Pontifícia Universidade Católica
do Rio de Janeiro

Mateus da Rocha Peixoto

Data-Driven & Cognitively Aware Supply Chain Management

Dissertação de Mestrado

Dissertation presented to the Programa de Pós–graduação em Engenharia de Produção of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia de Produção.

Advisor: Prof. Fernanda Araujo Baião Amorim

Rio de Janeiro
February 2025

Pontifícia Universidade Católica
do Rio de Janeiro

Mateus da Rocha Peixoto

Data-Driven & Cognitively Aware Supply Chain Management

Thesis presented to the Programa de Pós–graduação em Engenharia de Produção of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia de Produção. Approved by the Examination Committee:

Prof. Fernanda Araujo Baião Amorim
Advisor
Departamento de Engenharia Industrial – PUC-Rio

Prof. Renata Silva Sousa Guizzardi
University of Twente

Prof. Paula Medina Maçaira Louro
Departamento de Engenharia Industrial – PUC-Rio

Rio de Janeiro, February 7th, 2025

Mateus da Rocha Peixoto

Graduou-se em Engenharia de produção pela Pontifícia Universidade Católica (Rio de Janeiro) com domínio adicional em Análise de decisão e risco.

Bibliographic Data

da Rocha Peixoto, Mateus

Data-Driven & Cognitively Aware
Supply Chain Management / Mateus da Rocha Peixoto; advisor: Fernanda Araujo Baião Amorim. – 2025.

80 80 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2025.

Inclui bibliografia

1. Engenharia de Produção – Teses. 2. Engenharia Industrial – Teses. 3. Gestão de cadeia de suprimentos. 4. Incerteza dependente de decisão. 5. Otimização robusta. 6. Tomada de decisão multicritério. 7. Ajuste baseado em experiência. 8. Viés cognitivo. 9. Viés cognitivo de propensão de risco. I. Baião Amorim, Fernanda Araujo. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. III. Título.

CDD: 004

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoaento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

Gostaria de expressar a minha gratidão à PUC Rio e o CNPQ pelas bolsas fornecidas, tornando possível a dedicação total às atividades do mestrado. Estendo minha gratidão à CAPES pelo programa PRINT que financiou a minha ida para a universidade de Twente, Holanda.

Agradeço profundamente a minha orientadora, Fernanda Baião, pelo suporte incansável me ajudando a tornar em objetos concretos as ideias que surgram ao longo do mestrado. Gostaria de agradecer também aos professores do departamento de engenharia de produção que tiveram contribuições para a confecção da dissertação através dos trabalhos de suas disciplinas, em ordem cronológica são: Bruno Fânzeres, Luiz Fernando Cyrino, Rodrigo Caiado, Paula Maçaira, Antônio Márcio Thomé. Estendo também minha gratidão à professora Luciana Pessoa pelo o suporte com o programa do CAPES PRINT.

Faltam palavras para expressar a gratidão devida aos meus pais por todo o suporte que me deram de múltiplas formas ao longo desses dois anos.
Em primeiro lugar agradeço a Deus pela capacitação concedida, este documento é uma manifestação de Provérbios 16.3 "Dedique ao Senhor o trabalho de suas mãos e Ele estabelecerá os seus planos." portanto dedico esta dissertação ao Senhor Jesus Cristo.

Abstract

da Rocha Peixoto, Mateus; Baião Amorim, Fernanda Araujo (Advisor). **Data-Driven & Cognitively Aware Supply Chain Management**. Rio de Janeiro, 2025. 80p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Forecasting is of extreme importance for companies as it is the input for the S&OP process, an essential part of Supply Chain Management (SCM); however, considering the close involvement of humans in various moments that compose this activity, cognitive biases (e.g., risk-seeking) and their influence represent a threat to an organization's performance, with many potential risks to supply chains. This dissertation establishes a novel framework for data-driven and cognitively aware supply chain management. Multiple forecasting models were evaluated using the proposed framework to select the most satisfactory model considering the organization's strategic vision. This allows the SCM manager to perform judgmental adjustments, which are evaluated through an automatic risk-seeking bias detection system. The dissertation was experimentally assessed over simulated scenarios from real data about cardboard production during 2017 and 2023. The results evidenced the effectiveness of the proposal in addressing the complexity and intertwined objectives of different stakeholders within a supply chain. The automatic definition of a preference-based rational model defined by the SC manager, made it possible to detect risk-seeking biases using different thresholds for judgmental adjustments, thus mitigating the adverse effects of risk-seeking biases. In summary, it can be argued that the proposed framework represents an important advance regarding the implementation of the "Humachine" paradigm, integrating the positive elements of advanced statistical modeling with human expertise and context.

Keywords

Supply chain management; Decision-dependent uncertainty; Robust optimization; Multi-criteria decision-making; Judgmental adjustment; Cognitive bias; Risk-seeking cognitive bias.

Resumo

da Rocha Peixoto, Mateus; Baião Amorim, Fernanda Araujo. **Gestão de Cadeia de Suprimentos Orientada a Dados e Cognitivamente Consciente**. Rio de Janeiro, 2025. 80p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Previsão de demanda é de extrema importância para as empresas, pois serve como insumo para o processo de vendas e planejamento de operações (S&OP), uma parte essencial da Gestão de Cadeia de Suprimentos (SCM); entretanto, considerando o envolvimento próximo de humanos em diversos momentos que compõem essa atividade, os vieses cognitivos (e.g. propensão a risco) e sua influência representam uma ameaça ao desempenho organizacional, acarretando diversos riscos potenciais às cadeias de suprimentos. Esta dissertação estabelece um novo framework para a gestão da cadeia de suprimentos baseada em dados e ciente de aspectos cognitivos. Múltiplos modelos de previsão foram avaliados utilizando o framework proposto, de forma a selecionar o modelo mais satisfatório considerando a visão estratégica da organização. Isso permite que o gerente de SCM realize ajustes subjetivos, que são avaliados por meio de um sistema automatizado de detecção de viés de busca por risco. A dissertação foi experimentalmente avaliada em cenários simulados a partir de dados reais sobre a produção de papelão ondulado entre 2017 e 2023. Os resultados evidenciaram a eficácia da proposta em abordar a complexidade e os objetivos interdependentes dos diferentes stakeholders dentro de uma cadeia de suprimentos. A definição automática de um modelo racional baseado em preferências, definidas pelo gerente da cadeia de suprimentos, possibilitou a detecção de vieses de propensão a risco utilizando diferentes níveis de ajustes baseados em experiência, mitigando assim os efeitos adversos do viés cognitivo. Em síntese, pode-se argumentar que o framework proposto representa um avanço importante na implementação do paradigma “Humachine”, ao integrar os elementos positivos da modelagem estatística avançada com a expertise e o contexto providos por agentes humanos.

Palavras-chave

Gestão de cadeia de suprimentos; Incerteza dependente de decisão; Otimização robusta; Tomada de decisão multicritério; Ajuste baseado em experiência; Viés cognitivo; Viés cognitivo de propensão de risco.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ARSBD – Automatic Risk-Seeking Bias Detection
ARSBDS – Automatic Risk-Seeking Bias Detection System
BOM – Behavioral Operations Management
BSCM – Behavioral Supply Chain Management
BWE – Bullwhip Effect
BWM – Best Worst Method
CPM – Cumulative Prospect Theory
DDU – Decision-Dependent Uncertainty
DM – Decision Maker
DPM – Dual Process Modeling
DSS – Decision Support System
FPS – Finished Product Stock
FVA – Forecast Value Added
JA – Judgmental Adjustment
MCDM – Multicriteria Decision-Making
RMS – Raw Materials Stock
RO – Robust Optimization
RS – Risk-Seeking
RSC – Reverse Supply Chain
SC – Supply Chain
SCF – Supply Chain Forecasting
SCM – Supply Chain Management
SKU – Stock Keeping Unit
SL – Service Level
SPM – Single Process Modeling
S&OP – Sales and Operations Planning
OTD – On-Time Delivery
PC – Preferential Client

1
Introduction

Disruptive events such as Corona Virus Disease 2019 (COVID-19) affected global economics and society and caused great disturbance in demand forecasting. Supply Chain Management (SCM) - and in particular its component Sales & Operations Planning S&OP) process - links customers, manufacturers, and suppliers, and essentially deals with demand forecasting, and therefore is highly susceptible to disruption risks, which in turn cause impacts across several dimensions, including financial, lead time, demand changes, production, and performance (Moosavi et al., 2022). With the world plunged into uncertainty during the COVID-19 pandemic, a critical issue for senior management across organizations is stabilizing their supply chain to a consistent flow of components and materials.

Fairly, the complexity of SCM had been an increasingly critical topic for both academic and industrial stakeholders, even before the advent of the COVID-19 pandemic. The complexity of new tariff restrictions, port congestion, regional conflicts, and geopolitical events and disruptions due to international conflict is apparent, and it is clear that securing access to critical materials and resources is increasingly difficult and forecasting demand is even harder (Nikolopoulos et al., 2021). According to the Contingency Theory proposed by Sousa and Voss (2008), organizations must adapt their structures and processes to their environment, to achieve high performance. From this perspective, S&OP practices should be compatible with the manufacturing structure and environment (M.T. Thomé et al., 2014).

As a fundamental aspect of SCM, forecasting demand should ensure that businesses produce the appropriate type and volume of products, which is vital for sustaining profitability over an extended period (Wang, 2023). Demand uncertainty makes forecasting difficult and frustrating; meanwhile, managers who deal with forecasting should understand that the quality of a forecast is not determined solely by its accuracy, but also by how it is produced and used, as argued by (Kolassa et al., 2023) in the development of the Ideal Forecaster framework. In this context, forecasting is as much an organizational exercise as a statistical one, since the information is spread throughout the organization. In the same direction, the concept of Inventory Forecasting, as defined by Goltsos et al. (2022), emphasizes the necessity of considering the company strategic vision and the potential consequences of forecasting to other departments. Therefore, Supply Chain Forecasting (SCF) practitioners care about the variance of the forecast and forecast errors during the lead time and review periods, as this leads to excess inventories, the bullwhip effect (BWE), and uncertainty in decision-making (Wang and Disney, 2016). Kolassa et al. (2023) underlines that 71% of companies use judgmental input as part of their forecasting process. Therefore, the judgmental behavioral components may impact not only the business performance but also the performance of its partners.

Forecasting literature provides several examples of when judgmental adjustment led to better performance than purely using mathematical modeling (Fildes et al., 2009). Judgment and forecasting are fundamentally intertwined. Even in companies employing purely statistical forecasting methods, human judgment remains essential in selecting the appropriate statistical forecasting models, predictor variables, and datasets; in other words, the framing of the collected data belongs to the human agent (Perera et al., 2019). This underscores the undeniable role of human judgment in the forecasting process (Lawrence et al., 2006; Petropoulos et al., 2018; Perera et al., 2019).

1.1 Motivation and Research questions

Despite its practical importance, academic research in SCF, particularly regarding the interplay of human judgement with data-aware computational modeling, has often overlooked certain critical aspects and complexities. Even when there are robust theoretical developments, they are seldom translated into operational solutions or incorporated into cutting-edge decision support systems (DSS) (Syntetos et al., 2016). For example, the way information is presented can impact significantly upon the decisions of forecasting model selection, as shown by Reimers and Harvey (2024).

Human decision-makers bring their judgment, individual personalities, opinions, and biases to the decision-making process, and they are responsible for deciding how to use any analytically generated output. The introduction of human input brings along the benefits of context and creativity (Sanders and Wood, 2019), however, as evidenced by Behavioral supply chain mamangement (BSCM) literature, the complexity of SCM is often overwhelming, and human agents are unable to understand it fully (Fahimnia et al., 2019; Yang et al., 2021; Brauch et al., 2024). This amplifies the risks of distortions caused by their inherent cognitive biases (Sanders and Wood, 2019).

This reinforces the necessity of cutting-edge DSS, as defended by Syntetos et al. (2016), to fill the gap in efficiently combining forecasting model selection with judgmental adjustment. Current techniques are primarily concerned with accuracy and parsimony, disregarding operational constraints and the effects of human stakeholder decisions, since the forecast has many stakeholders who require it as input to their planning processes (Kolassa et al., 2023). Thus, an integrated view of forecasting with the company's strategic goals is crucial (Goltsos et al., 2022; Kolassa et al., 2023).

This work approaches SCM from the Humachine paradigm, defined by Sanders and Wood (2019) as the ideal incorporation of the beneficial aspects of human expertise with machine learning and advanced statistical modeling. In this direction, this dissertation comprises human and machine components, addressing specific limitations from both perspectives. From the human perspective, the proposal addresses scalability and the presence of cognitive biases during decision-making (Sanders and Wood, 2019)). From the machine perspective, the proposal addresses the lack of context of strict computer-based forecasting models (Sanders and Wood, 2019).

Our research questions are as follows.

1. How to increase the scalability of the human expert for a supply chain forecasting system?

2. How to avoid the impacts of cognitive bias in human judgmental adjustments in supply chain forecasting?

3. How to integrate a more holistic organizational context with supply chain forecasting supported by data-aware statistical models?

1.2 Objective

The main objective of this dissertation is to propose a methodology for supporting scalable supply chain forecasting that integrates mathematical modeling with human expertise while preventing cognitive biases by automatic detection.

This makes it relevant to analyze not only the characterization of uncertainty provided by each forecast, but also the consequences of the decision-making process using each forecasting model. For that, the methodology comprises a simulation step to approximate the decision-making using each forecasting model. The models are ranked based on their performance taking into account previously customized weights for each organizational dimension, and the most satisfactory is considered as the rational model.

1.3 Document structure

This dissertation is structured as follows. Chapter 2 provides foundations, defining the main concepts necessary for understanding the proposal, and presents the state of the art in the addressed topic. Chapter 3, Proposal, describes the proposed framework and subjacent methodology. Chapter 4, Results, presents and discusses the results of applying the proposal. Finally, Chapter 5, Contributions, summarizes the scientific contributions of the research.

2
Foundations and State of the Art

This section comprises a brief contextualization of the relevant topics used in this research, as well as a literature review of the state of the art in the literature for the problems addressed.

2.1
Conceptualization

This brief section will detail the most relevant theoretical concepts and methodologies adopted as inspiration for the proposed solution.

The BWE is a major problem in SC logistics since it distorts the information flow and affects subsequent entities (Wang and Disney, 2016) related to the system as they tend to make sub-optimal decisions, further propagating the distortion (Keliji et al., 2022). Therefore, there is a necessity to identify and mitigate biases within decision-making, to achieve this, the Humachine framework was adopted.

An ideal balance could be met by incorporating the beneficial aspects of human expertise and machine learning/advanced statistical modeling. The machine intelligence, despite its precision and scalability, however, is only as capable as the data allows it to be, and is completely incapable of creative thinking or absorbing context (Sanders and Wood, 2019). Regarding the machine limitations, the addition of human agents to the process will already contribute to the contextual information, enabling a better comprehension of the scenario, incapable of being gathered from simple statistical modeling. Complementary to that, humans bring their inherent creativity and domain knowledge with their inputs. Therefore the focus should be on the correction of human limitations.

However it is not only the human and machine components that must be integrated, the organization as a whole should integrate its many departments, especially forecasting, as defended by the Ideal Forecaster Framework was created by Kolassa et al. (2023) as the intersession of four core skills: Statistics, programming, business knowledge, and communication, as shown in Figure 2.1

- Statistics: Forecasters must develop their statistical thinking, a forecaster needs to grasp fundamental concepts such as likelihood and probability

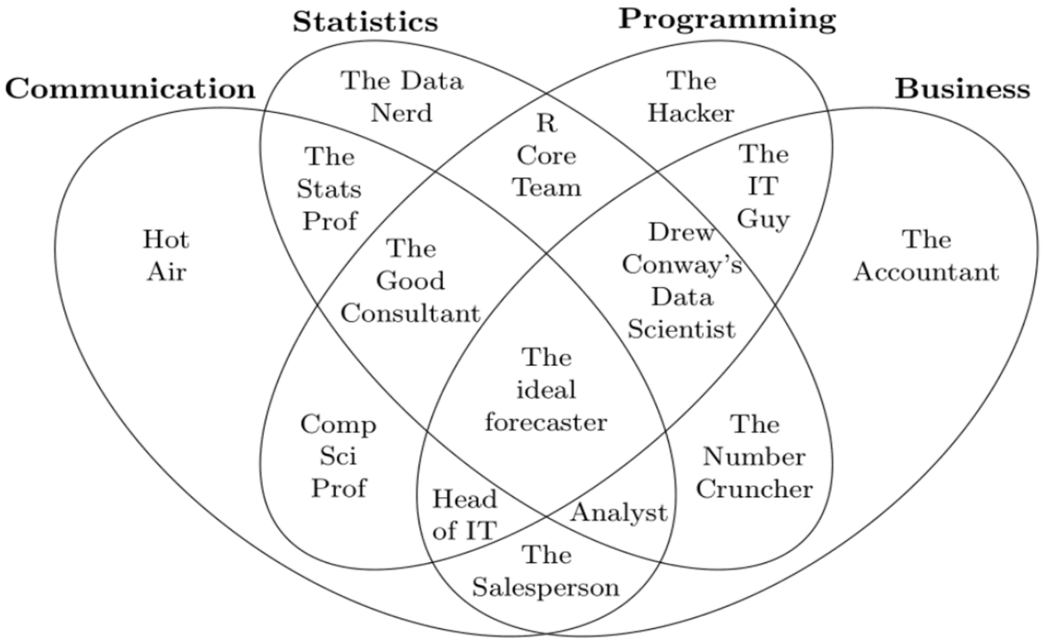

Figure 2.1: The Ideal Forecaster, adapted from Kolassa et al. (2023)

distributions. Additionally, a solid understanding of classical time series algorithms, such as ARIMA, as well as other forecasting methods, is crucial. The forecaster must always consider the bias-variance trade-off to achieve a balanced and effective model.

– Programming: Forecasting heavily relies on computing. The term "programming" in this context does not refer to traditional software development but rather to a broader proficiency in scientific computing. The specific tools a forecaster should be familiar with depend on the working environment. Familiarity with a company's proprietary forecasting software is crucial.

– Business: A forecast serves as a crucial tool for decision-making, whether it involves capacity planning for a factory or making long-term strategic decisions such as make-or-buy choices. Importantly, forecasting does not occur in a vacuum; it is deeply embedded in the business context. Often, understanding this context is more critical than possessing advanced knowledge in statistics or programming.

– Communication: Effective communication skills are indispensable for forecasters, complementing their statistical, technical, and domain expertise. Forecasters often find themselves communicating in varied circumstances with a diverse range of stakeholders, each with different levels of understanding and objectives.

Therefore, when creating or selecting a forecast, agents must comprehend the potential impacts of this decision, that being said, two crucial areas to be integrated are forecasting and inventory management. Goltsos et al. (2022) proposed the concept of Inventory Forecasting as a framework to provide this integration.

Inventory forecasting is described as the intersection of forecasting and inventory control. These works manipulate the characteristics of demand (or indeed forecasts) in pursuit of inventory and, ultimately, supply chain efficiencies. Robust optimization (RO) is among the areas described as promising for forecasting and inventory control integration. Robust optimization (RO) deals with uncertain variables by only looking at intervals without a need for further distributional information.

