively.

One can rewrite the cost function as follows,

$$J = \sum_{i=1}^{Hp} \left[\mathbf{r}(k+i) \; - \; \tilde{\mathbf{z}}(k+i)\right]^2_{\mathbf{Q}} \; + \; \left[\mathbf{u}(k+i-1)\right]^2_{\mathbf{R}}\,, \tag{B.8}$$

$$= \left[\mathcal{Z}^T(k) - \mathcal{R}^T\right] \mathcal{Q}_Q(k) \left[\mathcal{Z}(k) - \mathcal{R}(k)\right] + \mathcal{Q}_R(k)\, \mathcal{U}^T(k)\mathcal{U}(k), \tag{B.9}$$

where $\mathcal{R}(k) \;=\; \left[\tilde{\mathbf{r}}(1), \tilde{\mathbf{r}}(2), \,\cdots\,, \,\tilde{\mathbf{r}}(H_P)\right]^T$ is the vector of predicted references, and $\mathcal{Q}_Q(k)$ and $\mathcal{Q}_R(k)$ are weighing matrices of appropriate dimensions.

To find the minimum, one needs to take only the following partial derivative,

$$\mathcal{U}^* = \frac{\partial J}{\partial \mathcal{U}} = 2\,\mathcal{U}^T \left[\mathbf{R}_u^T \mathcal{Q}_Q\, \mathbf{R}_u \; + \; \mathcal{Q}_R\right] + 2\left[\mathbf{R}_u^T \mathcal{Q}_Q \left(\mathbf{R}_x \mathbf{x} \; + \; \mathbf{R}_d \mathcal{D} \; - \; \mathcal{R}\right)\right]^T = 0, \tag{B.10}$$

In turn,

$$\mathcal{U}^* = \arg\min\{\frac{1}{2}\,\mathcal{U}^T \mathbf{H}\, \mathcal{U} + \mathbf{c}^T \mathcal{U}\}, \qquad \text{subject to} \quad \mathbf{M}\,\mathcal{U} \;\leq\; \mathbf{\Lambda} \tag{B.11}$$

yielding,

$$\mathbf{H} = 2\left[\mathbf{R}_u^T \mathcal{Q}_Q\, \mathbf{R}_u \; + \; \mathcal{Q}_R\right], \tag{B.12}$$

$$\mathbf{c}^T = 2\left[\mathbf{R}_u^T \mathcal{Q}_Q \left(\mathbf{R}_x \mathbf{x} \; + \; \mathbf{R}_d \mathcal{D} \; - \; \mathcal{R}\right)\right]^T. \tag{B.13}$$

Regarding constraints, recall the problem at hand deals with seven types of constraint: input, output, state, pull, assignment, transformation, and scheduling constraints. For convenience and clarity these are listed next and are further discussed in the following.

$$\mathbf{x}(k) \geq 0, \tag{B.14}$$

$$\underline{\mathbf{z}} \;\leq \mathbf{z}(k) \;\leq\; \overline{\mathbf{z}}, \tag{B.15}$$

$$\underline{\mathbf{u}} \;\leq \mathbf{u}(k) \;\leq\; \overline{\mathbf{u}}, \tag{B.16}$$

$$\sum_i u_{ij_m}(k) \leq x_{i_m}(k-\tau_i), \tag{B.17}$$

$$\sum_{m_P} u_{ij_{m_P}}(k) \leq \lambda\, u_{ij_{m_T}}(k), \tag{B.18}$$

$$|u_{s_i}(k) - u_{s_i'}(k)|_{C_2^N} \leq 0, \tag{B.19}$$

$$\left[u_{s_{i_m}}(k) + u_{s_{i_n}}(k)\right]\Big|_{\forall\, m,n} \leq 0. \tag{B.20}$$

B.1 Input and output constraints

Starting by the input constraints, let n be the number of input variables. Thus, the constraints which are to hold at each time can be expressed as follows,

$$\underline{\mathbf{u}_n} \;\leq \mathbf{u}_n(k) \;\leq\; \overline{\mathbf{u}_n} \tag{B.21}$$

where $\underline{\mathbf{u}_n}$ and $\overline{\mathbf{u}_n}$ represent the minimum and maximum values, respectively.

