UNIVERSITÉ DE LILLE - CRIStAL
ÉCOLE DOCTORALE SCIENCES POUR L'INGÉNIEUR

Cadre Général et Méthodes d'Optimisation pour la Planification de Stock en Contextes Industriels

General Framework and Optimization Methods for Stochastic Replenishment Planning in Industrial Contexts

SAHU Rabin Kumar

Thèse préparée et soutenue publiquement le ***27 Août 2020****, en vue de l'obtention du grade de* Docteur en Informatique.

Jury:

M.	DAVY Manuel	Vekia SAS	Coencadrant
Mme.	DHAENENS Clarisse	Université de Lille	Directrice
M.	DOLGUI Alexandre	IMT Atlantique	Rapporteur
Mme.	DUCHIEN Laurence	Université de Lille	Présidente du Jury
M.	PENZ Bernard	INP Grenoble	Rapporteur
M.	VEERAPEN Nadarajen	Université de Lille	Coencadrant
M.	YALAOUI Farouk	Université de Technologie de Troyes	Examinateur

To Maa, Bapa and Itu.

Acknowledgements

Having a PhD was my dream, and without the support and motivation I received, it would not have been possible. Here, I take the opportunity to thank everyone involved during this journey who propelled me forward.

First of all, I extend my heartiest gratitude to my advisor Prof. Clarisse Dhaenens. She guided me throughout my research work and motivated me to remain focused. Her insights and immense subject knowledge helped me hone my work and produce something novel. She consistently inspired me to push boundaries and realize my true potential. Even more so, she cared for my wellbeing. I am lucky to have had her as my advisor and it could not have been better.

On an equal tone, I am also grateful to Dr. Manuel Davy, CEO, Vekia SAS, who was one of the cornerstones of this work. He recognized me and my potential, and motivated me to pursue my doctoral studies. Starting from identifying the problem to the insightful discussions on newer and cutting-edge technologies in the area of supply chain, I could gain knowledge in every aspect.

This dissertation would not have been complete without the critical review of the jury members. I am particularly thankful to Prof. Bernard Penz from INP Grenoble and Prof. Alexandre Dolgui from IMT Atlantique for their valuable time towards the review of my work as referees and for their suggestions. I am also thankful to Prof. Laurence Duchien for her acceptance to become the president of the jury, for her careful examination of my work and the seamless arrangement of the defense. I also extend my sincere gratitude to Prof. Farouk Yalaoui from the Université de Technologie de Troyes for and Dr. Nadarajen Veerapen from the Université de Lille for their valuable suggestions.

The last three years was a great learning opportunity for me. Being a CIFRE project, it was particularly helpful in gaining both academic and industrial knowledge at the same time. I would like to thank Hervé Lemai and Dr. Alexandre Gerussi from Vekia SAS for their relentless support and critical industrial evaluation of my work. This work is also fruitful because of Dr. Nadarajen Veerapen who provided me with all the needed support and discussed various new ways of addressing the targeted problems.

Last but not the least, I am also thankful to all the team members at Vekia SAS and all the members of ORKAD research team for their continuous support during the research period.

Contents

List of Figures x

List of Tables xiii

List of Abbreviations xiv

General Introduction of the Thesis xvi

I Preliminaries, State-of-the-Art, Classification and Evaluation 2

1 Context and Motivations 3

1.1 Introduction 3
1.2 About Vekia 4
1.2.1 Solutions 4
1.2.2 Vision 5
1.3 Context of Supply Chain Planning 7
1.3.1 Supply Chain Management 7
1.3.2 Decision Levels 8
1.3.3 Inventory Management 10
1.4 Motivations 11
1.4.1 Impacts of Inventory Management 11
1.4.2 Problem Diversity and Challenges 12
1.4.3 Limitations of Existing Methods 13
1.4.4 Supply Chain Planning in Future 13
1.5 Problem Statement and Research Methodology 14
1.6 Conclusions 15

2 State-of-the-Art and Classification of Inventory Optimization Problems 16

2.1 Introduction 16
2.2 State-of-the-Art: Inventory Optimization 17
2.2.1 General Inventory Optimization 17
2.2.2 Classification of Inventory Optimization Problems 19
2.2.3 Performance Assessment of IMSs 19

2.2.4 Single-Echelon Inventory Optimization . . . 20
2.2.5 Multi-Echelon Inventory Optimization . . . 23
2.2.6 Position of Our Work . . . 25
2.3 Classification Scheme for Inventory Optimization Problems . . . 26
2.3.1 Motivations and Objectives . . . 28
2.3.2 Classification Scheme . . . 28
2.3.3 Classification Examples . . . 31
2.3.4 Recent Practices in Inventory Management . . . 32
2.4 Conclusions . . . 34

3 Performance Assessment of Inventory Management Systems 35
3.1 Introduction . . . 35
3.2 Motivations and Objectives . . . 36
3.3 Preliminaries . . . 37
3.4 KPIs for Inventory Management Systems . . . 39
3.4.1 Financial KPIs . . . 40
3.4.2 Operational KPIs . . . 42
3.4.3 Service KPIs . . . 44
3.5 Computation Methodology . . . 45
3.6 Conclusions . . . 48