2.2 Related works

Supply Chain Management encompasses the planning and managing of all activities involved in sourcing and procurement, conversion, and all logistics management activities. Importantly, it also includes coordination and collaboration with channel partners, which can be suppliers, intermediaries, third-party service providers, and customers. Supply chain management integrates supply and demand management within and across companies. It is an integrated function with primary responsibility for linking major business functions and processes within and across companies into a cohesive, high-performing business model. There must be alignment between each firm's supply chain strategy and those of its supply chain partners, both internal and external. Thus, supply chain alignment results in a fit in terms of objectives, structures, and processes within and between different functions and members of a supply chain (Wong et al., 2012).

An optimal decision made by one supply chain member may well cause delivery delays and excessive inventories in another part of the supply chain, as demonstrated by Skipworth et al. (2015). Therefore, aligning business strategic views is of critical importance. Information sharing is also crucial to ensure the active collaboration between entities, however, according to Wong et al. (2012) it is rarely extended between first and second-tier suppliers. Houlihan (1985) proposed a definition of a supply chain that takes a more holistic view: SCM is about addressing the imbalances due to conflicting objectives in marketing, sales, manufacturing, and distribution by managing the trade-offs between supply policies, the economics of manufacturing, and complexity. Supply chains possess an inherent relation to trade-offs and multi-criteria aspects as exposed in Neto and Salomon (2022) modeling, by combining perspectives presented a deeper notion SC complexity can be met, the supply chain could also be approximately modeled as various agents acting in interconnected relationships (Lee and Kim, 2008). The relations between different entities among supply chains are not always equal, despite the advancements in agent modeling, there are several occurrences of both power or trust asymmetries that deeply influence how firms relate to one another as Beal Partyka (2022) demonstrated. Due to these asymmetries, managing the relations with preferential clients becomes paramount, since it defines the deterministic peak demand in the initial stage of fractional supply (Hüttinger et al., 2012). Therefore, as some members of the SC exert more influence than others, total collaboration becomes an idealistic solution for real-world problems as Thomas and Skinner (2010) proposed.

Dealing effectively with uncertainty in supply chain management poses a key challenge for companies, whether arising from external factors like natural disasters, political upheavals, major public health crises, or internal factors such as product quality issues and capacity constraints. The occurrence of uncertainty in supply chain management inevitably leads to increased risk and potential losses.

Thus, it is imperative for enterprises to devise effective strategies to address and mitigate these uncertainties, minimizing losses and maintaining operational continuity. The usage of blockchain technologies proved itself useful for mitigating specific disruptions, however, it does not apply to every situation and is dependent on the technological maturity level of the involved partners (Alkhudary et al., 2024). Despite technological advancements, there is no one-size-fits-all solution.

2.2.1 The role of the S&OP process

A well-executed S&OP process facilitates the sharing of crucial information related to demand forecasting. This process emphasizes the importance of marketing and sales teams communicating details such as upcoming product launches, new customer acquisitions, planned promotions, and other relevant data to those responsible for forecasting. Simultaneously, other departments like operations must provide essential input, such as inventory levels, available capacity, and similar data, to support the planning process. The output is a set of coordinated plans derived from the same input data: marketing creates a strategy for promotions and demand management, operations formulate production and procurement plans, finance develops cash flow projections and aligns them with other departments to communicate with investors and human resources drafts a personnel plan based on the same forecast data (Kolassa et al., 2023).

Supply chain management and S&OP are deeply interlinked since S&OP is a crucial tool for the coordination and alignment across vertical and cross-functional areas, the integration and continuous refinement of plans, along with the horizontal alignment across the supply chain involving both customers and suppliers. Several studies focused on operational improvements in specific areas, such as forecast, inventory management, the balance of the mix, the volume of products, and capacity resources. Some have concentrated on enhancing operations in specific areas, including forecasting, inventory management, balancing product mix and volume, and optimizing capacity resources. Trade-offs are also a frequent and inevitable occurrence, maximizing profits or customer satisfaction at minimum inventory and supply chain costs. However, the integration isn't always ideal, as shown by Tavares Thomé et al. (2012) only a few studies have addressed the integration of marketing practices like yield management and dynamic pricing, highlighting the limited or absent incorporation of financial goals and plans into S&OP practices.

S&OP typically operates on a monthly cycle within organizations, aiming to improve information sharing and plan coordination. It involves a cross-functional team, often including representatives from marketing, operations, finance, human resources, and a dedicated forecasting team if available.

2.2.2 Behavioral Operations Management

Predicated upon the works of Tversky and Kahneman (1974), behavioral operations management (BOM) takes the decision-maker as an agent of bounded rationality, and therefore it has a more empirical emphasis, focusing upon theory testing rather than on theory development as described by Moritz et al. (2014). Behavioral operations research is the study of attributes of human behavior and cognition that impact the design, management, and improvement of operating systems, and the interaction between such attributes and operating systems and processes. (Fahimnia et al., 2019).

One of BOM's key lines of research is developing mathematical models related to bounded rationality, however the inability to model human decision-making systematically is widespread in the literature, and the agents tend to continuously change their estimations (D'Urso et al., 2015). Behavioral operations research is the study of attributes of human behavior and cognition that impact the design, management, and improvement of operating systems, and the interaction between such attributes and operating systems and processes (Fahimnia et al., 2019).

A classical BOM approach would attempt to improve judgemental heuristics as a means of, correction (Fildes and Goodwin, 2007) and as industry 4.0 evolved more trust was placed on statistical models and DSS and theoretically it should have corrected all the problems, but human biases such as overconfidence and anchoring impeded the total collaboration between man and machine. However the objective of BOM is not to eliminate human involvement, several crucial elements in decision-making require human intervention, such as: Maintaining relations with business partners, supplying context, and the advantage of creativity to provide solutions for uncertain scenarios where statistical solutions may not be adherent to reality.

As technologies drastically evolved in Industry 4.0, new competitive advantages can be attained, however, if the operators and decision-makers cannot adapt to the new process, implementation may prove fruitless, such is the dilemma of the so-called Industry 5.0. Industry 5.0 is responsible for the integration of technology with human decision-makers and is currently the vanguard of behavioral operations research (Van Oudenhoven et al., 2023). Behavioral Supply Chain Management aims at understanding the decision-making of management using this understanding to generate interventions that improve supply chain operations (Fahimnia et al., 2019).

As BOM evolved it grew further from how the field's previous main focus of observing the impact and interactions of human agents with bounded rationality to then adopting a more holistic view, adopting aspects of SCM, since bad decisions propagate towards business partners and the emphasis on intervention, particularly present in DSS for judgmental decision making (Fahimnia et al., 2019).

Therefore, the forecasters' acceptance of the DSS may influence the information that will orient the entire operation can obtain the worst performance solely by the bias of decision makers involved in the process, making BOM essential for industry 5.0 since it will provide the tools necessary to address the problems derived from the inappropriate human integration of industry 4.0 (Fahimnia et al., 2019).

Therefore BSCM is the combination of BOM and SCM concepts, taking the objective of the "Humachine" by Sanders and Wood (2019) derived from BOM and regarding SCM, the main concern consists in understanding how the effects of the biases and inappropriate decision-making propagate thru-out the complex supply chain structure and develop mitigation strategies (Fahimnia et al., 2019).

Incentives are pivotal in molding perceptions within this domain, significantly impacting information processing, pertinent biases, and ultimately, the direction of quality advancement and its performance, therefore multiple studies confirm the impact of the framing upon the production management (Fahimnia et al., 2019), with transparency and feedback-oriented SC designs (information sharing) being a relevant contributing factor to mitigate adverse cognitive performance.

Recent emerging areas of concentration for BSM include, how SCM risks are perceived and how decision-makers respond to different types of risks when disruptions occur. Research indicates that proactive recovery actions after disruptions are more effective in stabilizing relationships, instead of reactive or deflation strategies (Reimann et al., 2017), this contributes to reducing uncertainty for the business partner, therefore the proactive information sharing could potentially avoid further propagation of the BWE, as it seems to be a consensus in the literature from the works of Wang and Disney (2016); Almeida et al. (2017); Brauch et al. (2024); Yang et al. (2021); De Almeida et al. (2015); Fahimnia et al. (2019) and Zhang (2004).

A wide range of research is conducted in the field from empirical studies to meta-theory developments, such as,Schorsch et al. (2017) proposing a meta-theory to help the BSCM field in opening up new perspectives by engaging in new and critical dialogues to grow and prosper, as the behavioral outcomes in BSCM could be described as relationship effectiveness, customer satisfaction, integration of demand and supply, and overall supply chain performance.

Ultimately, this research field aids organizations in enhancing performance by implementing robust and well-designed systems. It achieves this by considering relevant behavioral factors that influence practitioners' decisions, particularly during periods of uncertainty, when data-driven decision-making and statistical models may prove insufficient. (Goudarzi et al., 2023; Petropoulos et al., 2022)

Supply chain disruptions are closely linked to the bullwhip effect (BWE) (Smith and Fatorachian, 2023), the relation can be observed from the definition given by Monroe (2012): It has been observed that fluctuation and distortion of information increases as it moves up the supply chain, from retailers, manufacturers, to suppliers. This is called the bullwhip effect as inaccurate and distorted information travels up the chain like a bullwhip uncoiling.

It is also important to realize that information hardly will be perfectly transmitted comparing demand and it's correspondent order quantity, since operational constraints, variations in stock and eventual strategic variations on the company's DM will distort the information and the mere presence of the BWE leads to overstocking, potentially perpetuating a cycle of distortions (Zanddizari et al., 2019).

Despite being well known for almost 30 years, the bullwhip effect still represents a major challenge in SCM, the cognitive load required to fully comprehend the dynamics of a supply chain cannot be attained by humans, leading to a dissonance between reality and the mental model of the decision-makers. The analysis made by De Almeida et al. (2015) showed few studies have focused on addressing behavioral aspects to reduce the bullwhip effect. Fortunately, the literature evolved since, with significant works in this aspect such as Yang et al. (2021) and Brauch et al. (2024).

The bullwhip effect can be classified into four main causesBrauch et al. (2024):

– Inherent in the system structure
– related to uncertainty
– related to misaligned incentives
– related to inadequate cognition of the situation.

The system structure is mainly connected to lead time, defined as the duration between initiation and execution, along with inherent delays within the supply chain structure, which emerges as one of the most frequently cited causes of the bullwhip effect in the examined literature. It is argued that as the supply chain network expands from a low to a high number of echelons, the bullwhip effect demonstrates exponential growth. Various types of delays, such as material and information delays, are highlighted in the literature. System dynamics simulation suggests that an increase in material delay contributes to a rise in the bullwhip effect. Factors such as time dependencies, physical supply chain structure, complexity, feedback loops, and non-linearities are grouped under a single overarching category. This is not only due to their interconnected nature but also because they are seen as facilitators that create an environment conducive to the emergence of other causes of the bullwhip effect.

Supply chain uncertainty can stem from both, the demand and supply sides. Uncertainty is closely tied to the reality that not all pertinent information regarding future developments is consistently known, accessible, or available to all stakeholders within a supply chain. A growing concern about uncertain lead time, unpredictable stock-outs, and scarce resources are some of the triggers of ordering biases (Goudarzi et al., 2023) caused by poor information availability and quality, the lack of SC integration and information sharing is a major contributor to this aspect. Despite being able to be categorized differently, note they can be interlinked since lead time uncertainty could contribute to a cognitive or a system structure cause for the BWE.

Misaligned incentives can be summarized as a local optimization disregarding the supply chain as a whole, encompassing situations such as order batching. Once again this cause can be linked to a lack of integration, between the various participants.

Lastly, the fourth cause listed is the inadequate cognition of the situation, the elevated degree of complexity present in SCM leads to failure in human systems of thinking in totally understanding the situation analyzed and variables such as Cause-and-effect relationships, inventory, demand, and supply line information. Furthermore, irrational or sub-optimal decisions derived from the supply chain manager's bounded rationality can generate the BWE.

Regarding the last cause Yang et al. (2021) made an extensive systematic literature review contributing by analyzing the behavioral causes of the bullwhip effect, therefore demonstrating that human behaviors are a factor that cannot be ignored since human mental models are significant in dynamic decision-making. The author concluded that by studying, understanding, and analyzing mental models and the constraints to improving mental models, we can gradually alter mental models based on the enhanced understanding of cause-and-effect relationships of information feedback. Thus, we can make better decisions and optimal decisions eventually.

Sustainability is a relevant aspect in decision making in the context of BSCM (Fahimnia et al., 2019) despite the BWE being a known phenomenon since Lee et al. (1997) it has considerable aspects that were only realized recently, such as the severe sustainability importance of its mitigation in perishable products supply chains, it is a considerable challenge considering that there is the added uncertainty of the product's level of deterioration, found in the literature as one of the key contributing factors for the BWE (Durán Peña et al., 2021). The authors also stated that future research should study how the bullwhip effect affects human behavior in perishable product supply chains. Similarly, there is no research on the causes of the bullwhip effect in perishable products.

2.2.3 The role of forecasting

Despite being treated as an isolated department for the most part, the forecast is the starting point that will guide the whole S&OP process. Forecasting research frequently treats forecasting as an end goal, often overlooking the subsequent computational steps required to convert forecasts into replenishment decisions. This reflects how forecasting performance has typically been accessed only in the light of accuracy metrics, the forecasting literature has been focused on achieving gains against error metrics, with no regard for the potential operational consequences (Goltsos et al., 2022). This was reflected by Syntetos et al. (2010), which showed that depending on the accuracy metric used, a reduction of 1% in accuracy resulted in drastic inventory performance improvements by 10% and 15% reduction in costs, therefore disassociating accuracy directly with inventory performance.

Forecasting plays a crucial role in enterprises as the basis for the Sales and Operations Planning (S&OP) process, since its forecasts will serve as input for many other stakeholders. Statistical models are typically applied as a basis for forecasting and, in most cases, humanly intervened. Therefore, regardless of the forecasting model used, the decision-making process is subject to cognitive bias. Highly-experienced managers may fall into the trap of trusting their intuition and unconsciously making biased decisions, rather than carefully deliberating and reviewing all available data and options. The organization (and even the entire supply chain) may suffer the ripple effects of the biased decision, impacting the S&OP process (Ramos et al., 2022).

A complementary view is held by Kolassa et al. (2023), stating: "We treat forecasting as a predominantly statistical exercise. However, forecasting is as much an organizational exercise as a statistical one. Information is spread throughout the organization. The forecast has many stakeholders that require it as input to their planning processes. Thus, understanding demand forecasting requires understanding the statistical methods used to produce a forecast as well as how an organization creates a forecast and uses it for decision-making."

Both perceptions are perfectly in line with the literature search synthesis framework developed by Tavares Thomé et al. (2012), which demonstrates that forecasts are a critical input in the S&OP process.

Kolassa et al. (2023) further elaborates on how the integration could theoretically be achieved, however, he does not present a forecasting model selection methodology and how operational constraints and the influence of other departments should be considered by the forecasting department. The S&OP process can be viewed in terms of inputs and outputs. On the input side, a robust S&OP process facilitates the sharing of essential information, particularly demand forecasts. This requires marketing and sales teams to provide details on upcoming product launches, new customer acquisitions, planned promotions, and other relevant information to those responsible for forecasting. Concurrently, other departments, such as operations, must contribute critical data like inventory levels and available capacity.

The output of this process is a set of coordinated plans derived from the shared input. Marketing creates promotion and demand management strategies, operations develops production and procurement plans, finance prepares cash flow projections, and human resources devises staffing plans—all based on the same forecast data. An effective S&OP process ensures organizational alignment, allowing decision-makers to base their plans on comprehensive, shared information. If the process falters, functional silos can emerge, with departments withholding information or individuals skewing forecasts. This leads to inaccurate predictions, misaligned production capacities, and financial forecasts that fail to reflect actual plans, ultimately harming the organization's credibility.

There is a particular class of forecasting models that is of interest concerning the complementary aspects of human experts and statistical modeling, those would be the dynamic regression models, these models allow for the input of external variables into the time series predictions. According to Hyndman and Athanasopoulos (2018) the effects of holidays, competitor activity, changes in the law, the economy, or other external variables can explain some of the historical variation and lead to more accurate forecasts. The creativity of the human agent is paramount for the selection of candidates for exogenous variables.

As mentioned in the previous sections, forecasting is typically only concerned with accuracy, not taking into account how much it can impact the decision-making of other departments, this is in line with Goltsos et al. (2022) that stated that forecasts can affect activities such as budgeting, energy scheduling, and inventory control.

Goltsos et al. (2022) proposes a definition of Inventory forecasting as a perspective to integrate both areas, highlighting some relevant points to take into account. "We refer to a forecast as the (best possible) genuine expectation of how much demand is going to be for a particular SKU (often with sales as a proxy) Most often this refers to a point, mean demand forecasts, although forecasts of variance or higher moments, other quartiles or indeed the entire lead time demand distribution may be required. For the purposes of this article, we distinguish such forecasts of demand, to be used for SKU inventory management, from forecasts for other functions (e.g., marketing). We use the term "inventory forecasting" then to describe the intersection of these two areas, i.e. integrated literature of forecasting and inventory control. These works are manipulating characteristics of demand (or indeed of forecasts), in pursuit of inventory and ultimately supply chain efficiencies."

Forecast evaluation plays a critical role, especially when the selected forecasts are used to inform inventory management decisions. Point forecasts are evaluated using a range of accuracy metrics, but there is considerable discussion in the literature about the lack of consensus on which metrics are best. A further complication arises from the fact that different point forecasts may be optimized at different values, depending on the accuracy metric in question. Inventory costs are typically associated with service levels: 90%, 95%, and 99% are commonplace among the most used levels. Goltsos et al. (2022)

Combining the frameworks for Inventory Forecasting form Goltsos et al. (2022) and the Ideal Forecaster from Kolassa et al. (2023) as the theoretical foundation for the development of a robust forecasting model selector, taking into account the multi-faceted aspects that impact upon SCM.

Demand forecasting is critical as input for procurement, production, inventory, logistics, and overall decision-making of organizations. Various quantitative models have been developed and applied aiming to achieve more accurate product forecasts. Human judgment, either by itself or in combination with quantitative models, is a well-consolidated topic within forecasting literature. Arvan et al. (2019) Various judgment heuristics, or mental shortcuts, can be utilized in the judgmental forecasting process, 71% of companies use judgmental inputs as part of their forecasting process, and since forecasting is an organizational exercise that affects S&OP, it influences many stakeholders' decision-making process. The judgmental behavioral components can interfere with a business's performance and consequently its partners as well. Kolassa et al. (2023) The specific heuristic selected is influenced by the type of information accessible to the forecaster. In many cases, the pertinent information is solely stored in the forecaster's memory, with no possibility of direct integration into the statistical model. Petropoulos et al. (2022)

2.2.4 Judgmental adjustments

Judgment and forecasting are, by nature, deeply connected. In supply chains that rely solely on statistical forecasting techniques, human judgment still plays a critical role in choosing the appropriate models, methods, and predictors. Additionally, no matter how sophisticated the support system is, judgment cannot be eliminated from the process, as periodic performance reviews are paramount. Perera et al. (2019) judgmental forecasting has various meanings and processes within the literature, for instance: pure judgmental forecasting, judgmentally adjusting statistically derived forecasts, and combining forecasts (statistical and/or judgmental). For this dissertation the definitions adopted will be as follows: Judgmental forecasting: Refers to purely human-generated forecasts, with no statistical model involved; Judgmental adjustment: Judgmental adjustment made upon the output of statistical output, following the definition of Lawrence et al. (2006) "The judgmental adjustment of a statistical forecast comprises of two steps: first, determining whether the statistical output requires adjustment, and second, determining the size and direction of any adjustments required."; Combined Forecasts: The combination of forecasts to the context here presented will not refer to the combination of judgmental and statistical components but the combination of the two most accurate statistical models, as described by Wang et al. (2023).