To write this into the standard form of quadratic programming one needs only to rewrite $\mathbf{u}_n(k)$ as follows,

$$\underline{\mathbf{u}_n} \le \mathbf{u}_n(k) \Leftrightarrow -\mathbf{u}_n(k) \le \underline{\mathbf{u}_n} \Leftrightarrow [-1,\, 0,\, \cdots,\, 0]\; [\,\mathbf{u}_1(k),\, \cdots,\mathbf{u}_n(k)\,]^T \le -\underline{\mathbf{u}_n},$$

$$\mathbf{u}_n(k) \le \overline{\mathbf{u}_n} \Leftrightarrow [\,1,\, 0,\, \cdots,\, 0]\; [\,\mathbf{u}_1(k),\, \cdots,\, \mathbf{u}_n(k)\,]^T \le \overline{\mathbf{u}_n},$$

or in the following compact form

$$\begin{bmatrix} -1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \mathbf{u}(k) \le \begin{bmatrix} -\underline{\mathbf{u}_1} \\ \overline{\mathbf{u}_1} \\ -\underline{\mathbf{u}_2} \\ \overline{\mathbf{u}_2} \\ \vdots \\ -\underline{\mathbf{u}_n} \\ \overline{\mathbf{u}_n} \end{bmatrix} \Leftrightarrow \mathcal{F}_k\, \mathbf{u}(k) \le \mathbf{f}_k \tag{B.22}$$

.

Expanding now for an arbitrary time interval k_l yields,

$$\mathbf{F}\,\mathcal{U}(k) \le \mathbf{f}, \tag{B.23}$$

where

$$\mathcal{U}(k) = [\mathbf{u}(1), \mathbf{u}(2),\, \ldots,\, \mathbf{u}(k_l)]^T, \tag{B.24}$$

$$\mathbf{F} = \begin{bmatrix} \mathcal{F}_1 & 0 & \cdots & 0 \\ 0 & \mathcal{F}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathcal{F}_{k_l}, \end{bmatrix} \mathbf{f} = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \vdots \\ \mathbf{f}_{k_l} \end{bmatrix}. \tag{B.25}$$

Analogously, the following relation holds for the output constraints,

$$\begin{bmatrix} -1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \mathbf{z}(k) \le \begin{bmatrix} -\underline{\mathbf{z}_1} \\ \overline{\mathbf{z}_1} \\ -\underline{\mathbf{z}_2} \\ \overline{\mathbf{z}_2} \\ \vdots \\ -\underline{\mathbf{z}_n} \\ \overline{\mathbf{z}_n} \end{bmatrix} \Leftrightarrow \mathcal{G}_k\, \mathbf{z}(k) \le \mathbf{g}_k. \tag{B.26}$$

Expanding again for an arbitrary time interval k_l yields,

$$\mathbf{G}\,\mathcal{Z}(k) \le \mathbf{g}, \tag{B.27}$$

where

$$\mathcal{Z}(k) = [\mathbf{z}(1), \mathbf{z}(2), \ldots, \mathbf{z}(k_l)]^T, \tag{B.28}$$

$$\mathbf{G} = \begin{bmatrix} \mathcal{G}_1 & 0 & \cdots & 0 \\ 0 & \mathcal{G}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathcal{G}_{k_l} \end{bmatrix}, \; \mathbf{g} = \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_{k_l} \end{bmatrix}. \tag{B.29}$$

B.2 State constraints

State constraints require further manipulations. At each time step the following relation must hold,

$$\begin{bmatrix} -1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{bmatrix} \mathbf{x}(k) \leq \begin{bmatrix} -\underline{\mathbf{x}_1} \\ -\underline{\mathbf{x}_2} \\ \vdots \\ -\underline{\mathbf{x}_n} \end{bmatrix} \Leftrightarrow \mathcal{S}_k \, \mathbf{x}(k) \leq \mathbf{s}_k. \tag{B.30}$$

In turn, recall from equation (2.5) that

$$\mathbf{y}(k) = \mathbf{C}_y \mathbf{x}(k),$$

where $\mathbf{y}(k)$ refers to the collection of inventory levels of all nodes for time-instant k. Therefore, $\mathbf{C}_y = \mathbf{I}$, yielding $\mathbf{y}(k) = \mathbf{x}(k)$.

Rewriting equation (B.30),

$$\begin{bmatrix} -1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{bmatrix} \mathbf{y}(k) \leq \begin{bmatrix} -\underline{\mathbf{y}_1} \\ -\underline{\mathbf{y}_2} \\ \vdots \\ -\underline{\mathbf{y}_n} \end{bmatrix} \Leftrightarrow \mathcal{S}_k \, \mathbf{y}(k) \leq \mathbf{s}_k. \tag{B.31}$$

Expanding for an arbitrary time interval k_l yields,

$$\mathbf{S} \, \mathcal{Y}(k) \leq \mathbf{s}, \tag{B.32}$$

where

$$\mathcal{Y}(k) = [\mathbf{y}(1), \mathbf{y}(2), \ldots, \mathbf{y}(k_l)]^T, \tag{B.33}$$

$$\mathbf{S} = \begin{bmatrix} \mathcal{S}_1 & 0 & \cdots & 0 \\ 0 & \mathcal{S}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathcal{S}_{k_l} \end{bmatrix}, \; \mathbf{s} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_{k_l} \end{bmatrix}. \tag{B.34}$$

Thus, input, output, and state constraints can be written in the following compact forms, respectively:

$$\mathbf{F}\,\mathcal{U}(k) \leq \mathbf{f}, \tag{B.35}$$

$$\mathbf{G}\,\mathcal{Z}(k) \leq \mathbf{g}, \tag{B.36}$$

$$\mathbf{S}\,\mathcal{Y}(k) \leq \mathbf{s}. \tag{B.37}$$

The last step is to write constraints with respect to the same variable $\mathcal{U}$.

Starting with the output constraints and making use of the aforementioned prediction model (equation (B.1)), one can write the output constraints as follows,

$$\mathbf{G}\,\mathcal{Z}(k) \leq \mathbf{g}, \tag{B.38}$$

$$\mathbf{G}\,\mathbf{R}_u\mathcal{U}(k) \leq \mathbf{g} - \mathbf{G}\left(\mathbf{R}_x\mathbf{x} + \mathbf{R}_d\mathcal{D}\right). \tag{B.39}$$

In turn, the expression

$$\mathbf{S}\,\mathcal{Y}(k) \leq \mathbf{s}$$

requires another prediction model $\mathcal{Y}(k)$. Analogously to what was previously done, one can define $\mathcal{Y}(k)$ as follows,

$$\mathcal{Y}(k) = \mathbf{R}_{x_y}\mathbf{x}(k) \,+\, \mathbf{R}_{u_y}\mathcal{U}(k) \,+\, \mathbf{R}_{d_y}\mathcal{D}(k), \tag{B.40}$$

where,

$$\mathbf{R}_{x_y} = \left[\mathbf{C}_y\mathbf{A},\, \mathbf{C}_y\mathbf{A}^2,\, \cdots,\, \mathbf{C}_y\mathbf{A}^{Hp}\right]^T, \tag{B.41}$$

$$\mathbf{R}_{u_y} = \begin{bmatrix} \mathbf{C}_y\mathbf{B}_u & 0 & 0 & \cdots & 0 \\ \mathbf{C}_y\mathbf{A}\mathbf{B}_u & \mathbf{C}_y\mathbf{B}_u & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_y\mathbf{A}^{Hp-1}\mathbf{B}_u & \mathbf{C}_y\mathbf{A}^{Hp-2}\mathbf{B}_u & \mathbf{C}_y\mathbf{A}^{Hp-3}\mathbf{B}_u & \cdots & \mathbf{C}_y\mathbf{B}_u \end{bmatrix}, \tag{B.42}$$

$$\mathbf{R}_{d_y} = \begin{bmatrix} \mathbf{C}_y\mathbf{B}_d & 0 & 0 & \cdots & 0 \\ \mathbf{C}_y\mathbf{A}\mathbf{B}_d & \mathbf{C}_y\mathbf{B}_d & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_y\mathbf{A}^{Hp-1}\mathbf{B}_d & \mathbf{C}_y\mathbf{A}^{Hp-2}\mathbf{B}_d & \mathbf{C}_y\mathbf{A}^{Hp-3}\mathbf{B}_d & \cdots & \mathbf{C}_y\mathbf{B}_d \end{bmatrix}. \tag{B.43}$$

Hence,

$$\mathbf{S}\,\mathcal{Y}(k) \leq \mathbf{s} \tag{B.44}$$

$$\mathbf{S}\,\mathbf{R}_{u_y}\mathcal{U}(k) \leq \mathbf{s} - \mathbf{S}\left(\mathbf{R}_{x_y}\mathbf{x} + \mathbf{R}_{d_y}\mathcal{D}\right). \tag{B.45}$$

Equations (B.35) – (B.37) can therefore be written as follows,

$$\mathbf{F}\,\mathcal{U}(k) \leq \mathbf{f}, \tag{B.46}$$

$$\mathbf{G}\,\mathbf{R}_u\mathcal{U}(k) \leq \mathbf{g} - \mathbf{G}\left(\mathbf{R}_x\mathbf{x} + \mathbf{R}_d\mathcal{D}\right), \tag{B.47}$$

$$\mathbf{S}\,\mathbf{R}_{u_y}\mathcal{U}(k) \leq \mathbf{s} - \mathbf{S}\left(\mathbf{R}_{x_y}\mathbf{x} + \mathbf{R}_{d_y}\mathcal{D}\right). \tag{B.48}$$