Judgmental intervention should not be done without applying principles to help control for adjustments that may lead to inaccurate distortions, the three principles are described by Fildes and Goodwin (2007) as (i) when to use judgment (ii) how to use judgment and (iii) how to assess the effectiveness of judgment.

- Principle (i) states that the adjustments should be made upon quantitative approaches.
- Principle (ii) is the limitation of adjustments to quantitative forecasts.
- Principle (iii) determines that the adjustment should be made when events are expected in the future.

From these principles, it is implied that judgmental adjustments should only be applied to statistical forecasts when the manager has important contextual information about events not accounted for by the statistical method Fildes and Goodwin (2007).

Judgmental adjustments are usually required if additional information is only known by the expert, and not incorporated into the statistical models. The adjustments are done based on intuition, experience, domain knowledge, and context. Evidence indicates that most companies use a judgmental adjustment to some extent (Perera et al., 2019). Managers are capable of estimating the effects of special events like sales promotions, international conflicts, or challenges that statistical methods might face due to insufficient historical data. Research has shown that human judgment is vulnerable to various biases. For example, individuals frequently identify false patterns in random time series movements, which can result in detrimental modifications to otherwise dependable statistical forecasts, even in the absence of expected special events (Fildes and Goodwin, 2007).

When experiencing noise, forecasters tend to over-adjust, leading to inappropriate judgmental adjustments, (Goodwin, 2000). The type of magnitude of the adjustment is also significant for the performance (Goodwin and Wright, 1993). As demonstrated by Figure 2.2, despite the forecasting adjustment being overall positive in accuracy gains, it can be observed that positive adjustments tend to be more problematic than negative adjustments, this could indicate an optimism bias from the forecaster. The importance of forecasts in supply chain decision-making makes them susceptible to political influence, often leading to pressures to favor politically convenient forecasts over those that are reliable, this helps explain the tendency of optimism (Goodwin and Fildes, 2011) and at the same time contributes to the explanation of why the negative adjustments typically generated more accurate forecasts since something truly relevant must occur in context to justify their presence and overcome the optimism bias, or break through the political barriers involved with not reaching the sales target.

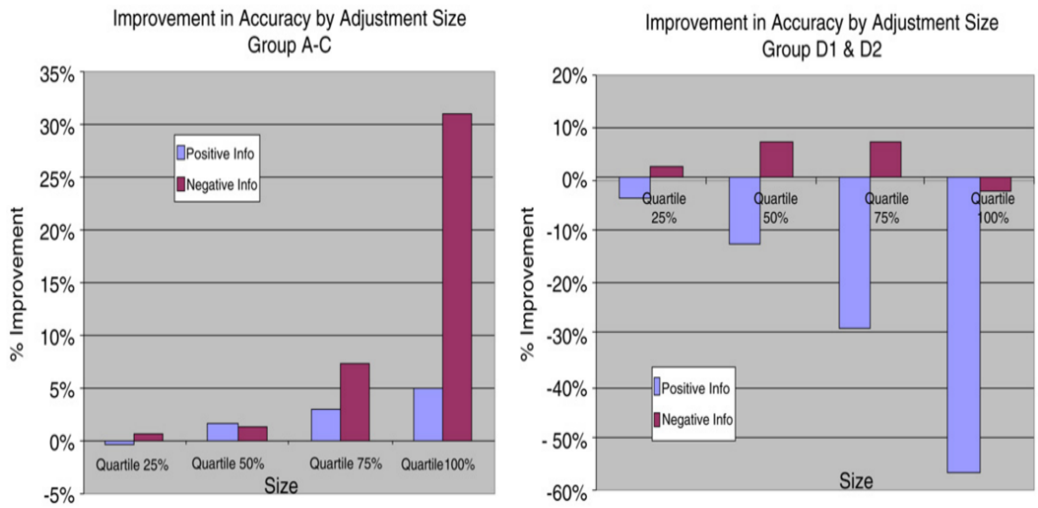

Figure 2.2: Goodwin effects adjustments

Judgmental adjustments to algorithmic computer-based forecasts can enhance accuracy by incorporating important extra information into forecasts, basically integrating context and the statistical prowess of the models (Lin, 2013). Despite judgmental adjustments to statistical product demand forecasts being able lead to better accuracy, however, there is a risk, in many companies, considerable time and effort are squandered on unneeded or detrimental interventions. Such adjustments should be sparingly applied and only when managers have trustworthy information about significant future events that the software is unaware of. Improved software design could result in more effective judgment, but current software frequently serves only to give an illusion of 'scientific' validity to what are essentially judgment-driven forecasts (Goodwin and Fildes, 2011).

Judgment and forecasting are fundamentally intertwined. Even in supply chains employing purely statistical forecasting methods, human judgment remains essential in selecting the appropriate statistical forecasting models, predictor variables, and datasets, in other words, the framing of the collected data belongs to the human agent. This underscores the undeniable role of human judgment in the forecasting process. Therefore, supply chain forecasting, the earliest stage of planning, already experiences human interference, as defended by Lawrence et al. (2006); Moritz et al. (2014); Petropoulos et al. (2018); Perera et al. (2019) and Petropoulos et al. (2022)

Human intervention can be introduced by either a single forecaster or a collaborative group, whether cross-functional or within the same function, to produce a forecast. Several studies praise the Delphi technique as an exceptionally adaptable group forecasting method, often resulting in superior forecasting accuracy's when compared to other group forecasting techniques such as dictator, consensus, and dialectic approaches (Perera et al., 2019). Decomposition is suggested as a strategy to potentially improve judgmental forecasting, by splitting the problem into sub-problems that the expert can handle (Goodwin and Wright, 1993), this aligns with the prevention of the cognitively caused bullwhip effect presented by Brauch et al. (2024). Therefore addressing overly complex problems with a computational system can be beneficial.

The addition of the human component brings forth cognitive biases that must be addressed to conciliate the best aspects of statistical and human components (Sanders and Wood, 2019).

Kahneman (2012) considers human judgment to be fundamentally fallible and defines cognitive bias as the systemic deviation of the rational alternative. Intuition, as a decision-making tool, has evolved to help us rapidly understand our environment, rather than to analyze all data and meticulously weigh alternatives. Managers can be tempted to rely on their first instincts, which can bias their decisions, rather than engaging in thorough deliberation and reviewing all relevant information. To enable effective judgment in forecasting, it is crucial to revisit initial impressions and allow further reasoning and information to potentially override the initial gut reaction (Moritz et al., 2014).

A helpful framework was devised by Kahneman (2012) describing human thinking in two systems:

1. System 1 is an effortless, involuntary, automatic, and quick way of thinking, such as the mental activity involved in passively processing speech or forming impressions from sensory inputs.

2. System 2 is an effortful, voluntary, directed, and slow way of thinking, such as the mental activity involved in conscious agency, making deliberate choices, or performing a calculation.

This proposition for human thinking is called the dual process model (DPM), a special case of multiple process models. Dual process models assume that we need to posit two qualitatively different processes to characterize human thinking. Despite the overall acceptance, there are some critics of DPM, favoring single process modeling (SPM). The dual vs single process model debate has not been resolved, it can be questioned whether the debate can be resolved, and even if it were to be resolved, it will not inform our theory development about the critical processing mechanism underlying human thinking. Trying to answer the core single vs dual process model debate is pointless for empirical scientists (De Neys, 2021).

There are critics of the DPM approach, some defend that systems are not discrete. Evidence indicates that many processes associated with both systems "crosscut". Systems 1 and 2 could be sequentially arranged, however, there isn't sufficient empirical evidence to validate either the default interventionist model or the parallel-competitive model of system interaction. System 2 despite its characterization as rational, can generate errors (Grayot, 2020). Despite the critics, no model proposal was able to resolve the divergences fully, for this work, DPM will be taken as a premise since SPM does not enable the detection of specific biases since it does not differentiate among error types, and the biases that this work is mainly concerned are overconfidence and risk-seeking, that are referred on the forecasting literature with definitions based on DPM. Therefore accepting DPM as a premise is the most coherent approach.

2.2.5 Cognitive biases in forecasting

Cognitive biases in time series forecasting are well documented. One such bias is the tendency to neglect to interpret new data within the broader context of the entire time series, a behavioral pattern known as system neglect. This bias leads forecasters to overreact to short-term market shocks while under-reacting to significant, long-term trends and shifts (Kolassa et al., 2023).

According to the works of Tversky and Kahneman (1974) establishing prospect theory, a cognitive bias is a systematic deviation from the rational alternative. However, due to the multifaceted nature of supply chain management, from the various distinct goals held simultaneously, accuracy, despite being the most utilized metric for forecasting model selection, is insufficient to assess the inventory performance, as Syntetos et al. (2010) pointed out. This is further emphasized in Goltsos et al. (2022). Therefore, a metric must be devised to act as an appropriate assessment of SCM performance, allowing to establish the rational alternative as the best-expected performer, and biases would be translated as deviations from the metric.

The proper use of deliberation in a system 2 oriented manner, despite demanding more effort from the expert, can reap great improvements in accuracy, as demonstrated by Fildes and Goodwin (2007) having individuals justify their judgments in writing brings multiple advantages, such as reducing the rate of unnecessary and detrimental adjustments to statistical forecasts from 85% to 35%. This change led to a marked improvement in the median absolute percentage error, which dropped from an average of 10.0% to 3.6%. There are several potential reasons for this improvement. People might become more accountable for their decisions, prompting them to be more careful. Additionally, the act of articulating explicit reasons may encourage deeper reflection on their reasoning.

The level of trust in the Decision Support Systems(DSS) outputs is of great importance for appropriate decision-making, regardless of what statistical or machine learning models are employed, if the user doesn't take the DSS output into account, the implementation would be useless Petropoulos et al. (2018).

There is evidence within the literature that inappropriate decision-making increases with complexity (Davis, 2018) and considering the complexity of the SCM domain, any human interventions on the S&OP process would be potentially subject to confirmation bias. As Brauch et al. (2024) pointed out, the cognitive strain upon a decision-maker is a relevant contributor to poor decision-making, and decision support systems (DSS) can contribute in this regard as shown by Hertel et al. (2019), however considering the complexity of the SCM domain, judgmental adjustments made upon the forecast itself would require an enormous amount of cognitive resources to account for factors such as operational constraints(production, stocking, delivery costs), the potential impact of the BWE, preferred client management, among others. Therefore, if executed in this stage of the process human interference may be prejudicial to the performance, due to a lack of trust in the DSS, as appointed by Petrou et al. (2012). This dissertation proposes that a system should be implemented to relieve the cognitive strain of the decision-maker, by accounting for the SCM factors, aiding in the perception of the consequences of the adjustment. (they are black boxes)

Human decision-makers often find statistical models challenging to comprehend,(perceiving them as black boxes) leading to lower trust in the method and a higher likelihood of discounting it. However, users of forecasting software were more likely to accept the forecast provided by the software when they could select the model from various alternatives (Lawrence et al., 2002). This further emphasizes the relevance of statistical proficiency described as one of the core characteristics of the Ideal Forecaster as proposed by Kolassa et al. (2023). This also prompts two relevant positions regarding the implementation of an integrated DSS

(i) Treat the problem in the light of a forecasting selection problem, with various options, and relevant criteria being used for selection instead of showing the user an arbitrary computer-generated output as decision suggestion.

(ii) Unless the decision-maker is extremely proficient with methods of explainable AI, machine learning models can be seen as black boxes, therefore classic statistical models should be employed, to avoid mistrust in the DSS.

Following the aforementioned approaches would contribute to creating a safety net against overconfidence bias, by constraining the models to be selected to only typical forecasting models. When observing only one metric (accuracy) it is easy to imagine that the forecaster may be overconfident, since he truly has an empiric understanding of the phenomena that cannot be trained into the statistical models. However, as the complexity of SCM manifest itself in a more complex manner, the cognitive strain to evaluate all the consequences may be so severe that the agent may truly recognize the need for a computer to execute the task.

Therefore, given the complexity of the domain, the chances that the manager would adequately comprehend all the consequences of his intervention is extremely unlikely, if the consequences of the adjustment can be appropriately estimated it may reveal the manager limitations, increasing chances of acceptance.

The concept of risk that has gained widespread acceptance in the risk analysis field is the one articulated by Eugene Rosa who defined risk as “a situation or event where something of human value has been put at stake and where the outcome is uncertain”. According to Rosa, this definition embodies the three critical and sufficient criteria to describe risk. Firstly, risk concerns a potential scenario that may affect someone’s interests, whether in a negative manner. Secondly, risk is characterized by uncertainty about whether this scenario will manifest in the future; therefore, if an outcome is guaranteed, it does not constitute a risk. Lastly, risk pertains to a possible condition or state of reality (Sales et al., 2018).

Therefore, taking Kahneman (2012) definition of bias as a basis, and combining it with the definition of risk supported by Sales et al. (2018), the risk-seeking bias as understood by this paper is the cognitive bias that would be manifest in the propensity of an agent to deviate from the rational alternative by systematically underestimating the impact of potential negative scenarios associated with his decision-making. Note that this description is very similar to the optimism bias, widely discussed within forecasting literature as exemplified by Goodwin and Fildes (2011).

As indicated by 2.2, the literature on judgmental adjustments of statistical forecasts reveals that large negative adjustments, and even negative adjustments in general, typically improve forecast accuracy, whereas positive adjustments generally decrease it. Although these adjustments are common in the industry, companies have yet to adopt authorization protocols that permit negative adjustments without restriction, while mandating explicit justification or approval for positive ones (Syntetos et al., 2016).

This further justifies the particular interest of this work in risk-seeking bias (RSB), however as shown by Syntetos et al. (2010) accuracy loss does not directly correlate with decreases in supply chain performance. The operational implications of RSB, particularly in the context of positive adjustments, are illustrated in the work of Zanddizari et al. (2019), which demonstrates the causal relationship between overstocking and the bullwhip effect (BWE). This suggests that a higher tendency to overstock can be perceived as a precursor to the future emergence of the BWE. Additionally, unlike stock-outs, overstocking can introduce seemingly random fluctuations in the order quantity time series, significantly raising the likelihood of propagating the BWE throughout the supply chain due to increased variability.

Cognitive biases and overconfidence in the judgmental forecasting capabilities can significantly impact not only stock management but also the relations between suppliers (Mannes and Moore, 2013).

The most utilized quantitative methods for mathematical modeling are structural equations and linear regressions (Beal Partyka, 2022) it is understandable since the domain is intrinsically connected to strategic decision-making, that most methods selected are descriptive by nature. However, from the perspective of modeling, the main criticism observed was the lack or little use of more advanced mathematical modeling and robust optimization techniques, (Oliveira et al., 2016) despite the advancements in agent theory to a certain extent (Lee and Kim, 2008).

The bounded rationality of the decision-makers, which leads to sub-optimal decisions derived from cognitive biases. This indicates the current inadequacy of the mathematical modeling such as regressions for decision support, since cognitive overload arises when the agent must resolve overly complex problems, not addressed by the modeling, and likely leading to distorted demand information.(Yamini, 2023; Brauch et al., 2024)

2.2.6 Supply Chain Management and Robust Optimization

SCM is an extremely complex endeavor, total information sharing is unfeasible, and yet decisions are made on the highest level without robust analysis since regressions and structural equations cannot provide it. The literature indicates a lack of modeling maturity. Robust optimization can prove to be particularly usefully in asymmetric scenarios where there is coercion and no information sharing from a dominant business partner (Chen et al., 2023). The proper alignment of strategic views is crucial, however, little to no consideration was found regarding partner-dependent alignment, or how one member could adapt and rethink the strategic goals of the company as different partners alter their strategies (Fayezi et al., 2012).

Decision-dependent uncertainty(DDU) optimization could bring a better potential and characterization of uncertainty as a solution, nevertheless, it would require human intervention to modify the operational constraints to be used in daily operations, so the solutions generated are adherent to reality. DDU already proved useful for strategic supply chain decisions such as factory location and installed capacity (Zhao and You, 2019).

Based on the review conducted by Goltsos et al. (2022) some promising areas for integration of forecasting and inventory control were detected, among them are forecast evaluation and robust optimization. Therefore RO, defined as: Robust optimization deals with uncertain variables by only looking at intervals without a need for further distributional information. This way the technique can produce results irrespective of the true underlying distribution that generates the data. RO offers the advantage of computational tractability compared to alternative methods. Its independence from demand distribution assumptions makes the RO approach well-suited for inventory control applications under demand uncertainty.

Decision-dependent uncertainty (DDU) consists of a class of problems where decisions are optimized over a time horizon under uncertainty, and such decisions influence the time of information discovery for a subset of the uncertain parameters (Goel and Grossmann, 2006). In other words, it is akin to a robust optimization problem approach where a series of sequential decisions affect not only one another but also the uncertainty surrounding the next decision. Notably, despite the dynamic nature of supply chains, Decision-Dependent Uncertainty (DDU) has been applied sparingly in SCM issues.

Instead, DDU methods are more prevalent in energy planning optimization problems such as in Ma et al. (2016). However, the literature on DDU, including works like Yin et al. (2019), provides a solid understanding of the underlying modeling and methodologies. In practice, the nature of uncertainty is decision-dependent, with its resolution influenced by earlier investment decisions aimed at obtaining more accurate information. Investment decisions in wind energy planning are shaped by existing capacity, this is akin to various supply chain factors such as cash flow, stock levels, machinery capacity, workforce, and logistics also being inherently decision-dependent, despite this, DDU has rarely been used in SCM problems.

One example of DDU applied to the SCM domain is in the work of Zhao and You (2019) in the paper the authors make two decisions related to SC design, the decisions were the location of the facilities and optimal capacity, having resilience as the objective function. The conclusion shows that the formulation proved efficient in reducing the BWE. This work utilized a formulation to benefit a single company by turning a regular decision problem into a decision-dependent uncertainty decision problem, thus creating a "local DDU problem" by introducing the concept of fractional delivery, to be better detailed in the framework section.