B.3 Pull constraints

Previously, in Section 2.2 it was shown that it was possible to build a dynamical model that updates pull constraints each time the controller is updated with the system's state as follows,

$$\mathbf{p}(k+1) = \mathbf{A}_\tau\mathbf{p}(k) + \mathbf{B}_\tau\Delta\mathbf{y}(k) + \mathbf{\Omega}\mathbf{B}_{u_\tau}\mathbf{u}(k-1), \tag{B.49}$$

$$\mathbf{P}(k) = \mathbf{\Omega}_\tau^T\mathbf{p}(k) \tag{B.50}$$

where

$$\mathbf{A}_\tau = diag\left(\mathbf{A}_{\tau_1},\, \mathbf{A}_{\tau_2},\, \cdots,\, \mathbf{A}_{\tau_{n_m}}\right) \in \mathbb{R}^{[n\times n_m(\tau_i+1)]\times[n\times n_m(\tau_i+1)]}, \tag{B.51}$$

$$\mathbf{B}_\tau = diag\left(\mathbf{B}_{\tau_1},\, \mathbf{B}_{\tau_2},\, \cdots,\, \mathbf{B}_{n_m}\right) \in \mathbb{R}^{[n\times n_m(\tau_i+1)]\times n_m}, \tag{B.52}$$

$$\mathbf{\Omega} = diag\left(\mathbf{\Omega}_1,\, \mathbf{\Omega}_2,\, \cdots,\, \mathbf{\Omega}_{n_m}\right) \in \mathbb{R}^{[n\times n_m(\tau_i+1)]\times n_m}, \tag{B.53}$$

$$\mathbf{B}_{u_\tau} = f\left(\mathbf{B}_u\right)^1 \in \mathbb{R}^{n_m\times n_m}, \tag{B.54}$$

$$\mathbf{A}_{\tau_{\mathbf{n_m}}} = \begin{bmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{n(\tau_i+1)\times n(\tau_i+1)}, \tag{B.55}$$

$$\mathbf{B}_{\tau_{\mathbf{n_m}}} = [0,\, \cdots,\, 0,\, 1]^T \in \mathbb{R}^{n(\tau_i+1)\times 1}, \tag{B.56}$$

$$\mathbf{\Omega}_{\mathbf{n_m}} = [1,\, 0,\, \cdots,\, 0]^T \in \mathbb{R}^{n(\tau_i+1)\times 1}, \tag{B.57}$$

$$\mathbf{p}(k) = [\mathbf{x}_1(k-\tau), \cdots,\, \mathbf{x}_{n_m}(k-\tau)]^T \in \mathbb{R}^{[n\times n_m(\tau_i+1)]\times 1}, \tag{B.58}$$

$$\mathbf{x}_{n_m}(k-\tau) = [x_{1_m}(k-1), \cdots,\, x_{i_m}(k-\tau_i)]^T \in \mathbb{R}^{n(\tau_i+1)\times 1}, \tag{B.59}$$

$$\Delta\mathbf{y}(k) = \mathbf{y}(k) - \mathbf{y}(k-1) \in \mathbb{R}^{[n\times n_m(\tau_i+1)]\times 1}, \tag{B.60}$$

Following an analogous procedure to what was previously presented, one can develop for a generic prediction horizon, H_p, as follows,

$$\mathcal{P}(k) = \mathbf{R}_{x_\tau}\mathbf{p}(k) + \mathbf{R}_{u_\tau}\gamma(k) + \mathbf{R}_{d_\tau}\mathcal{U}_\tau(k), \tag{B.61}$$

where

$$\mathbf{R}_{x_\tau} = \left[\mathbf{C}_\tau\mathbf{A}_\tau,\, \mathbf{C}_\tau\mathbf{A}_\tau^2,\, \cdots,\, \mathbf{C}_\tau\mathbf{A}_\tau^{Hp}\right]^T, \tag{B.62}$$

$$\mathbf{R}_{u_\tau} = \begin{bmatrix} \mathbf{C}_\tau\mathbf{B}_{u_\tau} & 0 & 0 & \cdots & 0 \\ \mathbf{C}_\tau\mathbf{A}\mathbf{B}_{u_\tau} & \mathbf{C}_\tau\mathbf{B}_{u_\tau} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_\tau\mathbf{A}_\tau^{Hp-1}\mathbf{B}_{u_\tau} & \mathbf{C}_\tau\mathbf{A}_\tau^{Hp-2}\mathbf{B}_{u_\tau} & \mathbf{C}_\tau\mathbf{A}_\tau^{Hp-3}\mathbf{B}_{u_\tau} & \cdots & \mathbf{C}_\tau\mathbf{B}_{u_\tau} \end{bmatrix}, \tag{B.63}$$

$$\mathbf{R}_{d_\tau} = \begin{bmatrix} \mathbf{C}_\tau\mathbf{\Omega}\mathbf{B}_{u_\tau} & 0 & 0 & \cdots & 0 \\ \mathbf{C}_\tau\mathbf{A}_\tau\mathbf{\Omega}\mathbf{B}_{u_\tau} & \mathbf{C}_\tau\mathbf{\Omega}\mathbf{B}_{u_\tau} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_\tau\mathbf{A}_\tau^{Hp-1}\mathbf{\Omega}\mathbf{B}_{u_\tau} & \mathbf{C}_\tau\mathbf{A}_\tau^{Hp-2}\mathbf{\Omega}\mathbf{B}_{u_\tau} & \mathbf{C}_\tau\mathbf{A}_\tau^{Hp-3}\mathbf{\Omega}\mathbf{B}_{u_\tau} & \cdots & \mathbf{C}_\tau\mathbf{\Omega}\mathbf{B}_{u_\tau} \end{bmatrix}, \tag{B.64}$$

$$\mathcal{U}_\tau = [\mathbf{u}(k-1),\, 0,\, \cdots,\, 0]^T, \tag{B.65}$$

$$\gamma = [\Delta\mathbf{y}(k),\, 0,\, \cdots,\, 0]^T, \tag{B.66}$$

Putting everything together, one can write the constraint in the quadratic-programming standard form as follows,

$$\mathbf{F_p}\mathcal{U}(k) \leq \mathbf{R}_{x_\tau}\mathbf{p}(k) + \mathbf{R}_{u_\tau}\gamma(k) + \mathbf{R}_{d_\tau} + \mathcal{U}_\tau(k), \tag{B.67}$$

or in a more compact form,

$$\mathbf{F_p}\mathcal{U}(k) \leq \mathcal{P}(k). \tag{B.68}$$

B.4 Assignment constraints

To better understand the definition of assignment constraints the following example. Assume there is one commodity type, P1, that can be moved by two different types of transport agents, T1 and T2, while P2 can only be transported by T2. One can then write these constraints as follows:

$$u_{ijP1} + u_{ijP2} - \lambda_{T2}u_{ijT2} \leq 0, \tag{B.69}$$

$$u_{ijP1} - \lambda_{T1}u_{ijT1} \leq 0, \tag{B.70}$$

where λ_w refers to the capacity of transport agent w. The previous system of equations can yet be written in a more compact form as follows,

$$\begin{bmatrix} 1 & 1 & -\lambda_{T1} & 0 \\ 0 & 1 & 0 & -\lambda_{T2} \end{bmatrix} \begin{bmatrix} u_{ijP1} \\ u_{ijP2} \\ u_{ijT1} \\ u_{ijT2} \end{bmatrix} \leq \mathbf{0}, \tag{B.71}$$

$$\mathbf{F_a}\,\mathcal{U} \;\leq\; \mathbf{f_a}. \tag{B.72}$$

Thus, for a generic assignment constraint given as

$$\sum_{m_P} u_{ij_{m_P}}(k) \leq \lambda\, u_{ij_{m_T}}(k),$$

one has simply to choose appropriate $\mathbf{F_a}$ and $\mathbf{f_a}$ matrices.

B.5 Transformation constraints

Recall that transformation constraints were previously defined as follows,

$$|u_{s_i}(k) - u_{s_i'}(k)|_{C_2^N} \leq 0, \tag{B.73}$$

to represent the set of combinatorial conditions of the form $|u_{s_i}(k) - u_{s_i'}(k)| \leq 0$ that would be necessary for the flow of all complementary raw materials to simultaneously assume the same value.

Making use of a simple example as means of illustration, consider the situation where there are only two raw complementary materials, u_{s_i} and $u_{s_i'}$. In such a case, the previous definition yields,

$$u_{s_i}(k) - u_{s_i'}(k) \leq 0, \tag{B.74}$$

$$-u_{s_i}(k) + u_{s_i'}(k) \leq 0. \tag{B.75}$$

For a generic system, one can write transformation constraints in a more compact form as follows,

$$\mathbf{F_t}\,\mathcal{U} \;\leq\; \mathbf{f_t}, \tag{B.76}$$

where $\mathbf{F_t}$ and $\mathbf{f_t}$ are matrices of appropriate dimensions.

B.6 Scheduling constraints

Analogously to transformation constraints one can expand the scheduling constraint definition given as,

$$\left[u_{s_{i_m}}(k) + u_{s_{i_n}}(k)\right]\Big|_{\forall\, m,n} \leq 0 \tag{B.77}$$

to a more compact form

$$\mathbf{F_s}\,\mathcal{U} \;\leq\; \mathbf{f_s}, \tag{B.78}$$

where $\mathbf{F_s}$ and $\mathbf{f_s}$ are matrices of appropriate dimensions.