The local DDU approach was already validated in a paper accepted by SBPO 2024: Supply Chain Management in a Weak Dominance Coopetition Environment under Decision-Dependent Uncertainty. The paper used a slightly less sophisticated local DDU formulation than the one to be presented in the framework section. The formulation was showed to be effective in mitigating the BWE even while having no direct communication between parties, counting only with second-tier supplier communication. For the problem's formulation, agents can communicate information instantly, but the flow of products occurs between standard lead time intervals, therefore creating an artificial stage-like interval between supply chain members. In other words, to meet their respective demand, each agent would need to order at least double their expected consumption, thus compensating for the products still in transit.

Thus, the local DDU-RO was established as an appropriate choice for decision-making simulation within the SCM domain, since it can address the various complex variables that a human agent would likely struggle to conciliate, thus DDU-RO is a good tool for decision support and prevention against cognitive overload and the aforementioned consequences associated to it.

2.2.7 Multi-criteria decision-making and Human agent integration

The goal of mono-criterion modeling (i.e. robust optimization) is to create a maximization problem without constraints, where the optimal solution represents the best possible choice. However, certain aspects of decision-making are difficult to quantify in terms of cost. An alternative approach in single-criterion modeling is to treat non-monetary factors as constraints. While including these criteria in the objective function or constraints may be theoretically acceptable, it can be detrimental to the decision-making process. Doing so limits the flexibility of the decision-maker and leads to rigid choices. Another limitation of optimization modeling is its lack of realism from a human perspective (Aouni and Laflamme, 2014). As already demonstrated in the previous sections, SCM is an endeavor concerned with multiple diverging aspects, and therefore trade-offs are inherent to its nature, this was well captured by the modeling of Neto and Salomon (2022). Therefore a multi-criteria approach makes itself a solid candidate to assist as decision support for SCM decisions.

Applying a multi-criteria approach consists of basically two steps, first, the relevant criteria must be defined and weighted, and the second step is about defining the ranking algorithm, thus the most satisfactory alternative can be selected by simply intuiting the data.

The Best-Worst Method (BWM) is an effective tool for solving multi-criteria decision-making (MCDM) problems, as it provides the weights of the criterion. It begins by selecting the best (most important) and worst (least important) criteria, which are then compared with all other criteria in pairwise comparisons. These comparisons are used in a max-min problem to calculate the criteria weights, ensuring that the maximum difference between the weight ratios and comparisons is minimized. A similar approach is used to determine the weights of the alternatives. These weights are then aggregated to rank the alternatives, and the best option is selected. A consistency ratio is introduced to evaluate the reliability of the comparisons, and a nonlinear min-max model is employed to minimize the maximum difference resulting in potential optimal solutions.

BWM was selected as an alternative to the classic AHP (analytic hierarchy process) weighting method for MCDM problems, AHP is also a pairwise comparison-based method. Statistical results show that BWM performs significantly better than AHP concerning the consistency ratio, and the other evaluation criteria: minimum violation, total deviation, and conformity. Therefore the main advantages of utilizing BWM are its lower requirement of comparison data and more consistent comparisons, leading to more reliable results (Rezaei, 2015). The process of comparison can be better encapsulated by Figure 2.3, as presented by the method's creator.

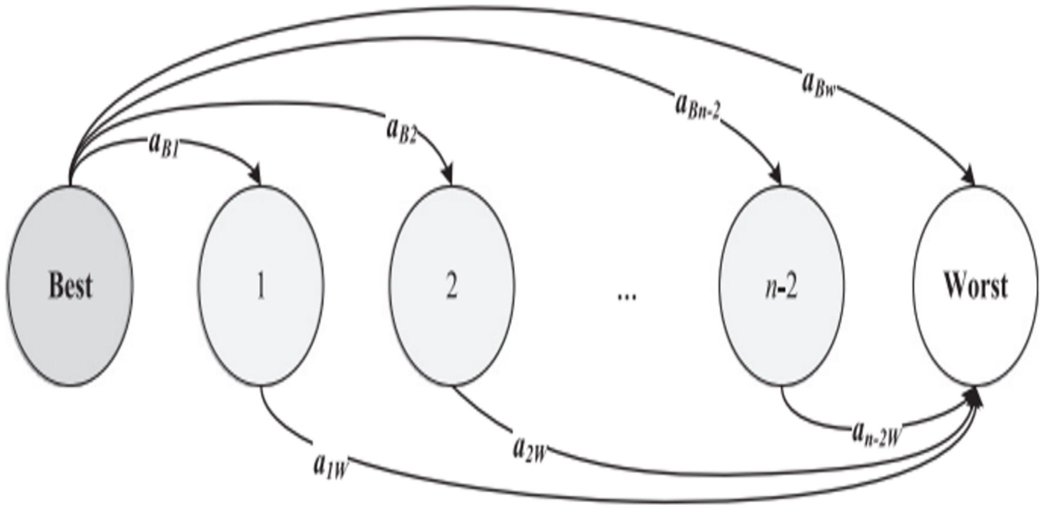

Figure 2.3: Best Worst Method pairwise comparison, adapted from Rezaei (2015)

Relating the ranking aspect of the MCDM solution, two ranking MCDM methods that have widespread use within the literature are TOPSIS and TODIM.

The TODIM method employs a nonlinear version of Cumulative Prospect Theory (CPT), reflecting the same value function shape as the gains and losses function used by Kahneman (2012). In this method, gains and losses are evaluated relative to a specified reference point. TODIM is constructed based on a global value measurement approach as outlined by prospect theory. It combines all gain and loss measures across different criteria into a single multi-attribute value function, which then ranks alternatives from highest to lowest. The Preference Selection Index (PSI) method, by contrast, does not require the determination of criteria weights; instead, it assesses the overall preference value using statistical principles, ranking alternatives according to their preference selection indexes. Thus this method was constructed to have a robust response against human inappropriate perception of gain or loss described within CPT (Gomes et al., 2013).

The TOPSIS method has a simple process that makes it easy to program and use, and it is a prevalent method among MCDM researchers. It is also a very efficient technique with widespread application. From the work of Madanchian and Taherdoost (2023) the main advantages and disadvantages can be summarized as:

- Advantage: TOPSIS can represent the rationale of human choice.
- Advantage: Scaler value of TOPSIS can account simultaneously for the best as well as worst alternatives.
- Advantage: It has a simple computation process that makes programming on a spreadsheet possible.
- Disadvantage: It sometimes cannot precisely determine uncertain data due to the possibility of vague human judgments when the information is insufficient.
- Disadvantage: An important issue to consider is the rank reversal phenomenon, which occurs when the ranking of alternatives shifts after the addition or removal of an alternative or criterion. In contrast, the ranking index method evaluates alternatives based on their distances from the negative and positive ideal points, without taking into account the weights or relative significance of these distances.

Both disadvantages are somewhat addressed by the combination of TOPSIS with the TODIM method, since TODIM can mitigate the rank reversal effect and the robust weighting process of BWM also assists with the vagueness factor from human judgment.

3
Proposed Framework

This research proposes a framework and a subjacent decision support system (DSS) for SCF, following the Humachine paradigm. The proposal integrates data-driven computational methods, such as statistical forecasting models and robust optimization, with cognitively-aware techniques, in the form of a multi-criteria decision-making methodology and an ontology-driven approach for the automatic detection of risk-seeking cognitive biases founded on the Prospect Theory (Kahneman, 2012). This system allows SC managers to relate the impact of their decisions to the company's strategic goals. The DSS enhances managers' cognitive awareness through interactions with the automatic bias detection system, thus, preventing poor decisions influenced by risk-seeking biases.

The proposal addresses the research questions presented in Chapter 1, integrating computational support with expert insight, while avoiding the effects of risk-seeking cognitive biases, in a Data-Driven Cognitively-Aware Supply Chain Management. The cognitive biases must be precisely defined in order to be detected in decisions, this can be achieved by using the ontology for intuitive decision-making developed by Ramos et al. (2024). Computational ontologies are a means to formally model the structure of a system, i.e., the relevant entities and relations that emerge from its observation, and which are useful to our purposes(Guarino et al., 2009).

A deeper understanding of conceptual models can be grasped by analyzing the ontological nature of the relations that appear in the model, clarifying their cardinality constraints and the distinctions within types, considering alternative conceptualizations, recognizing patterns, and visualizing possible interpretations (Guizzardi and Guarino, 2024). This is precisely why an ontology is an appropriate tool to comprehend how cognitive biases work, and from this comprehension, the automatic detection mechanism can be devised. This work used the same mechanism as Ramos et al. (2024) that follows Kahneman (2012) definition.

In summary, the proposed framework selects the most appropriate time series for the expert's intervention, and then selects a model to be used as a rational reference point. This model selection is done by estimating how different forecasting techniques interfere with decision-making, considering the companies' operational constraints and strategic business perspectives, the rational model can be established. Once attained, deviation from the model is classified as a risk-seeking (RS) decision, thus creating an automatic risk-seeking bias detection system, the third sub-process.

Figure 3.1 illustrates the underlying methodology for the proposed framework, represented in the form of a process following the BPMN notation (Kocbek et al., 2015).

The methodology begins with the Dimensionality reduction sub-process, to mitigate potential losses in financial performance when the SKUs are grouped into clusters. Clustering algorithms typically are applied with no operational aspects taken into account, this creates potential vulnerabilities. A way to verify the impact of the dimensionality reduction can contribute to performance, allowing for an accurate assessment of the relevance of each SKU. Once assessed, expert involvement can be employed for the most necessary items.

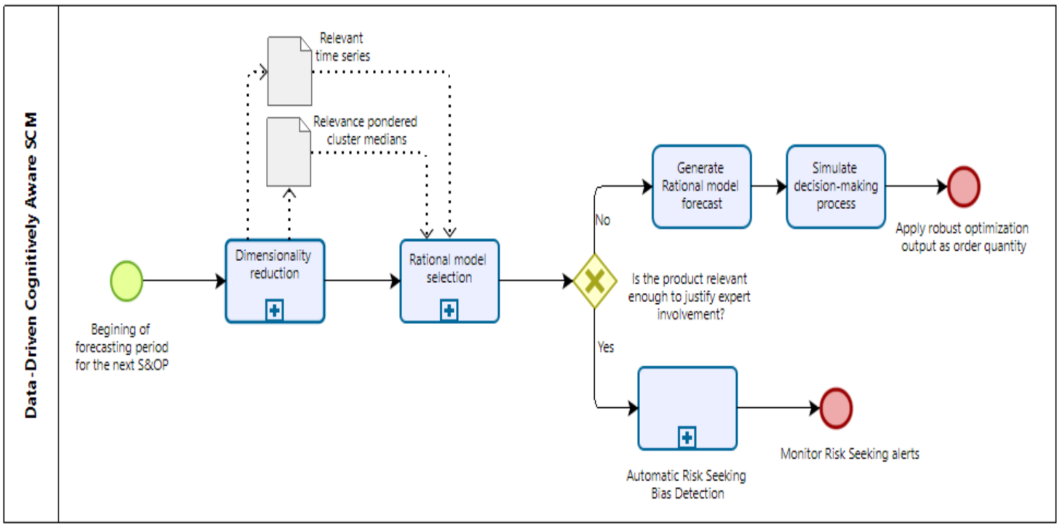

Figure 3.1: The underlying methodology for the proposed Data-Driven & Cognitively-Aware Supply Chain Forecasting Framework

3.1 Dimensionality Reduction

The dimensionality is reduced by generating relevance-weighted cluster medians as approximations for grouped time series. However, clustering SKUs carries risks, as strategically important items near cluster borders may result in poor approximations that harm performance.

Time series poorly approximated by the cluster median—those causing performance loss compared to the naïve model—are identified for expert intervention.

The performance of the clusters is measured using forecast value-added analysis (FVA) (Goodwin et al., 2017), defined as the process of comparing more complex forecasting methods with the accuracy of naïve forecasts. FVA compares the current performance with the naive model performance, to verify if the use of a more complex technique has a justified performance that justifies its use. After the clusters are validated, a forecasting model is selected for each cluster median within the computational time restrictions, maximizing accuracy. For every forecast, a heuristic optimization problem determines the order quantity for all the SKUs contained by the cluster to maximize profit. The remaining time series, inappropriate for the clustering, are ranked by relevance, and the most accurate forecasting model within the computational constraints is used, when the computational time ends, the remaining time series have the naïve model attributed as forecast.

Figure 3.3 illustrates the overall sub-process of dimensionality reduction, where each step is detailed in the following subsections.

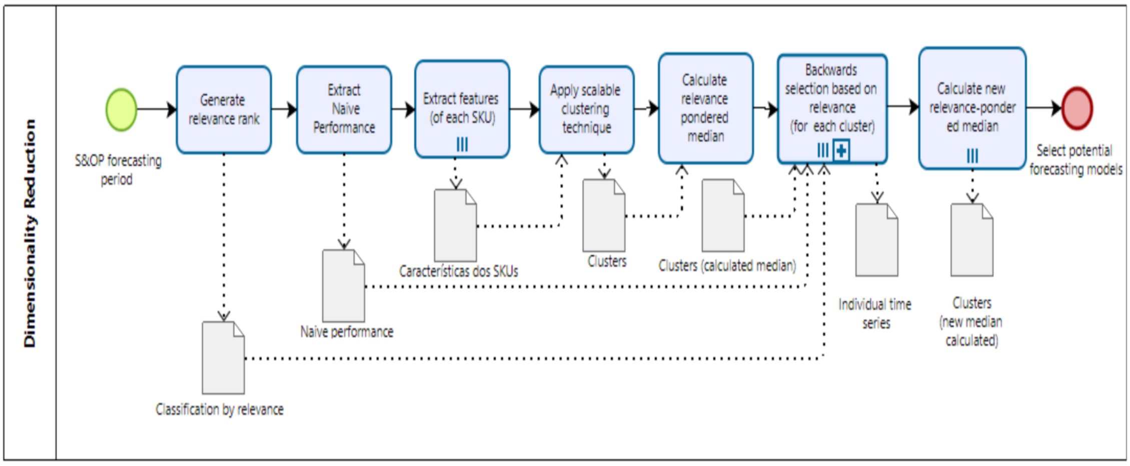

Figure 3.2: Dimensionality Reduction

3.1.1 Generate relevance rank

The relevance rank is a multi-dimensional evaluation and is highly dependent on the strategic goals surrounding the company's inventory management policies. As an example, a set of criteria that tends to be common would be the Impact of lack, defined as the perception of the impact caused by the omission of the product; Lead-time, the longer most relevant the product since the acquisition is harder and demands better planning; Unitary cost of acquisition and absolute profit. Using the set of criteria selected and implementing multi-criteria methods a rank of relative importance can be attained, if new products are included the ranking should be re-run. The high number of SKUs essentially guarantees the convergence of the Vikor method, which would be one of the most recommended methods since it is robust to the inclusion of new products (Opricovic and Tzeng, 2004).

3.1.2 Extract naïve performance

The process is relatively simple, take the period of interest and use the naïve performance measured by the profit generated by each SKU, the sum of the cluster members multiplied by their profit and relevance rank is the metric to be used for the FVA analysis.

3.1.2.1 Extract features

The features of each of the time series are extracted, to make the clustering solution feasible since the proposal of Wang et al. (2006) many programs were developed capable of extracting hundreds of potentially relevant characteristics, due to the extremely high volume of data, a scalable clustering algorithm is paramount. Feature extraction can be costly in computational time due to the many features it uses for each item to be analyzed.

This operation, however, only needs to be done once, in future interactions, the features are extracted only for new items that are added. The feature extraction approach is compatible with parallel computing, assisting with potential computational time constraints.

3.1.3 Apply scalable clustering technique

Due to their high scalability, the algorithms selected for use are K-means and HDBSCAN. Therefore, there are two approaches: Distance-based by K-means, while HDBSCAN is density-based. However, any scalable clustering algorithm could be employed. HDBSCAN possesses a small number of intuitive parameters and few assumptions about data distribution, it is ideally suited to exploratory data analysis and is capable of dealing with data of different scales (McInnes and Healy, 2017). Since the clusters will be used to replace the orders of the time series, algorithms based on shape would not be recommended, such as dynamic time warping, which is commonly applied to time series considering the risk of time series of divergent scale to present similar behavior.

3.1.4 Calculate relevance pondered median

Considering the hypothesis that characteristics-based clustering addresses differences in scale by collecting the features of each data point, it is expected that the clusters to be formed by time series on a similar scale. The scale is relevant for the median of the cluster median pondered on the relevance given to each product. The median is used to generate the forecast and optimization output to replace all the cluster's SKUs.

3.1.5 Backwards selection based on relevance

As described in Figure 3, this step is the main contribution of the proposed methodology, instead of simply accepting the generated cluster, an FVA analysis occurs to guarantee that the cluster is not generating unwanted consequences from the dimensionality reduction.

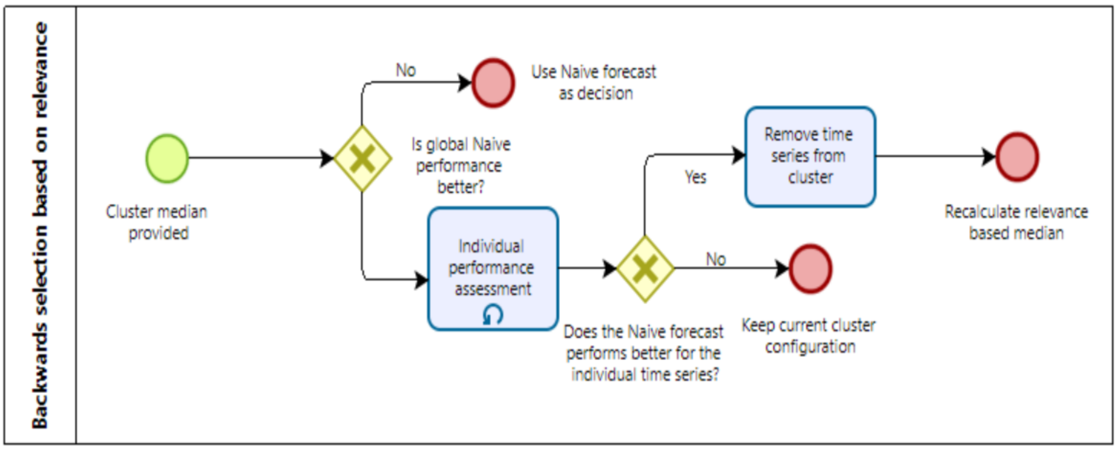

Figure 3.3: Backwards selection based on relevance

3.1.5.1 Individual performance assessment

After the FVA analysis (sum of naïve performance of cluster members against cluster performance) ends, if the cluster is worse than the naïve performance, the cluster is dissolved since computational time is being spent to generate a worse solution. If the cluster has a satisfactory behavior, then each time series is checked individually, and those that perform insufficiently well are removed, the process repeats iteratively until completion.

3.1.5.2
Remove time series from cluster

If the performance of any time series is worse within its cluster than the naïve performance of the individual time series, it is removed from the cluster.

3.1.5.3
Recalculate relevance based median

For each new cluster generated, a new median is calculated, leading to an even better approximation of the remaining time series behavior.

3.1.6
Select potential forecasting models

For each cluster median, a forecasting model is selected, due to the diverse nature of the clusters, various models should ideally be tested, from state space models, classic time series forecasting, and regressions to dynamic regressions and machine learning models. all respecting the computational time constraints, like section 4.1.2.1 shows this step can be computationally intensive but would only run all the alternatives the first time, using the selected model for future iterations.