However, since transformation and scheduling constraints are so analogously defined and refer to manufacturing activities, it is suggested both constraints be concatenated into the same matrices yielding,

$$\mathbf{F_t}\,\mathcal{U} \;\leq\; \mathbf{f_t}. \tag{B.79}$$

B.7 Conclusion

Assembling the constraint equations, one can express the constraints in its matrix form as follows,

$$\begin{bmatrix} \mathbf{S}\,\mathbf{R}_{u_y} \\ \mathbf{G}\,\mathbf{R}_{\mathbf{u}} \\ \mathbf{F} \\ \mathbf{F}_{\mathbf{p}} \\ \mathbf{F}_{\mathbf{a}} \\ \mathbf{F}_{\mathbf{t}} \end{bmatrix} \mathcal{U}(k) \leq \begin{bmatrix} \mathbf{s} - \mathbf{S}\ \left(\mathbf{R}_{x_y}\mathbf{x} + \mathbf{R}_{d_y}\mathcal{D}\right) \\ \mathbf{g}\ -\ \mathbf{G}\ \left(\mathbf{R}_{x}\mathbf{x} + \mathbf{R}_{d}\mathcal{D}\right) \\ \mathbf{f} \\ \mathbf{P}_{\mathbf{x}\mathbf{u}_{\mathbf{p}}}\ \left(\mathbf{R}_{x_y}\mathbf{x} + \mathbf{R}_{d_y}\mathcal{D}\right) \\ \mathcal{P}(k) \\ \mathbf{f}_{\mathbf{a}} \\ \mathbf{f}_{\mathbf{t}} \end{bmatrix}, \tag{B.80}$$

or even more compactly,

$$\mathbf{M}\,\mathcal{U}(k) \leq \mathbf{\Lambda}. \tag{B.81}$$

Appendix C

Conditions for optimal solution

In an optimisation problem formulated as a quadratic programming (QP) problem, the uniqueness of solution depends on the Hessian matrix being at least positive semi-definite [44]. The present chapter presents the conditions in which the devised modelling approach results in QP formulations with only one optimal solution. Section C.1 opens with important definitions and mathematical results, which are then used in Section C.2 to analyse the resulting models of the proposed approach and determine the necessary conditions for a unique optimal solution to be verified.

C.1 Definitions

Definition C.1.1. *A matrix* $\mathbf{X} \in \mathbb{R}^{n\times n}$ *is said to be symmetric, if and only if the following relationship holds true,*

$$\mathbf{X} = \mathbf{X}^T. \tag{C.1}$$

Definition C.1.2. *A matrix* $\mathbf{X} \in \mathbb{R}^{n\times n}$ *is said to be diagonal, if the entries outside its main diagonal are all zero.*

Definition C.1.3. *A matrix* $\mathbf{X} \in \mathbb{R}^{n\times n}$ *is said to be positive definite, if and only if the following relationship holds true,*

$$\mathbf{v}^T\mathbf{X}\mathbf{v} > \mathbf{0}, \quad \forall \mathbf{v} \in \mathbb{R}^{n\times 1}. \tag{C.2}$$

Definition C.1.4. *A matrix* $\mathbf{X} \in \mathbb{R}^{n\times n}$ *is said to be positive semi-definite, if and only if the following relationship holds true,*

$$\mathbf{v}^T\mathbf{X}\mathbf{v} \geq \mathbf{0}, \quad \forall \mathbf{v} \in \mathbb{R}^{n\times 1}. \tag{C.3}$$

Proposition C.1.1. *The sum of symmetric matrices, results in a symmetric matrix.*

Proof. Let $\mathbf{E}$, and $\mathbf{F}$ be two arbitrary matrices $\in \mathbb{R}^{2\times 2}$. Then,

$$\mathbf{E} + \mathbf{F} = \begin{bmatrix} a & b \\ b & c \end{bmatrix} + \begin{bmatrix} d & e \\ e & f \end{bmatrix} = \begin{bmatrix} a+d & b+e \\ b+e & c+f \end{bmatrix} \tag{C.4}$$

□

Proposition C.1.2. *The product of a symmetric matrix by a constant different than* 0*, results in a symmetric matrix.*

Proof. Letting $\alpha \in \mathbb{R}\backslash 0$, and $\mathbf{T} \in \mathbb{R}^{2\times 2}$, yields

$$\alpha \mathbf{T} = \alpha \begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} \alpha a & \alpha b \\ \alpha b & \alpha c \end{bmatrix}. \tag{C.5}$$

□

Proposition C.1.3. *If* $\mathbf{W}$ *is a diagonal matrix, and* $\mathbf{T}$ *is an arbitrary matrix of appropriate size, then* $\mathbf{T}^T\mathbf{W}\mathbf{T}$ *results in a symmetric matrix.*

Proof. Let $\mathbf{W}$ and $\mathbf{T}$ be defined as follows,

$$\mathbf{W} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \in \mathbb{R}^{2\times 2}, \mathbf{T} = \begin{bmatrix} \alpha & \beta \\ \theta & \omega \end{bmatrix} \in \mathbb{R}^{2\times 2}. \tag{C.6}$$