3.2
Simulate decision-making process

An optimization problem can be solved using integer variables for the order quantity given the forecasting model prediction as a characterization of uncertainty. The objective of the optimizer is to fine-tune the order quantity for the specific cluster. This way the computational time that would be used in the complex individual forecast of all possible time series can be replaced by an extremely fast robust optimization problem to boost performance further. There is an extremely relevant restriction to be added to this step, the cost of acquisition must be considered, otherwise, the optimizer would converge on overstocking from a myopic potential sales objective function. This all implies that the proposal does not guarantee an optimal solution, precisely because if from the hypothesis of the problem, there is no computational time available to appropriately forecast all SKUs, there won't be enough time to solve the optimization problem for the order quantity as well. Therefore, the optimization problem is used as a heuristic approach, that only allows for decisions that would be beneficial.

It is important to note that the proposed framework does not ensure dimensionality reduction but prevents the negative effects of automated or unsupervised dimensionality reduction. This is precisely what happens when classic clustering techniques are applied, based purely on time series behavior and with no concern for the operational performance.

3.2.1 S&OP meeting

During the S&OP meeting many of the isolated time series, i.e. the time series that are not allocated to any cluster, are addressed by the expert supply chain manager, maximizing the utility of the limited time available, and for the remaining time series the naïve model is employed. all-time series within this step, with enough computational time available, proceed to the rational model selection afterwards. This process will simulate the S&OP meeting in the sense that it will produce the same output, the order quantities, with the same inputs of operational constraints and distribution of potential events in the form of a forecast.

The S&OP meeting contemplates the strategic goals of the company, a forecast, and managers' insight to attempt the most satisfactory decision for the business, in that sense the optimizer receives the forecast distribution and the automatic cognitive bias detection system evaluates if the suggestions of the managers should be considered or if they are too risk seeking, in the end resulting in an order quantity. The S&OP meeting simulation is simply the solution derived from the DDU RO.

3.3 Rational Model Selection

This sub-process, described by Figure 3.4, analyzes several model alternatives and selects one of them as rational considering the input of the many different stakeholders, using a Multi Criteria Decision-Making (MCDM) methodology as an approximation of the perceived utility of each output from decisions guided by the specific forecasting model in question. The approximated utility function is attained by employing the Best Worst Method to appropriately weight each criterion. The performance is calculated using DDU RO as a proxy and support for decision-making while considering the outcome that would occur considering past events.

This sub-process comprises four steps; (i) Generate forecasts: selects possible forecast model alternatives, run the prediction for the next iteration for the desired confidence interval, turn each prediction into normal distribution; (ii) Simulate decision-making process: feeds the previously obtained distributions as characterizations of uncertainty for the DDU robust optimization,

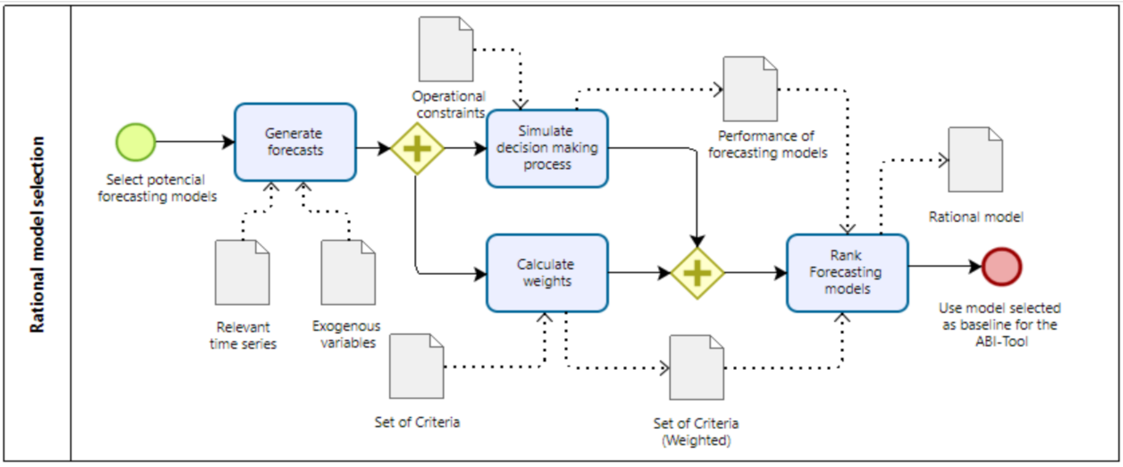

Figure 3.4: Framework for Rational Model Selection

collects the expected values for the relevant criteria; (iii) Calculate weights: discovers the SC manager preferences, attained from the best-worst method and; (iv) Rank forecasting model: selects the most satisfactory model according to the expert by solving the MCDM problem using the joined TOPSIS-TODIM values, therefore we can select the rational/baseline model to be employed.

3.3.1 Generate forecasts

The explainability of the outcome is one of the most relevant factors when taking the models to be evaluated into account. Understanding demand forecasting requires understanding the statistical methods used to produce a forecast, as well as how an organization creates a forecast and uses it for decision-making (Kolassa et al., 2023). In this step we preferred statistical methods over machine learning- based methods for forecasting, since the former typically provides a higher explainability than the latter for the SC manager. The statistical forecasting methods employed were: Historic average, Holt, Sarima, Exponential smoothing (AutoETS), Naïve, Seasonal naïve, Dynamic regression with ARIMA and exponential smoothing based errors, and a combination of the two most accurate models, following the orientations presented by Wang (2023), which states: One prefers to exclude component forecasts that perform poorly and to combine only the top performers. Therefore, only the most accurate models are be combined.

For every forecast made from each model evaluated, a prediction is obtained in the form of a normal distribution of potential events, and therefore a characterization of uncertainty. The characterization of uncertainty is given as input to the DDU RO process to function as a suggested decision of order quantity and provide the needed information for the criteria to feed the MCDM ranking.

3.3.2 Simulate decision-making process

The basis for the application derived from the article Resilient Supply Chain Design and Operations with Decision-Dependent Uncertainty Using a Data-Driven Robust Optimization Approach (Zhao and You, 2019). In the paper, the authors make two decisions related to SC design, and the location of the facilities, and then they optimize for capacity and resilience. The conclusion shows that the formulation proved efficient in reducing the BWE. A new formulation to benefit a single company by turning a regular decision problem into a decision-dependent uncertainty decision problem by introducing the concept of fractional delivery. This approach was done by using two expeditions(deliveries) (zI, zF) and a constraint of delivery interval between expeditions for zF as proportional to how much of the demand was answered in zI. Uncertainty is defined as ξ for the formulation. One of the benefits of this approach is that zI can be treated as a preferential client. The objective of the model is optimization for profit while following the mentioned constraints, the model also contemplated lead time uncertainty, conversion time uncertainty, as well as different stocking costs for raw materials and finished products. This way decisions are made regarding: How much should be ordered(X), How much raw material should be converted in each period (ConvI, ConvF), if any product should be converted and stored for the next cycle Finished Product Stock(FPS), and we may evaluate consequences such as the Remainder of finished products in the previous decision cycle, costs associated with storage and vehicle fleet expenses. The formulation of the constraints is summarized as follows:

- zI must be dispatched within 2 days
- zF to be dispatched depends on the proportion of demand met in zI (ξI - zI = ξF)
- zI + zF = FPS + Raw Material
- Stock(RMS)|Conversion Time (ConvT)
- X = RMS for tI; FPS = Remainder + RMS|ConvTime
- Converted Prime Matter for zF (ConvF) in t-1 act as FPS for zIt, make to stock.
- The profit for both zI and zF has a minimum value, acting as a cash flow constraint.

For each provided forecast/characterization of uncertainty, the performance of the 5 criteria is collected, using the DDU RO solution.

3.3.3
Calculate weights

Based on the previous statement of Kolassa (2023) on how the internal departments may influence decision-making for the SCF model selection: Accuracy, cost, profit, BWE, OTD, and service level. The weighting process of these criteria was done using the BWM method. All this set of criteria, except for accuracy, are also supported by Gostos (2022) The selection of the BWM as a means to collect the SCM experts weighting for the criteria was derived from the conclusion of the propositional paper by Jafar Razei (2015), from which follows: BWM is a vector-based method that requires fewer comparisons compared to matrix-based MCDM and the final weights derived from BWM are highly reliable as it provides more consistent comparisons compared to AHP.

The BWM works by collecting the expert input, relating the most relevant criteria and least relevant criteria in pairwise comparison, and solving an optimization problem to determine the remaining comparisons and weights (Rezaei, 2015).

3.3.4
Rank forecasting models

After the preferences have been appropriately weighed by the BWM, now based on the performance of each criterion obtained by the DDU RO solution, we apply classical MCDM methods to determine the most satisfactory and therefore, the rational model. TOPSIS and TODIM were the methods used, to increase robustness, the methods were combined. TODIM applies min-max normalization, that is, the best alternative is assigned a rating of one and the least preferable, zero, while TOPSIS evaluates in regards to the best possible criteria for each model. Min-max normalization was applied to the TOPSIS output, therefore bringing both methods to the same scale, this way they could be combined in a sum and subsequent Min-max normalization to adjust the scale allowing the rank of the most satisfactory model to be used, that is, the rational model.

Ideally, a company's strategic outlook should be long-term and remain stable with minimal changes. However, the volatility of certain situations, as highlighted by Covid-19, has proven otherwise. The presented process offers a valuable approach for selecting the most suitable model during regime shifts. Adjustments can be made rapidly by modifying the BWM inputs and recalculating rankings in just a few seconds since the forecast generation and decision-making simulation processes do not need to be re-run.

3.4 Automatic risk seeking bias detection

Once the rational model is established, there is a reference to classify the decisions according to the risk-seeking bias, decisions that deviate from the DDU solution of the rational model and have a lower expected value trigger a risk-seeking alert and the consistent deviation characterize risk-seeking bias propensity on the part of the decision-maker. This is shown in Figure 3.5.

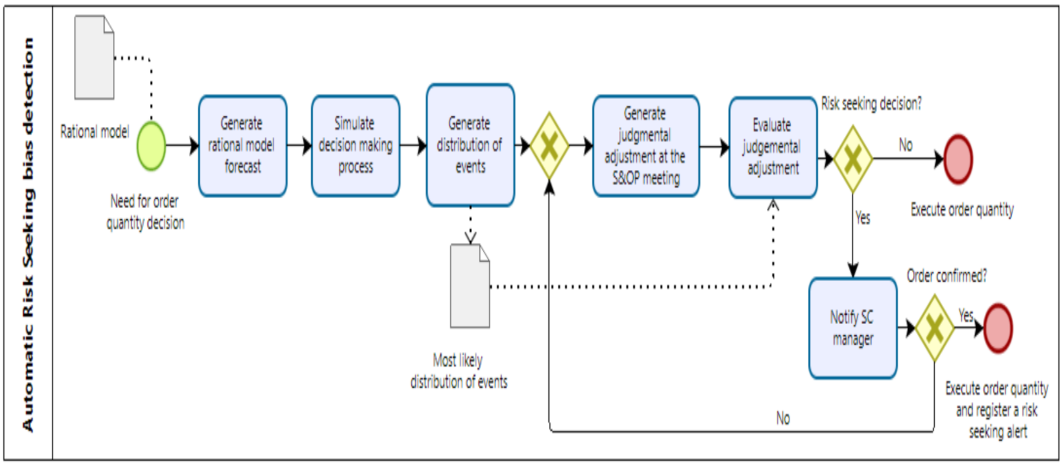

Figure 3.5: Framework for Automatic risk-seeking bias

3.4.1 Generate rational model forecast

Once the rational model is identified, it is the only one that needs to be generated. However, periodic re-evaluation of the models is recommended to ensure continued alignment with the company's objectives over time.

3.4.2 Simulate decision-making process

The computational power needed for the final stage of the solution is significantly reduced, as the DDU RO problem only needs to be solved for the rational model.

3.4.3 Generate distribution of events

The key element in evaluating judgmental adjustments is analyzing the event distribution to determine if the adjustment is the result of risk-seeking (RS) behavior. For this, the most accurate forecasting model serves as a representation of uncertainty in the DDU RO evaluation phase. It is crucial to recognize that this distribution may or may not match the rational model. If the rational model is identified as the most accurate, further problem-solving becomes unnecessary, as any judgmental adjustment would invariably be classified as RS, given its deviation from the optimal solution. Otherwise, the most accurate model is used to estimate the distribution of events, assessing the impact of the deviation.

3.4.4 Generate judgmental adjustment

Judgmental adjustments to algorithmic computer-based forecasts can enhance accuracy by incorporating important extra information into forecasts, basically integrating context and the statistical prowess of the models (Lin, 2013). The final criterion of the decision is made by the forecaster, therefore the alterations are made before the S&OP process. In the forecasting literature definition of judgmental adjustment, the main concern is to generate a more accurate forecast. However, this process potentially ignores how it may impact other departments' decision-making processes and the objectives of the business as a whole. To promote stronger cross-departmental integration, the forecasting models were assessed through a multi-criteria framework, thus mitigating the potential occurrence of a less satisfactory decision considering company goals. During the S&OP meeting, while observing the expected performance of the rational model, the relevant departments must discuss and conclude a new, potentially better order quantity.

Cognitive biases, especially risk-seeking behavior, can result in suboptimal decisions. For this reason, the ARSBDS is most effectively applied during the S&OP process to mitigate poor decision-making. Judgmental adjustments in this context should be seen as tactical decisions, contributing to a decision support system, not merely as a forecasting tool. ARSBDS can detect risk-seeking alternatives by analyzing differences in expected profit, allowing it to manage the risk-seeking nature of overstocking while keeping service levels acceptable.

3.4.5 Evaluate judgmental adjustment

Considering the most accurate and the rational models are not the same, a simplified version of the DDU problem is solved using the most accurate model's characterization of uncertainty for both decisions. The new agility of the process comes from the fact that since the order quantity is already predetermined, the only decisions to be made are the distribution of deliveries to preferential and regular clients. The solution is then obtained for the rational and judgmental adjusted models if the expected profit of the judgmentally adjusted model is lower than the rational model solution. The system notifies the supply chain manager of the potential risk-seeking decision. A risk-seeking alert is registered if the supply chain manager insists on executing the decision. Depending on the frequency of alerts registered, a systemic pattern of deviation from the rational model would establish the supply chain manager as an agent with a high propensity to RS bias, following the definition of Kahneman (2012).

3.5 Proposed Architecture

An architecture was devised to demonstrate how the Dimensionality reduction and Rational model selection sub-processes are integrated, as shown by Figure 3.6.

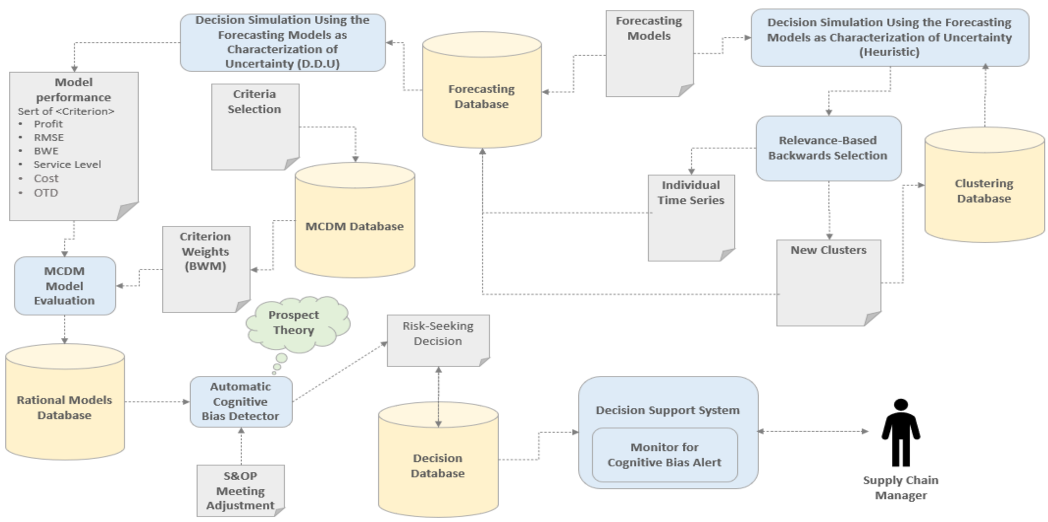

Figure 3.6: Architecture for Data-Driven & Cognitively-Aware Supply Chain Management Framework

The proposed architecture aims to solve the main concerns about adding human integration in the S&OP decision-making process: firstly, the limited scalability. By clustering as many time series together without loss in operational performance, the decision-maker can more effectively focus on the most relevant individual time series. The second element is the insertion of context while controlling for risk-seeking bias, a Rational model is selected following MCDM techniques to reflect the strategic views of the company for every individual time series. Based on prospect theory, the risk-seeking bias can be detected given the adjustments made from the rational model, thus solving the problem of risk-seeking cognitive bias in decision-making distortion.

4
Results

This Chapter details the implementation of the proposed framework and the results of its evaluation in two distinguished scenarios. First, the method proposed for dimensionality reduction was evaluated using M5, a well-known forecasting competition dataset from the literature (Makridakis et al., 2022), regarding its scalability. Second, the automatic detection of risk-seeking cognitive bias was evaluated using real data about cardboard production in Brazil, from 2017 to 2023 (IPEA, 2024), with regard to its potential to increase the robustness of the system during a pandemic period, when the demand for this product suffered a huge variation. Both evaluations are detailed in the following Sections, together with their corresponding results.

4.1
Evaluating Scalability

To validate Dimensionality reduction, the M5 forecasting competition dataset was selected (Makridakis et al., 2022). The dataset contains 30.500 SKUs of real data divided among three states each with various Walmart stores and products divided into various categories. There is complementary information such as the day of the week and the occurrence of special events and prices however, the most important information is the metric of relevance given for each product, the metric would be used to evaluate accuracy considering the importance of each item. This value can be used as a proxy of the relevance rank. The dataset contains a hierarchical time series, that is the sum of all the SKUs of a given category, being the sum among all regions or product types, for the experiment conducted these special time series were treated as if they were any other.

We performed experiments varying the clustering algorithm, including K-means and HDBSCAN, this way clustering based on distance and density. The only requirement for the algorithms is high scalability, any other scalable clustering algorithm could have been employed.

The K-means scenarios varied k (the number of clusters) in the range [5, 10, 20, 25, 50, 75, 100, 250, 500, 1000]. Most scenarios resulted in a clustering with a low silhouette coefficient (ranging from 0 and 0.04, mostly close to

0), with the best scenario being the one with k=75 clusters. However, when further evaluating each clustering scenario, they performed worse than the naive model.

The HDBSCAN algorithm was assessed by varying the min_sample parameter in the range [2, 5, 10, 20, 25, 50, 75, 100]. Most of the resulting clusters had a silhouette score between - 0.09 and 0.04.

Therefore, further testing progressed with the two best results as described in Table 4.1:

Algorithm	Min sample	Noise points	k	SC
HDBSCAN	2	2,350	5	0.28
HDBSCAN	5	894	28	0.11

Table 4.1: The specification of the parameters (min_sample, noise_points, number of clusters, or k) and the internal evaluation metric value (silhouette coefficient, or SC) for the top-2 scenarios using HDBSCAN clustering algorithm

The interaction that resulted in 28 clusters had great variance among each group, establishing clusters with SKUs of widely different scales. The difference among the maximum and minimum difference of the SKU clusters was in a range from 20 to 700 units of difference, all clusters analyzed resulted in a loss of performance when compared to the Naive model.

Noise points	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
2,350	20	6	6	28,101	7

Table 4.2: Number of SKUs per cluster, resulting from the application of HDBSCAN Clustering

As it can be perceived from Table 2 there is a high concentration of SKUs in cluster 3, if all the clusters were to be accepted the reduction of dimensionality would be 93%. The silhouette score obtained was 0.28, which indicates an inefficient set of clusters, however, this section aims to observe the financial impacts of applying dimensionality reduction techniques, to only reduce dimensionality when beneficial.

4.1.1 Forecast Value-Add Analysis

As Table 4.3 demonstrates, cluster 3 fails to be cohesive enough to be approved by the FVA. Then, we further evaluated the performance of two additional scenarios (including cluster 3 or excluding it) and compared the cluster profitability of each scenario against Naive performance. Since the information related to acquisition and stocking costs was unavailable in the dataset for this comparison, an arbitrary value was defined as one-third of the sale price for the total cost of overstocking a product, considering both the stocking and acquisition penalties.

FVA	Naive	Clustering solution
Overall performance	103,786,193	75,232,299
Clusters (0, 1, 2, 4)	71,302	392,875
Cluster (0)	36,565	201,474
Cluster (1)	10,863	60,440
Cluster (2)	10,327	58,853
Cluster (4)	13,547	72,108
Cluster (3)	103,714,891	74,839,424

Table 4.3: FVA analysis

Unfortunately, the decision of not adopting cluster 3 limits the dimensionality reduction that could be achieved using our proposed method, since over 99% of the dataset was not addressed by the clusters. However, there is still a slight reduction for 7% of the SKUs of the dataset. Supposing that the forecast of each SKU takes 30 seconds (and considering that the 7% of the SKUs in the adopted clusters account for 2,135 items), the time consumed by forecasting all the items in the M5 dataset would be reduced from 4.24 to 3,94 hours. Thus, our proposed method for the dimensionality reduction proved beneficial since for a low computational power it generated a profit of over 320 thousand dollars over the Naive performance, a value equivalent to over 5.5 times the previous return. This gains represent an overall profit increase of approximately 0.31%., which for high-scale scenarios (such as the full Walmart database, which contains over 350 million SKUs) can contribute up to a significant amount of dollars. Regardless of the failure in a reduction for a majority of the SKUs, this experiment showed that clustering can lead to losses in operational performance, alerting for the lack of integration between areas.

In a nutshell, this experiment presented a mechanism to check for the operational consequences of clustering, which if not properly monitored can lead to negative effects. Thus, we argue that our proposed dimensionality reduction method serves as a checking point against poor clustering, supporting the decision on which series should be forecast with more complex models.

4.2 Evaluating Robustness

Both Rational model selection and Automatic risk-seeking bias detection were evaluated on top of the same time series dataset, comprising data about

Brazilian cardboard production. We extract the period from January 2017 to March 2023 for our analysis, therefore displaying the robustness of the system to the regime change created by the COVID-19 crisis.

To evaluate the robustness provided by our proposed method for Automatic risk-seeking bias detection, we used real data about Brazilian cardboard production, from January 2017 to March 2023, being the chosen test period, provided in IPEA (2024). This period comprised the COVID-19 pandemic, in which the demand for this item suffered huge variations, leading to the bullwhip effect and its resulting impacts. The time series had monthly measurements on a scale of tons of cubic meters, the training data started at the year 2000 and was used to make monthly 1 step ahead predictions for the test period. The time series presented well-delimited periodicity every 12 months, the same can be said about the exogenous variables used in the dynamic regression forecasting. Figures 4.1 displays the time series and its analysis, the time series has a very well defined 12 month seasonality, regarding it's trend there is a slope around 2008, most likely caused by the financial crises at the time and after it the trend grows rapidly, especially after the 2020 pandemic.

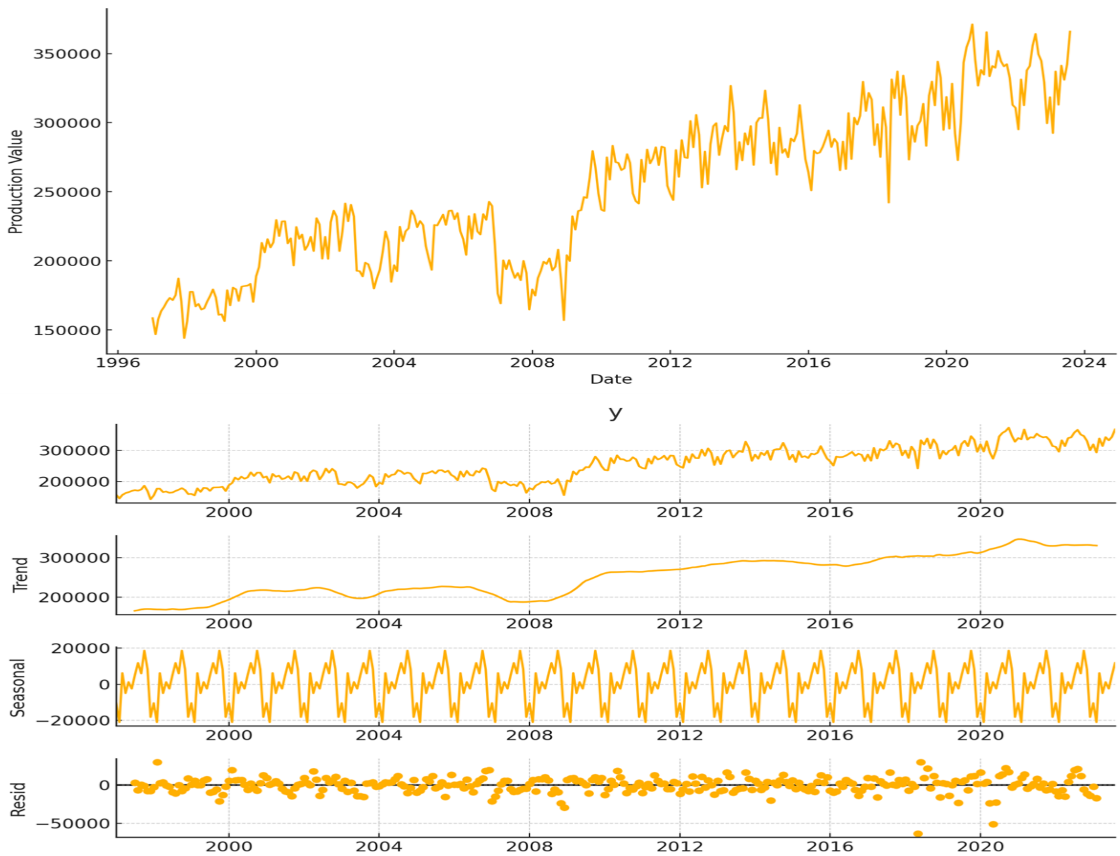

Figure 4.1: Brazilian cardboard production

The performance of all models was evaluated for all the criteria mentioned earlier. The models selected for evaluation were: Naïve as a benchmark, Holt, ETS/AutoETS, Sarima/AutoARIMA, Dynamic Regression using ARIMA structure, and Dynamic Regression using ETS structure. The set of exogenous variables selected was not available in real-time, making them stochastic regressors, for this specific instance, the information lag was two months, and therefore a two-step prediction needed to be employed to estimate the values, this is not ideal since it introduces uncertainty into the model. The list of variables, alongside the best predictors for each during the 77-point period, can be shown in figures 4.2, 4.3, and 4.4.

- Brazilian industrial production index -ETS - MAPE 6.5%
- Brazilian non-durable imported goods - Sarima - MAPE 4.3%
- Brazilian semi-durable imported goods- Holt - MAPE 13.2%

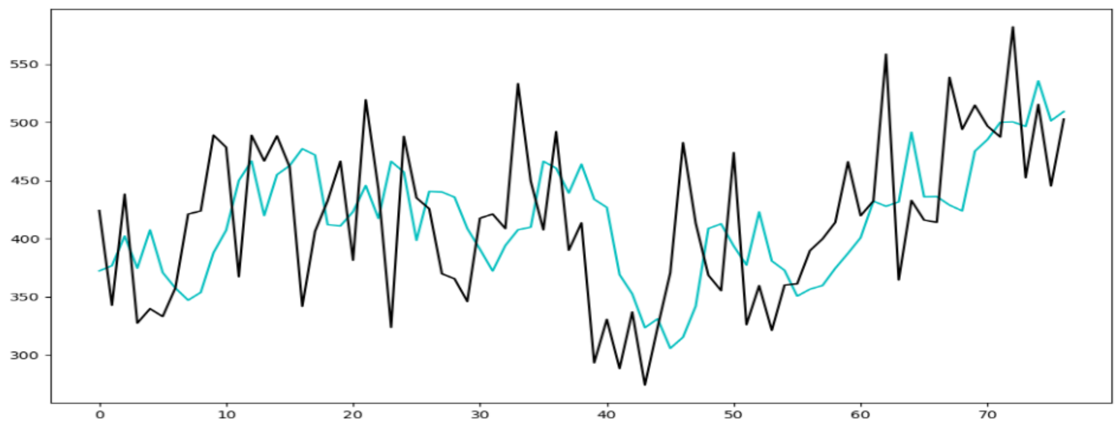

Figure 4.2: Brazilian semi-durable imported goods

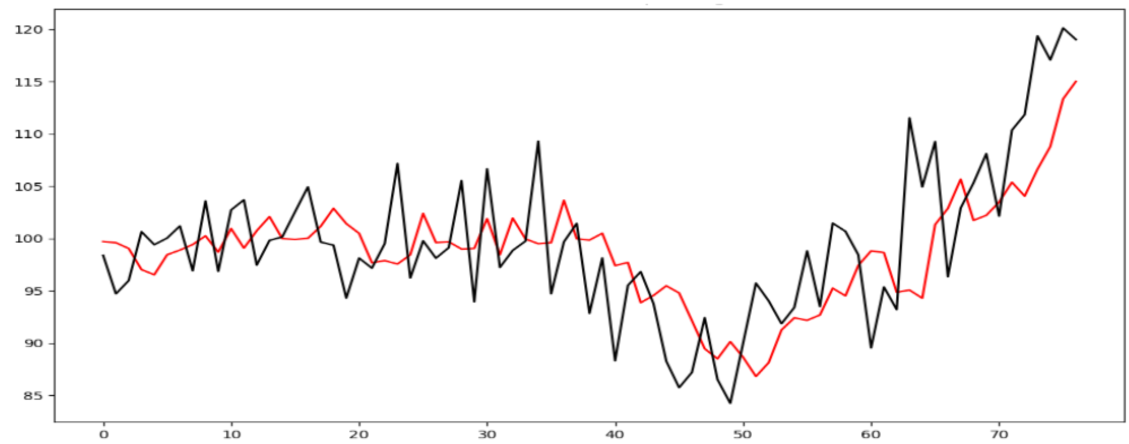

Figure 4.3: Brazilian non-durable imported goods

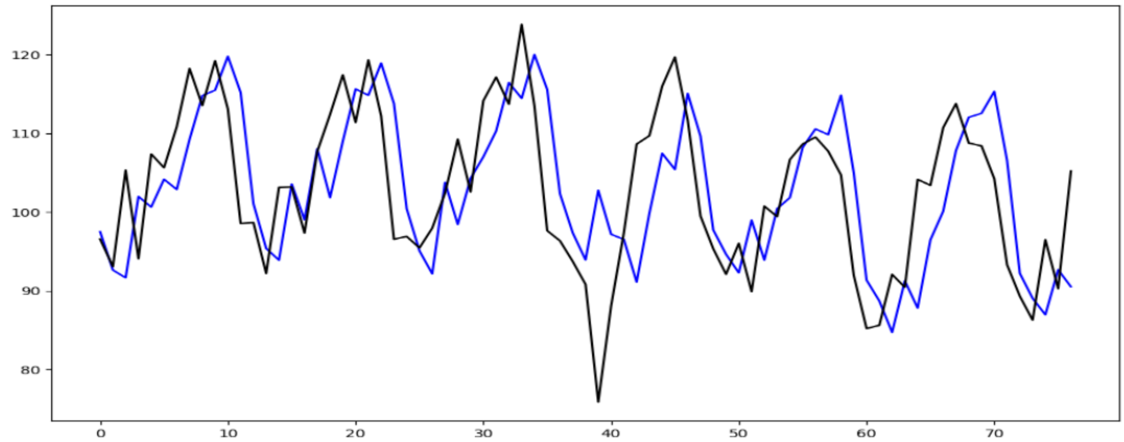

Figure 4.4: Brazilian Industrial Production

The reasoning for selecting this set of exogenous variables was the change in behavior of Brazilians' online shopping habits during and after the pandemic, the assumption was that the increase of non-durable and semi-durable imported goods, while monitoring the industrial production index would provide a better reaction for the regime changes during the test period as well as the historic correlation between all time series.

After collecting the forecasts and their distributions of uncertainty, the DDU RO problem can be solved sequentially for each instance and forecasting model, the BWM weight calculation process was done by taking the mean of the experts' responses from a convenience sample. Therefore, the performance of the models can be presented in Table 4.4:

Some interesting considerations can be attained from Table 4.4, the overall satisfactory OTD performance can be explained by the cash flow constraints, motivating a higher service level in the first iteration (zI) regardless

Table 4.4: Forecasting models performance

Weights	**5.64%**	**18.60%**	**12.40%**	**12.40%**	**18.27%**	**32.69%**	**100%**
Criteria	**RMSE** (-)	**BWE** (-)	**Cost** (-)	**Profit** (+)	**Service Level** (+)	**OTD** (+)	**Evaluation** (+)
AutoARIMA	18,953	1,882	28,477	6,839	96.56%	96.55%	0.56
AutoETS	17,569	618	28,807	6,617	97.78%	97.78%	0.81
Holt	11,149	1,601	29,932	7,293	95.86%	98.72%	0.83
Combined	12,165	575	29,199	6,733	94.29%	98.58%	1.00
Dyn Reg ARIMA	15,034	715	29,142	6,675	98.22%	98.22%	0.88
Dyn Reg ETS	22,613	3,269	27,828	6,923	95.86%	95.86%	0.09
Naive	25,491	3,881	27,880	7,236	96.61%	96.61%	0.00

of the model selected. The Naive model had one of the best performances considering the financial criteria, which was unexpected. The level of accuracy was not directly correlated with the mitigation of the BWE and neither was the complexity of the model utilized since the worst BWE performers were the least and one of the most complex models. Considering the variance of the original time series, 681, it was possible to analyze if the usage of each model as a characterization of uncertainty resulted in the distortion of information, by measuring the variance of the order quantity time series, which is the same as the DDU RO solution. Any models that present a higher BWE than the original demand time series would be theoretically inappropriate as a model for SCF, since it is distorting information for the following members. This is a very relevant finding, since considering the outcomes presented, following classical forecasting evaluation techniques focused on accuracy, would lead to the selection of the Holt model, which is an inappropriate model having over three times as much variance as the original time series.

On the other hand, our framework led to the selection of a proper model for this instance. However, it should be highlighted that the outcome is sensitive to the BWM preferences. Note that the rational model was the Combined model, which was derived from Wang et al. (2023) doing the average of the two most accurate forecasting models for this instance: Holt and the Dynamic regression with ARIMA structure. The most accurate model was the Holt model, and therefore it was used to generate the distribution of events for the automatic risk-seeking bias detection.

4.3 Automatic risk seeking bias detection

To demonstrate the automatic risk-seeking bias detection (ARSBD) process, arbitrary adjustments were made generating increases of 5%, 10%, and 20% upon the optimal decision provided by the rational model solution. To evaluate the performance of the system, two sets of DDU RO problems were compared: the judgmental adjustment (JA) and our proposal (ARSBD). JA represents the solution for a case without the ARSBD intervention as a mediator, therefore the order quantity applied was always the judgmental adjustment suggested (i.e. the RS arbitrary adjustments mentioned previously). ARSBD, on the other hand, needs to solve both the rational and the judgmental adjustment problems using the (most accurate) Holt model prediction as the distribution of potential events. For each iteration, the system chooses which decision should be implemented based on the higher potential profit, this way generating a solution in-between the rational and the JA.

Since what is being evaluated are the effects of the decisions taken and not the accuracy of the model of origin, especially considering the difficulty of defining an approximation for accuracy of the judgmental adjustment heuristic in real life, the RMSE criterion was excluded and the weights of the remaining criteria were distributed proportionally. The performance of the models in this new weighting scenario can be seen in Table 4.5.

Table 4.5: Judgmental adjustment and ARSBD performance

Weights	**22.26%**	**14.03%**	**14.03%**	**21.81%**	**21.81%**	**100%**
Criteria	**BWE (-)**	**Cost (-)**	**Profit (+)**	**Service Level (+)**	**OTD (+)**	**Evaluation (+)**
Rational	575	29,199	6,733	94.29%	98.58%	0.932
JA (5%)	789	30,689	4,435	97.73%	97.72%	0.632
ARSBD (5%)	660	29,500	7,021	98.72%	98.72%	0.967
JA (10%)	844	31,950	2,280	98.72%	98.72%	0.435
ARSBD (10%)	580	29,414	7,011	95.90%	98.72%	1.000
JA (20%)	913	34,283	-3,536	98.73%	98.72%	0.048
ARSBD (20%)	588	29,500	7,023	95.86%	98.72%	0.993
Holt	1,601	29,932	7,293	95.86%	98.72%	0.363

Table 4.5 shows that, despite being optimized for profit, by using the same distribution as a basis for the distribution of events and obtaining the highest performance in most criteria, the Holt model is less satisfactory than the Rational and ARSBD-mediated adjustment. This confirms the complex

multi-criteria nature of this problem and further emphasizes the need for a more holistic approach, that is, operational decision-making should consider more aspects than the myopic view based on profit, typically experienced in businesses. The most prevalent trade-off is between the BWE, profit, and cost.

The better performance of the intervention over the judgmental adjustment was already expected. However, the response output by the ARSBD system is also more satisfactory than the rational model. Therefore, the methodology proposed is capable not only of mitigating the adverse effects of the adjustment but also providing a better solution than the purely statistical rational model.

Another important aspect is the indirect role of the system in the mitigation of the BWE. Since the BWE of JAs are higher than the variation of the original series (681), while the ARSBDs' are within an acceptable range, preventing information distortion from this link to the other members of the supply chain.

It is important to consider that the results detailed in Table 4.5 demonstrate the system effectiveness against the optimistic overconfidence bias, as pointed by Van Oudenhoven et al. (2023), since the judgmental adjustments with a lower expectation of profit would be filtered out.