Then,

$$\mathbf{T}^T\mathbf{W}\mathbf{T} = \begin{bmatrix} \alpha & \beta \\ \theta & \omega \end{bmatrix}^T \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \theta & \omega \end{bmatrix} = \begin{bmatrix} \alpha a & \theta b \\ \beta a & \omega b \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \theta & \omega \end{bmatrix} = \tag{C.7}$$

$$= \begin{bmatrix} \alpha^2 a + \theta^2 b & \alpha\beta a + \omega\theta b \\ \alpha\beta a + \omega\theta b & \beta^2 a + \omega^2 b \end{bmatrix} \tag{C.8}$$

□

Theorem C.1.1. *If* $\mathbf{T}$ *is a symmetric matrix given as follows,*

$$\mathbf{T} = \begin{bmatrix} \lambda & \epsilon \\ \epsilon & \tau \end{bmatrix} \in \mathbb{R}^2, \tag{C.9}$$

where $|\lambda| \geq |\tau| \geq |2\epsilon| \vee |\tau| \geq |\lambda| \geq |2\epsilon|$*, then* $\mathbf{T}$ *is positive definite.*

Proof. Let $\mathbf{v}$ be an arbitrary non-null vector, defined as $\mathbf{v} = [x, y]^T$. Then,

$$\mathbf{v}^T\mathbf{T}\mathbf{v} = [x, y]^T \begin{bmatrix} \lambda & \epsilon \\ \epsilon & \tau \end{bmatrix} [x, y] = x^2\lambda + 2xy\epsilon + y^2\tau. \tag{C.10}$$

Without loss of generality, assuming $\lambda \geq \tau$, and $|x| \geq |y|$ yields,

$$\lambda(x^2 + y^2) + 2|x||y||\epsilon| \geq x^2\lambda + 2|x||y||\epsilon| + y^2\tau. \tag{C.11}$$

For $|\lambda| \geq |2\epsilon|$, it follows that,

$$0 \leq |x||y||2\epsilon| \leq x^2|\lambda| < (x^2 + y^2)|\lambda| \Rightarrow \tag{C.12}$$

$$\Rightarrow |\lambda|(x^2 + y^2) - 2|x||y||\epsilon| > 0. \tag{C.13}$$

□

Corollary C.1.1.1. *If $\mathbf{T}$ is a symmetric matrix satisfying Theorem C.1.1, and $\mathbf{W}$ is a diagonal matrix of appropriate dimensions, then $\mathbf{X} = \mathbf{T} + \mathbf{W}$ is:*

1. *positive definite if and only if $|a| > -(|\lambda| + |\epsilon|)$,*

2. *positive semi-definite if and only if $|a| \geq -(|\lambda| + |\epsilon|)$,*

where a is the greatest element in the diagonal of matrix $\mathbf{W}$, and λ and ϵ are the greatest elements in and outside the diagonal of matrix $\mathbf{T}$, respectively.

Proof. Let $\mathbf{T}$ and $\mathbf{W}$ be two arbitrary matrices $\in \mathbb{R}^{2\times 2}$, and $\mathbf{v}$ a generic non-null vector $\in \mathbb{R}^{2\times 1}$ defined as follows,

$$\mathbf{T} = \begin{bmatrix} \lambda & \epsilon \\ \epsilon & \tau \end{bmatrix}, \mathbf{W} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \mathbf{v} = [x, y]^T, \tag{C.14}$$

where $|\lambda| \geq 2|\epsilon|$.

Then, $\mathbf{v}^T\mathbf{X}\mathbf{v}$ yields,

$$\mathbf{v}^T\mathbf{X}\mathbf{v} = [x, y]^T \begin{bmatrix} \lambda + a & \epsilon \\ \epsilon & \tau + b \end{bmatrix} [x, y] = \tag{C.15}$$

$$= x^2(\lambda + a) + 2xy\epsilon + y^2(\tau + b). \tag{C.16}$$

Without loss of generality, one can take $|x| \geq |y|$, $|\lambda| \geq |\tau|$ yielding,

$$2x^2[|\lambda| + |a|] + 2x^2|\epsilon| \geq x^2(|\lambda| + |a|) + 2|x||y||\epsilon| + y^2(|\tau| + |b|) \geq 0 \Rightarrow \tag{C.17}$$

$$\Rightarrow |\lambda| + |a| + |\epsilon| \geq 0 \Rightarrow \tag{C.18}$$

$$\Rightarrow |a| \geq -(|\lambda| + |\epsilon|). \tag{C.19}$$

In turn,

$$2x^2[|\lambda| + |a|] + 2x^2|\epsilon| > 0 \Rightarrow |a| > -(|\lambda| + |\epsilon|). \tag{C.20}$$