Additional benefits were observed along the investigation of the research questions, which present significant complementary findings. In a variation of the scenario analyzed in the previous section, having small differences such as slightly different sell price, cost, and cash flow constraints, however, still relating to the same base time series and models. The objective of the analysis was to observe the effects of preferential client acquisition, exemplified by a 15% increase in real demand, which enables a variance reduction since the increase introduces a partially deterministic component for the demand uncertainty. Another point of interest was the theoretical effect of introducing a 15% recovery reverse supply chain from what was sold on the previous instance. This introduces an entirely deterministic component that reduces the need for the manufacturing of the current instance, and makes the uncertainty surrounding the two-stage manufacturing decision (how much should be delivered to preferred client zI and regular clients zF) less volatile. Therefore, 4 scenarios were devised for comparison:

1. Standard - Base scenario
2. RSC - Reverse Supply Chain Scenario, 15% of the demand met in $t-1$ is delivered as finished products to be used at the beginning of instant t.

3. PC - Preferred Client scenario, there is a 15% increase in the real demand, reflected by a deterministic increase of 15% in the forecast that describes the uncertainty at time t, thus the relative variance becomes far less significant facilitating decision-making.

4. PC & RSC - PC and RSC scenarios happen simultaneously.

We compared the effects of each scenario on the BWE, particularly if the forecasting models could generate an appropriate variation upon the order quantity time series. The results are in Figures 4.5 and 4.6, that show the variance of the order quantity time series for each of the models alongside the variance of the demand.

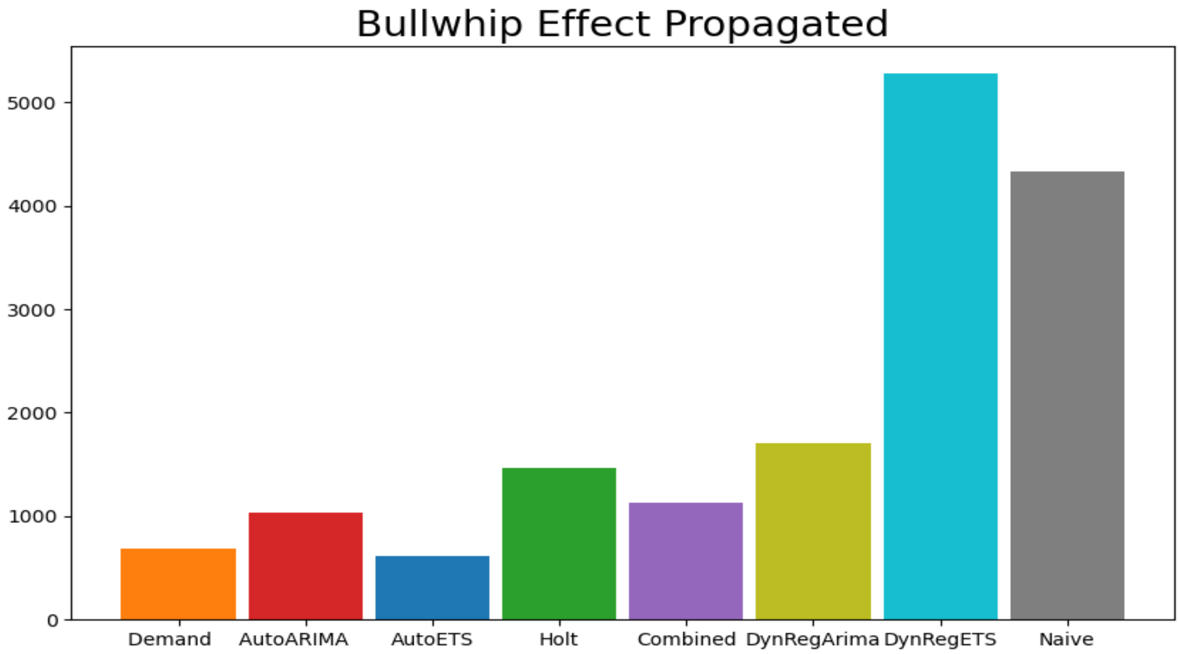

Figure 4.5: Standard scenario BWE propagation

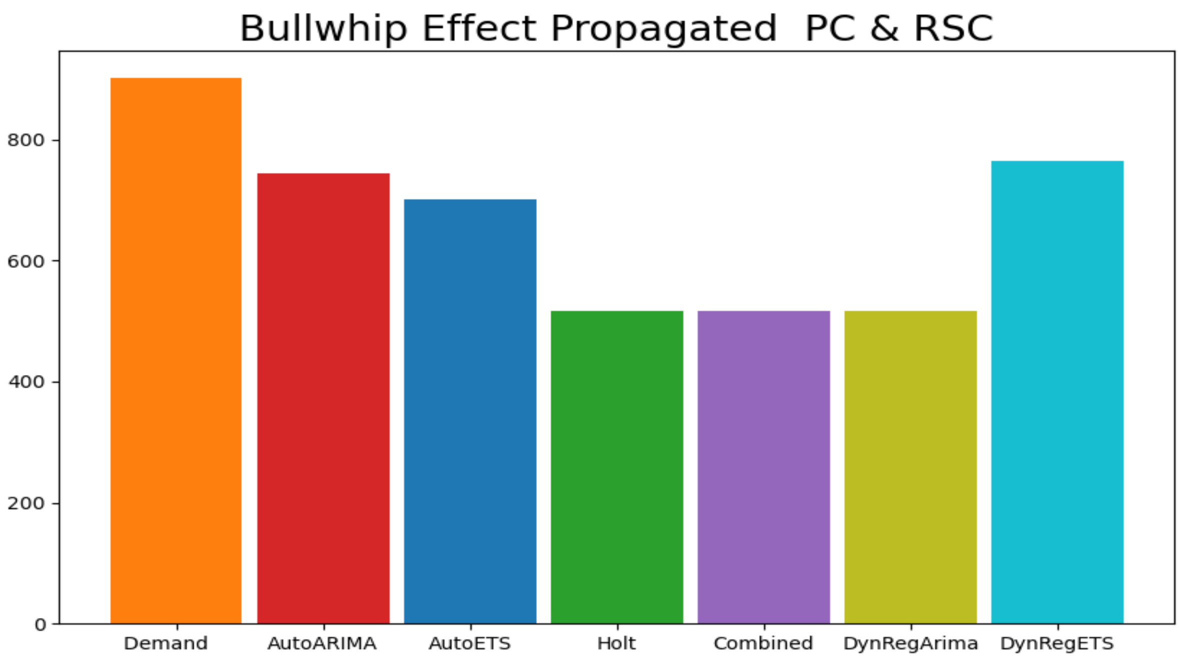

Figure 4.6: PC & RSC scenario BWE propagation

Notice that despite generating an increase in the demand variation, which is expected, since the acquisition of the new PC caused a demand upward shift, all forecasting models had less variance than the original series. Therefore all models were considered appropriate, following Monroe (2012) definition of the bullwhip effect.

The effectiveness of the proposed formulation for BWE dispersion can be demonstrated by the comparative performance of order quantity outputs for the Standard and RSC + PC scenarios using the ETS-error-based dynamic regression:

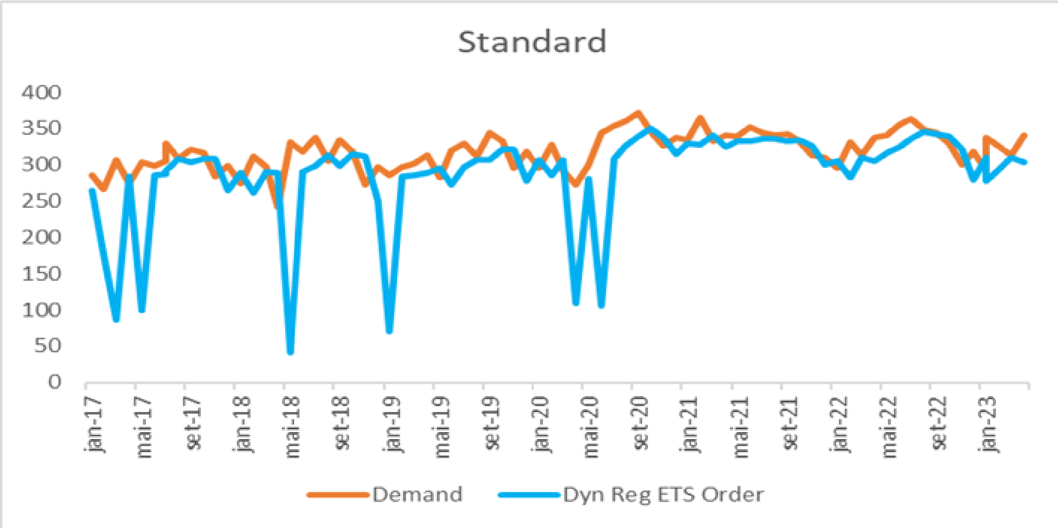

Figure 4.7: Standard scenario demand and Dynamic regression ETS order quantity

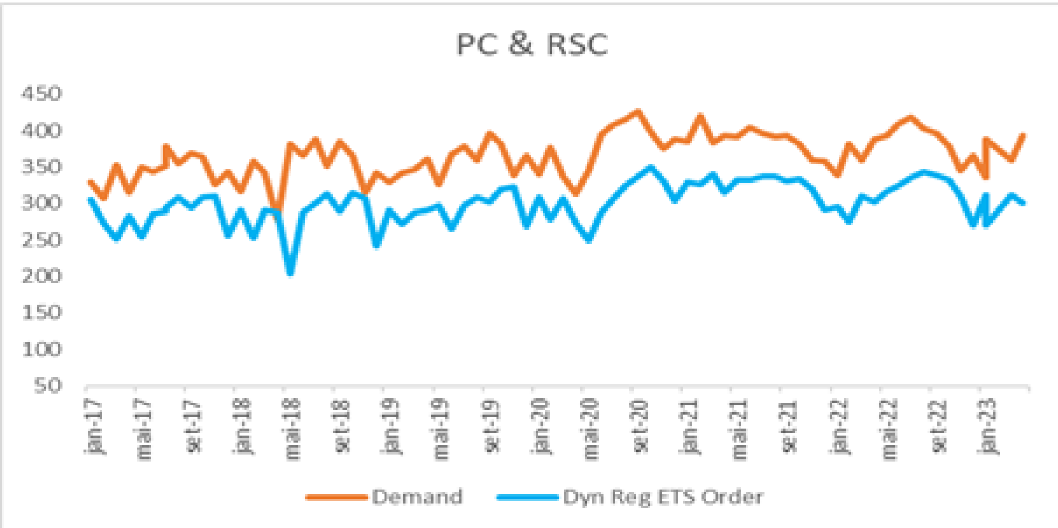

Figure 4.8: PC & RSC scenario demand and Dynamic regression ETS order quantity

The downward distortions presented on the series of orders over time in Figure 4.7 can be explained by previous instances where the demand was overestimated, resulting in overstocking, which led to an order lower than usual on the next iteration.

As can be seen in Figure 4.8, the time series of orders turns almost into a linear transformation of the real demand time series, with the difference in the level being attributed to the extra demand generated by PC, in most cases supplied by the RSC. Leading to negligible, if any, information distortion in variance and behavior.

The overall impact of the scenarios can be assessed by the summary Table 4.3. Service level was not statistically significantly affected, and it would not be fair to compare it between scenarios of distinguished demand levels.

Standard	**PC**
Average Profit Increase from RSC	
+4.1%	+8.9%
Average Cost Reduction from RSC	
-11.2%	-11.8%
Average BWE Reduction from RSC	
-51.4%	-32.8%

Table 4.6: Scenario transition benefits

5
Conclusions

This dissertation aimed to tackle the main limitations and consequences of the integration of human agents into a supply chain management decision-making process. The main limitations pointed out by Sanders and Wood (2019) are non-scalability and cognitive biases, for the scalability approach, the usage of time series clustering techniques such as dynamic time warping is well documented, however, the clustering techniques do not consider the operational impacts of the clustering itself.

Thus an approach was devised for a relevance-based clustering that, despite being inefficient for dimensionality reduction, proved itself valuable as a prevention mechanism against the negative operational consequences of applying clustering algorithms, concerned mostly with metrics such as the silhouette score.

The time series judged as too relevant to be appropriately placed within a cluster would be evaluated by the supply chain manager. For them, a rational model would be defined, the concept of a rational model was proposed as the model that would lead to the most satisfactory outcome in a multi-criteria analysis, since supply chain management cannot be constrained to a single basic metric as profit, cost, accuracy, or the BWE.

Therefore, the following research questions were addressed.

1. How to increase the scalability of the human expert for a supply chain forecasting system? To increase the scalability of the human expert, a new methodology to validate the effect of clustering was implemented, allowing for the measurement of the operational consequences created by the usage of the cluster for decision-making.

2. How to avoid the impacts of cognitive bias in human judgmental adjustments in supply chain forecasting? The impact of the cognitive biases in judgmental adjustments in supply chain forecasting was addressed by implementing an ontology-driven automatic risk seeking bias detection mechanism, that also informs the user of the likely consequences of their biased decision by solving a robust optimization problem.

3. How to integrate a more holistic organizational context with supply

> chain forecasting supported by data-aware statistical models? The more holistic organizational context if integrated through the usage of a rational model, the forecasting model that would lead to the most satisfactory outcome considering the companies strategic goals, this occurs by evaluating the decisions the forecasting models would generate when used to solve a robust optimization problem, whose the likely outcome is evaluated in a multi-criteria approach.

This new forecasting model selection methodology was achieved by observing the performance of simulated decisions within a DDU RO profit-driven problem and measuring the performances according to the MCDM weights for the relevant individual criteria. This new technique not only demonstrated potential undesired consequences such as BWE propagation out of traditional accuracy-driven forecasting model selection. The application of this new method provided an integrated view of the positive impacts of reverse supply chain logistics and preferential client acquisition with information sharing. The proposed methodology for forecasting model selection can be beneficial in a multitude of ways and contexts.

Once the rational model was established, the ARSBD system could be implemented, as a means of detecting the presence of risk-seeking biases based on prospect theory (Kahneman, 2012) using profit as a decision metric for the probability distribution of scenarios. The ontology utilized had critical importance in the definition of the cognitive biases, thus allowing for the formulation of a precise and effective automatic detection mechanism

The results demonstrate that the system intervention was capable of improving the performance beyond the expert judgmental adjustment alone and also surpassing the usage of the pure statistical model. Further improvements in the decision simulations could have been achieved by using the satisfaction metric as the objective function of the DDU RO problem, however, due to time constraints this research could not yet be concluded, this step would guarantee optimality and possibly reveal optimal decisions the algorithms employed may have missed for being within a reduced profit subset. However, for this second implementation, it would be required to write the multi-criteria weights of each criterion as the objective function of the robust optimization problem, this would be challenging since while being computed the decision-dependent parameters are in an indeterminate state, therefore operations required for normalization cannot be executed, which is crucial for MCDM. It is hard to estimate the impact of these changes to be done within future works but the current implementation presents a considerably robust and consistent performance.

In a nutshell, the main differences of the proposed method and current literature are in the multi-criteria approach to decision-making in supply chains, while utilizing robust optimization, thus mitigating the limitations of mono and multi criteria techniques. Furthermore, this more holistic view of a forecasting model, evaluating it beyond accuracy is a direction recently inclined within forecasting literature. The automatic detection of biases represents a novel way to address the cognitive biases inherent to human decision-makers' inputs as judgmental forecasts, and the proposed system would demonstrate the impact of the decisions. Finally, the evaluation of the impact caused in the usage of clustering techniques is a relevant aspect brought by the relevance-based clustering, thus preventing loss in performance in the pursuit of scalability.

Our proposed framework was evaluated in two experiments: the first used M5, a competition dataset from the literature, and the second used real data about cardboard production in Brazil, from 2017 to 2023. The experimental results demonstrated that the system intervention was capable of improving the performance beyond the expert judgmental adjustment alone and also surpassing the usage of the pure statistical model. The framework implementation also had a considerably robust and consistent performance.

5.1 Limitations and Future work

Although we obtained positive results in evaluating the application of the proposed framework (and its underlying methodology) in two relevant scenarios, we acknowledge some limitations of this research such as the limited capacity for generalization. A limitation of this work is presented in the assumption of proper forecasting models given the behavior of the time series analyzed. Therefore, if the output of a given forecasting model results in a normal distribution of potential events but this is not consistent with the behavior of the time series, such as would be the case for intermittent time series, the assumptions of the robust optimizer will be false. That being said, a basic understanding of time series analysis is required to apply the framework, even if cases using inappropriate model candidates are likely to result in poor performance. The framework contemplated an analysis of the impact of using certain forecasting models and the simulated human interference as input for robust optimizers serving as decision-making proxies, that being said the evaluation of the models is done after the fact to determine what model should be used in the future, despite the forecasts being done at every step there is no evaluation of the models until the end, therefore a satisfaction-driven step by step model selection was not contemplated in the formulation, but could be easily adapted.

Some directions of future work could prove beneficial as a complementary approach to the developments presented:

- A satisfaction-driven robust optimization formulation, as aforementioned, can reveal potential optimal satisfaction-wise solutions, and be of considerable use for multi-criteria-related problems, such as wallet composition problems that take into account sector diversification, and similarly structured problems such as multi-criteria supplier selection.
- Implement the satisfaction-driven analysis of the forecasting model as a step by step feature.
- The comparison of the framework integration along multiple agents may be able to better demonstrate the impact of the MCDM approach to forecasting model selection.
- The integration of other biases, especially the anchoring bias, could be of great significance since, alongside overconfidence, are some of the most well-documented biases within forecasting.
- Many cognitive biases have a moderation relation to knowledge level, including both of the previously mentioned. The lack of statistical capabilities of most managers to fully grasp the forecasting models is a well-documented aspect, tools like ontologies can provide the needed context and serve as an educational tool for such cases.
- The implementation and testing of the framework as a DSS used in real world cases.

5.2 Scientific Contributions and Awards

Among the contributions throughout the research are the following papers:

1. IPSERA 2024 - Data-Driven & Cognitively Aware Supply Chain Management. Authors: Mateus Peixoto; Fernanda Baião. (At this stage the judgmental adjustment step was not yet implemented.)

2. IPSERA 2024 - Multicriteria forecasting model selection. Authors: Mateus Peixoto; Guilherme Saboya; Rodrigo Caiado.

3. 3. IPSERA 2024 - Relevance of forecasting model selection and reverse supply chain performance under uncertain times. Authors: Mateus Peixoto; Bruna Santiago; Fernando Oliveira. This work was developed during the Time Series Analysis course lectured by Professor Fernando Oliveira, which led to the selection for the International Institute of Forecasters Student Award 2023.

4. ENEGEP 2024 - Towards a Framework for Data-Driven & Cognitively-Aware Supply Chain Management. Authors: Mateus Peixoto; Fernanda Baião.

5. SBPO 2024 - Supply Chain Management In a Weak Dominance Coopetition Environment Under Decision-Dependent Uncertainty. Authors: Mateus Peixoto; Bruno dos Santos; Leonardo Blois. This paper was selected among the 5 finalists for the Roberto Diéguez Galvão award for best paper in English.