□

Corollary C.1.1.2. *If $\mathbf{W}$ is a diagonal matrix, in which all entries are positive, and $\mathbf{T}$ is an arbitrary matrix of appropriate dimension, then $\mathbf{T}^T\mathbf{W}\mathbf{T}$ is at least positive semi-positive.*

Proof. It follows directly from Proposition C.1.3 and Theorem C.1.1. □

C.2 System analysis

Recall that the optimisation problem at hand is defined as follows,

$$\min_{\tilde{u}} J = \frac{1}{2}\,\mathbf{u}^T\mathbf{H}\mathbf{u} \;+\; \mathbf{c}^T\mathbf{u}, \tag{C.21}$$

$$\mathbf{M}\,\mathbf{u} \;\leq\; \mathbf{\Lambda}, \tag{C.22}$$

where

$$\mathbf{H} = 2\left[\mathbf{R}_u^T \mathcal{Q}_Q\,\mathbf{R}_u \;+\; \mathcal{Q}_R\right], \tag{C.23}$$

$$\mathbf{R_u} = \begin{bmatrix} \mathbf{C}_y\mathbf{B}_u & 0 & 0 & \cdots & 0 \\ \mathbf{C}_y\mathbf{A}\mathbf{B}_u & \mathbf{C}_y\mathbf{B}_u & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_y\mathbf{A}^{Hp-1}\mathbf{B}_u & \mathbf{C}_y\mathbf{A}^{Hp-2}\mathbf{B}_u & \mathbf{C}_y\mathbf{A}^{Hp-3}\mathbf{B}_u & \cdots & \mathbf{C}_y\mathbf{B}_u \end{bmatrix}, \tag{C.24}$$

$$\mathbf{A} = diag(a_1, \cdots, a_m) \in \mathbb{R}^{m\times m}, \tag{C.25}$$

$$\mathbf{B}_u \in \mathbb{R}^{m\times m-1}, \tag{C.26}$$

$$\mathbf{C} = \mathbf{I} \in \mathbb{R}^{m\times m}, \tag{C.27}$$

$$\mathcal{Q}_Q = diag(q_1, \cdots, q_{2m}) \in \mathbb{R}^{2m\times 2m}, \tag{C.28}$$

$$\mathcal{Q}_R = diag(\rho_1, \cdots, \rho_{2m}) \in \mathbb{R}^{2m\times 2m}. \tag{C.29}$$

Since the problem is formulated in the quadratic programming form, and the constraints are written in the form of inequalities, a necessary and sufficient condition for the uniqueness of optimal solution to be guaranteed is that the Hessian matrix, $\mathbf{H}$, must be positive semi-definite [44]. In turn, $\mathbf{H}$ explicitly depends on matrices $\mathbf{A}$, $\mathbf{B}_u$, $\mathbf{C}$, $\mathcal{Q}_Q$ and $\mathcal{Q}_R$. Thus, the following paragraphs will focus on determining in which conditions can one guarantee a unique optimal solution for the problem at hand.

Letting $\mathbf{A} = diag(a, b) \neq \mathbf{0}$, $\mathbf{B} = [c, d]^T \neq \mathbf{0}$, and $\mathcal{Q}_Q = diag(g, h, i, j) \neq \mathbf{0}, \forall_{a,b,c,d,g,h,i,j} \in \mathbb{R}$ yields,

$$\mathbf{R}_u^T \mathcal{Q}_Q\,\mathbf{R}_u = \begin{bmatrix} cd & ac & bd & \\ 0 & 0 & c & d \end{bmatrix} \begin{bmatrix} g & 0 & 0 & 0 \\ 0 & h & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & j \end{bmatrix} \begin{bmatrix} c & 0 \\ d & 0 \\ ac & c \\ bd & d \end{bmatrix} \tag{C.30}$$

$$= \begin{bmatrix} c^2g + d^2h + a^2c^2i + b^2d^2j & ac^2i + bd^2j \\ ac^2i + bd^2j & c^2i + d^2j \end{bmatrix} \tag{C.31}$$

According to Theorem C.1.1, for $\mathbf{R}_u^T \mathcal{Q}_Q\,\mathbf{R}_u$ to be positive definite it suffices that $\mathcal{Q}_Q$ be a diagonal matrix in which all non-null entries are positive. Furthermore, it follows directly from Corollaries C.1.1.1 and C.1.1.2 that for diagonal matrices $\mathcal{Q}_Q$ and $\mathcal{Q}_R$ defined such that all non-null entries are positive, then $\mathbf{H}$ is positive, regardless of $\mathbf{A}$ and $\mathbf{B}$. Therefore, a sufficient condition to guarantee the uniqueness of optimal solution is that $\mathcal{Q}_Q$ and $\mathcal{Q}_R$ must be diagonal matrices in which all non-null entries are positive.