Alongside the following contributions: (i) The definition of a rational reference point for supply chain order quantity decisions, (ii) an automatic ontology-driven robust decision-dependent system for risk seeking bias detection, (iii) a mechanism to measure the potential operational impact of clustering algorithms as dimensionality reduction, (iv) a new methodology to evaluate forecasting models, (v) evidence for the role of preferred client communication and reverse supply chain in the mitigation of the bullwhip effect, (vi) a new formulation for decision-dependent uncertainty robust optimization applied to supply chain management decision-making problems, (vii) an analysis of the effectiveness of coopetition and information sharing in the mitigation of the bullwhip effect, (viii) and an experiment proving that the integration of statistical systems and human insight can result in better outcomes than either one isolated, as long as the cognitive biases involved can be mitigated.

Bibliography

Alkhudary, R., Queiroz, M. M., and Féniès, P. (2024). Mitigating the risk of specific supply chain disruptions through blockchain technology. *Supply Chain Forum: An International Journal*, 25(1):1–11.

Almeida, M. M. K. D., Marins, F. A. S., Salgado, A. M. P., Santos, F. C. A., and Silva, S. L. D. (2017). <b>The importance of trust and collaboration between companies to mitigate the bullwhip effect in supply chain management. *Acta Scientiarum. Technology*, 39(2):201.

Aouni, B. and Laflamme, S. (2014). From mono-criterion to multi-criteria decision aid: a necessary but unfinished evolution in operational research. *International Journal of Applied Decision Sciences*, 7(2):123.

Arvan, M., Fahimnia, B., Reisi, M., and Siemsen, E. (2019). Integrating human judgement into quantitative forecasting methods: A review. *Omega*, 86:237–252.

Beal Partyka, R. (2022). Supply chain management: an integrative review from the agency theory perspective. *Revista de Gestão*, 29(2):175–198.

Brauch, M., Mohaghegh, M., and Größler, A. (2024). Causes of the bullwhip effect: a systematic review and categorization of its causes. *Management Research Review*.

Chen, L., Dong, T., Peng, J., and Ralescu, D. (2023). Uncertainty Analysis and Optimization Modeling with Application to Supply Chain Management: A Systematic Review. *Mathematics*, 11(11):2530.

Davis, A. M. (2018). Biases in Individual Decision-Making. In Donohue, K., Katok, E., and Leider, S., editors, *The Handbook of Behavioral Operations*, pages 149–198. Wiley, 1 edition.

De Almeida, M. M. K., Marins, F. A. S., Salgado, A. M. P., Santos, F. C. A., and Da Silva, S. L. (2015). Mitigation of the bullwhip effect considering trust and collaboration in supply chain management: a literature review. *The International Journal of Advanced Manufacturing Technology*, 77(1-4):495–513.

De Neys, W. (2021). On Dual- and Single-Process Models of Thinking. *Perspectives on Psychological Science*, 16(6):1412–1427.

D'Urso, D., Di Mauro, C., Chiacchio, F., and Compagno, L. (2015). Modelling Human Behaviour in Newsvendor Game. *IFAC-PapersOnLine*, 48(3):610–615.

Durán Peña, J. A., Ortiz Bas, , and Reyes Maldonado, N. M. (2021). Impact of Bullwhip Effect in Quality and Waste in Perishable Supply Chain. *Processes*, 9(7):1232.

Fahimnia, B., Pournader, M., Siemsen, E., Bendoly, E., and Wang, C. (2019). Behavioral Operations and Supply Chain Management–A Review and Literature Mapping. *Decision Sciences*, 50(6):1127–1183.

Fayezi, S., O'Loughlin, A., and Zutshi, A. (2012). Agency theory and supply chain management: a structured literature review. *Supply Chain Management: An International Journal*, 17(5):556–570.

Fildes, R. and Goodwin, P. (2007). Against Your Better Judgment? How Organizations Can Improve Their Use of Management Judgment in Forecasting. *Interfaces*, 37(6):570–576.

Fildes, R., Goodwin, P., Lawrence, M., and Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting*, 25(1):3–23.

Goel, V. and Grossmann, I. E. (2006). A Class of stochastic programs with decision dependent uncertainty. *Mathematical Programming*, 108(2-3):355–394.

Goltsos, T. E., Syntetos, A. A., Glock, C. H., and Ioannou, G. (2022). Inventory – forecasting: Mind the gap. *European Journal of Operational Research*, 299(2):397–419.

Gomes, L. F. A. M., Machado, M. A. S., and Rangel, L. A. D. (2013). Behavioral multi-criteria decision analysis: the TODIM method with criteria interactions. *Annals of Operations Research*, 211(1):531–548.

Goodwin, P. (2000). Improving the voluntary integration of statistical forecasts and judgment. *International Journal of Forecasting*, 16(1):85–99.

Goodwin, P. and Fildes, R. (2011). Forecasting in supply chain companies: Should you trust your judgment? *OR Insight*, 24(3):159–167.

Goodwin, P., Petropoulos, F., and Hyndman, R. J. (2017). A note on upper bounds for forecast-value-added relative to naïve forecasts. *Journal of the Operational Research Society*, 68(9):1082–1084.

Goodwin, P. and Wright, G. (1993). Improving judgmental time series forecasting: A review of the guidance provided by research. *International Journal of Forecasting*, 9(2):147–161.

Goudarzi, F. S., Bergey, P., and Olaru, D. (2023). Behavioral operations management and supply chain coordination mechanisms: a systematic review and classification of the literature. *Supply Chain Management: An International Journal*, 28(1):140–161.

Grayot, J. (2020). Dual process theories in behavioral economics and neuroeconomics: a critical review. *Review of Philosophy and Psychology*, 11(1):105–136.

Guarino, N., Oberle, D., and Staab, S. (2009). What Is an Ontology? In Staab, S. and Studer, R., editors, *Handbook on Ontologies*, pages 1–17. Springer Berlin Heidelberg, Berlin, Heidelberg.

Guizzardi, G. and Guarino, N. (2024). Explanation, semantics, and ontology. *Data & Knowledge Engineering*, 153:102325.

Hertel, G., Meeßen, S. M., Riehle, D. M., Thielsch, M. T., Nohe, C., and Becker, J. (2019). Directed forgetting in organisations: the positive effects of decision support systems on mental resources and well-being. *Ergonomics*, 62(5):597–611.

Houlihan, J. B. (1985). International Supply Chain Management. *International Journal of Physical Distribution & Materials Management*, 15(1):22–38.

Hyndman, R. and Athanasopoulos, G. (2018). *Forecasting: Principles and Practice*. OTexts, Australia, 2nd edition.

Hüttinger, L., Schiele, H., and Veldman, J. (2012). The drivers of customer attractiveness, supplier satisfaction and preferred customer status: A literature review. *Industrial Marketing Management*, 41(8):1194–1205.

IPEA, B. (2024). IPEADATA Cardboard.

Kahneman, D. (2012). *Thinking, fast and slow*. Penguin psychology. Penguin Books, London.

Keliji, P. B., Aghajani, H. A., Movahedi, M. M., and Shayannia, S. A. (2022). The Analysis of the Role of Bullwhip Effects on the Four-Level Supply Chain in Industry Using Statistical Methods. *Discrete Dynamics in Nature and Society*, 2022:1–16.

Kocbek, M., Jost, G., Hericko, M., and Polancic, G. (2015). Business process model and notation: The current state of affairs. *Computer Science and Information Systems*, 12(2):509–539.

Kolassa, S., Rostami-Tabar, B., and Siemsen, E. (2023). *Demand Forecasting for Executives and Professionals.* Chapman and Hall/CRC, Boca Raton, 1 edition.

Lawrence, M., Goodwin, P., and Fildes, R. (2002). Influence of user participation on DSS use and decision accuracy. *Omega*, 30(5):381–392.

Lawrence, M., Goodwin, P., O'Connor, M., and Önkal, D. (2006). Judgmental forecasting: A review of progress over the last 25years. *International Journal of Forecasting*, 22(3):493–518.

Lee, H. L., Padmanabhan, V., and Whang, S. (1997). Information Distortion in a Supply Chain: The Bullwhip Effect. *Management Science*, 43(4):546–558.

Lee, J.-H. and Kim, C.-O. (2008). Multi-agent systems applications in manufacturing systems and supply chain management: a review paper. *International Journal of Production Research*, 46(1):233–265.

Lin, V. S. (2013). Improving Forecasting Accuracy by Combining Statistical and Judgmental Forecasts in Tourism: . *Journal of China Tourism Research*, 9(3):325–352.

Ma, S., Fildes, R., and Huang, T. (2016). Demand forecasting with high dimensional data: The case of sku retail sales forecasting with intra- and inter-category promotional information. *European Journal of Operational Research*, 249(1):245 – 257. Cited by: 106; All Open Access, Green Open Access.

Madanchian, M. and Taherdoost, H. (2023). A comprehensive guide to the TOPSIS method for multi-criteria decision making. *Sustainable Social Development*, 1(1).

Makridakis, S., Spiliotis, E., and Assimakopoulos, V. (2022). M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting*, 38(4):1346–1364.

Mannes, A. E. and Moore, D. A. (2013). A Behavioral Demonstration of Overconfidence in Judgment. *Psychological Science*, 24(7):1190–1197.

McInnes, L. and Healy, J. (2017). Accelerated Hierarchical Density Based Clustering. In *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, pages 33–42, New Orleans, LA. IEEE.

Monroe, R. W. (2012). Nada R.Sanders. Supply Chain Management: A Global Perspective. Hoboken, NJ: John Wiley and Sons. Hardcover, 428 pages, copyright 2012, ISBN: 978-0-470-14117-5, US $199.95; E-book, copyright 2011, ISBN: 978-0-470-91395-6, US $119.50. *Transportation Journal*, 51(4):506–508.

Moosavi, J., Fathollahi-Fard, A. M., and Dulebenets, M. A. (2022). Supply chain disruption during the COVID-19 pandemic: Recognizing potential disruption management strategies. *International Journal of Disaster Risk Reduction*, 75:102983.

Moritz, B., Siemsen, E., and Kremer, M. (2014). Judgmental Forecasting: Cognitive Reflection and Decision Speed. *Production and Operations Management*, 23(7):1146–1160.

M.T. Thomé, A., Soucasaux Sousa, R., and F.R.R.S. Do Carmo, L. (2014). Complexity as contingency in sales and operations planning. *Industrial Management & Data Systems*, 114(5):678–695.

Neto, J. M. and Salomon, V. A. P. (2022). Multi-criteria Analysis of Disruption Risks for Supply Chains Due to Pandemics. In Qudrat-Ullah, H., editor, *Understanding the Dynamics of New Normal for Supply Chains*, pages 121–137. Springer International Publishing, Cham. Series Title: Understanding Complex Systems.

Nikolopoulos, K., Punia, S., Schäfers, A., Tsinopoulos, C., and Vasilakis, C. (2021). Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions. *European Journal of Operational Research*, 290(1):99–115.

Oliveira, J. B., Lima, R. S., and Montevechi, J. A. B. (2016). Perspectives and relationships in Supply Chain Simulation: A systematic literature review. *Simulation Modelling Practice and Theory*, 62:166–191.

Opricovic, S. and Tzeng, G.-H. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. *European Journal of Operational Research*, 156(2):445–455.

Perera, H. N., Hurley, J., Fahimnia, B., and Reisi, M. (2019). The human factor in supply chain forecasting: A systematic review. *European Journal of Operational Research*, 274(2):574–600.

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Ben Taieb, S., Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., Browell, J., Carnevale, C., Castle, J. L., Cirillo, P., Clements, M. P., Cordeiro, C., Cyrino Oliveira, F. L., De Baets, S., Dokumentov, A., Ellison, J., Fiszeder, P., Franses, P. H., Frazier, D. T., Gilliland, M., Gönül, M. S., Goodwin, P., Grossi, L., Grushka-Cockayne, Y., Guidolin, M., Guidolin, M., Gunter, U., Guo, X., Guseo, R., Harvey, N., Hendry, D. F., Hollyman, R., Januschowski, T., Jeon, J., Jose, V. R. R., Kang, Y., Koehler, A. B., Kolassa, S., Kourentzes, N., Leva, S., Li, F., Litsiou, K., Makridakis, S., Martin, G. M., Martinez, A. B., Meeran, S., Modis, T., Nikolopoulos, K., Önkal, D., Paccagnini, A., Panagiotelis, A., Panapakidis, I., Pavía, J. M., Pedio, M., Pedregal, D. J., Pinson, P., Ramos, P., Rapach, D. E., Reade, J. J., Rostami-Tabar, B., Rubaszek, M., Sermpinis, G., Shang, H. L., Spiliotis, E., Syntetos, A. A., Talagala, P. D., Talagala, T. S., Tashman, L., Thomakos, D., Thorarinsdottir, T., Todini, E., Trapero Arenas, J. R., Wang, X., Winkler, R. L., Yusupova, A., and Ziel, F. (2022). Forecasting: theory and practice. *International Journal of Forecasting*, 38(3):705 – 871. Type: Review.

Petropoulos, F., Kourentzes, N., Nikolopoulos, K., and Siemsen, E. (2018). Judgmental selection of forecasting models. *Journal of Operations Management*, 60(1):34–46.

Petrou, P., Demerouti, E., Peeters, M. C. W., Schaufeli, W. B., and Hetland, J. (2012). Crafting a job on a daily basis: Contextual correlates and the link to work engagement. *Journal of Organizational Behavior*, 33(8):1120–1141.

Ramos, E. d. C., Campos, M. L. M., and Baião, F. (2024). ABI Approach: Automatic Bias Identification in Decision-Making Under Risk based in an Ontology of Behavioral Economics. Version Number: 1.

Ramos, P., Oliveira, J. M., Kourentzes, N., and Fildes, R. (2022). Forecasting Seasonal Sales with Many Drivers: Shrinkage or Dimensionality Reduction? *Applied System Innovation*, 6(1):3.

Reimann, F., Kosmol, T., and Kaufmann, L. (2017). Responses to Supplier-Induced Disruptions: A Fuzzy-Set Analysis. *Journal of Supply Chain Management*, 53(4):37–66.

Reimers, S. and Harvey, N. (2024). Bars, lines and points: The effect of graph format on judgmental forecasting. *International Journal of Forecasting*, 40(1):44–61.

Rezaei, J. (2015). Best-worst multi-criteria decision-making method. *Omega*, 53:49–57.

Sales, T. P., Baião, F., Guizzardi, G., Almeida, J. P. A., Guarino, N., and Mylopoulos, J. (2018). The Common Ontology of Value and Risk. In Trujillo, J. C., Davis, K. C., Du, X., Li, Z., Ling, T. W., Li, G., and Lee, M. L., editors, *Conceptual Modeling*, volume 11157, pages 121–135. Springer International Publishing, Cham. Series Title: Lecture Notes in Computer Science.

Sanders, N. R. and Wood, J. D. (2019). *The Humachine: Humankind, Machines, and the Future of Enterprise*. Routledge, 1 edition.

Schorsch, T., Wallenburg, C. M., and Wieland, A. (2017). The human factor in SCM: Introducing a meta-theory of behavioral supply chain management. *International Journal of Physical Distribution & Logistics Management*, 47(4):238–262.

Skipworth, H., Godsell, J., Wong, C. Y., Saghiri, S., and Julien, D. (2015). Supply chain alignment for improved business performance: an empirical study. *Supply Chain Management: An International Journal*, 20(5):511–533.

Smith, C. and Fatorachian, H. (2023). COVID-19 and Supply Chain Disruption Management: A Behavioural Economics Perspective and Future Research Direction. *Journal of Theoretical and Applied Electronic Commerce Research*, 18(4):2163–2187.

Sousa, R. and Voss, C. A. (2008). Contingency research in operations management practices. *Journal of Operations Management*, 26(6):697–713.

Syntetos, A., Babai, M., Davies, J., and Stephenson, D. (2010). Forecasting and stock control: A study in a wholesaling context. *International Journal of Production Economics*, 127(1):103–111.

Syntetos, A. A., Babai, Z., Boylan, J. E., Kolassa, S., and Nikolopoulos, K. (2016). Supply chain forecasting: Theory, practice, their gap and the future. *European Journal of Operational Research*, 252(1):1–26.

Tavares Thomé, A. M., Scavarda, L. F., Fernandez, N. S., and Scavarda, A. J. (2012). Sales and operations planning: A research synthesis. *International Journal of Production Economics*, 138(1):1–13.

Thomas, R. and Skinner, L. (2010). Total Trust and Trust Asymmetry: Does Trust Need to Be Equally Distributed in Interfirm Relationships? *Journal of Relationship Marketing*, 9(1):43–53.

Tversky, A. and Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, 185(4157):1124–1131.

Van Oudenhoven, B., Van De Calseyde, P., Basten, R., and Demerouti, E. (2023). Predictive maintenance for industry 5.0: behavioural inquiries from a work system perspective. *International Journal of Production Research*, 61(22):7846–7865.

Wang, J. (2023). Demand forecasting in SCM for automobile industry based on time-series modeling. In Jin, S. and Dai, W., editors, *Second International Conference on Statistics, Applied Mathematics, and Computing Science (CSAMCS 2022)*, page 125, Nanjing, China. SPIE.

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

Wang, X., Hyndman, R. J., Li, F., and Kang, Y. (2023). Forecast combinations: An over 50-year review. *International Journal of Forecasting*, 39(4):1518–1547.

Wang, X., Smith, K., and Hyndman, R. (2006). Characteristic-Based Clustering for Time Series Data. *Data Mining and Knowledge Discovery*, 13(3):335–364.

Wong, C., Skipworth, H., Godsell, J., and Achimugu, N. (2012). Towards a theory of supply chain alignment enablers: a systematic literature review. *Supply Chain Management: An International Journal*, 17(4):419–437.

Yamini, S. (2023). Impact of loss aversion on the newsvendor problem: a literature review and insights for future researchers. *OPSEARCH*, 60(4):1926–1950.

Yang, Y., Lin, J., Liu, G., and Zhou, L. (2021). The behavioural causes of bullwhip effect in supply chains: A systematic literature review. *International Journal of Production Economics*, 236:108120.

Yin, X., Cheng, L., Wang, X., Lu, J., and Qin, H. (2019). Optimization for Hydro-Photovoltaic-Wind Power Generation System Based on Modified Version of Multi-Objective Whale Optimization Algorithm. *Energy Procedia*, 158:6208–6216.

Zanddizari, M., Tavakkoli-Moghaddam, R., and Azaron, A. (2019). Modeling stock-out loss and overstocking loss generated by bullwhip effect. *Scientia Iranica*, 26(3):1913 – 1924. Cited by: 2; All Open Access, Bronze Open Access.

Zhang, X. (2004). The impact of forecasting methods on the bullwhip effect. *International Journal of Production Economics*, 88(1):15–27.

Zhao, S. and You, F. (2019). Resilient supply chain design and operations with decision-dependent uncertainty using a data-driven robust optimization approach. *AIChE Journal*, 65(3):1006–1021.