II Problem Definitions and Solution Methods 49

4 Problem Definitions and Modular Framework 50
4.1 Introduction . . . 50
4.2 Solution Design Framework . . . 51
4.2.1 The Supply Chain Network and Planning Problems . . . 51
4.2.2 Modular Inventory Management Framework . . . 52
4.3 Single-Item Inventory Optimization . . . 53
4.3.1 Single-item Replenishment Problem (SRP) . . . 53
4.3.2 Single-item Replenishment Problem with Supplier Selection (SRPSS) . 54
4.4 Multi-Item Inventory Optimization . . . 55
4.4.1 Joint Replenishment Problem (JRP) with Single Supplier . . . 56
4.4.2 Joint Replenishment Problem with Supplier Selection (JRPSS) . . . 56
4.4.3 Promotional Joint Replenishment Problem (PJRP) . . . 57
4.5 Industrial Extensions . . . 58
4.6 Conclusions . . . 58

5 Sampling-based Replenishment Planning in Single-Item Inventory Systems 59
5.1 Introduction . . . 59
5.2 Context and Motivations . . . 61
5.2.1 Context . . . 61

5.2.2 Motivations . . . 61
5.3 Related Literature . . . 62
5.3.1 Stochastic Inventory Optimization . . . 62
5.3.2 Sampling-Based Stochastic Optimization . . . 63
5.3.3 Sampling-Based Stochastic Inventory Optimization . . . 63
5.4 Preliminaries . . . 64
5.5 Sampling-based Optimization Models . . . 65
5.5.1 Expected Cost Approach (ECA) . . . 67
5.5.2 Robust Cost Approach (RCA) . . . 67
5.5.3 Immediate Expected Cost Approach (IECA) . . . 68
5.6 Heuristic Approaches . . . 70
5.6.1 Enumerative Search Heuristic . . . 70
5.6.2 Convexity Results and Optimized Heuristic . . . 71
5.7 Experimental Protocol . . . 73
5.8 Results of Numerical Experiments . . . 75
5.8.1 Approximation Accuracy for the OH-IECA . . . 75
5.8.2 Tests on Synthetic Problem Instances . . . 76
5.8.3 Tests on Benchmark Problem Instances . . . 79
5.8.4 Tests on Non-stationary Demand . . . 80
5.9 Extensions . . . 81
5.9.1 Addressing Batch Size Constraint . . . 81
5.9.2 Robust Solution . . . 82
5.9.3 Addressing Minimum Order Quantity . . . 83
5.10 Conclusions . . . 84

6 Sampling-based Replenishment Planning in Non-Stationary Multi-Supplier Inventory Systems 86
6.1 Introduction . . . 86
6.2 Preliminaries . . . 87
6.3 SRP with Non-Stationary Demand (Single-Supplier) . . . 88
6.3.1 Ordering Options . . . 89
6.3.2 Reduced-state Dynamic Program (RDP) . . . 90
6.4 SRP with Supplier Selection (SRPSS) . . . 91
6.4.1 Context and Motivations . . . 91
6.4.2 Problem Formulation . . . 93
6.4.3 Common Supplier Selection . . . 93
6.4.4 Dynamic Supplier Selection . . . 94
6.4.5 RDP Framework for the SRPSS . . . 95
6.5 Numerical Experiments . . . 97
6.5.1 Experimental Protocol . . . 97
6.5.2 Numerical Results . . . 100
6.6 Conclusions . . . 102

7 Promotional Replenishment Planning in Multi-item Inventory Systems 104
7.1 Introduction 104
7.2 Problem Description 106
7.3 Literature Review 107
7.4 Mathematical Formulation 108
7.5 Methodology 111
7.5.1 Inventory Classification and Minimum Service Level Allocation 111
7.5.2 Multi-Objective Optimization with probabilistic forecast 112
7.6 Case Study 1 113
7.6.1 Input Data and Experimental Setting 113
7.6.2 Results and Interpretation 115
7.7 Problem Reformulation 116
7.7.1 Optimization Model 117
7.7.2 Complexity Analysis 117
7.8 Metaheuristic Approach 118
7.8.1 Initial Solution 119
7.8.2 Local Search 120
7.8.3 Guided Search Heuristic 121
7.8.4 Perturbation and Iterated Local Search 123
7.9 Numerical Study for the Metaheuristic 125
7.9.1 Comparative Performances on GH1 and GH2 125
7.9.2 Performance of IGSH 128
7.10 Case Study 2 131
7.10.1 Data and Experimental Protocol 131
7.10.2 Results and Discussion 132
7.11 Conclusions 133

8 Extensions, Industrialization, Conclusions and Perspectives 135
8.1 Summary of the Main Contributions 135
8.2 Extensions and Industrialization 137
8.3 Perspectives 137

Appendices 153
A Proof for Property 1 154
B Explanation of Equation 5.25 156
C Additional Results from Chapter 5 160
D Additional Heuristic for Chapter 7 162
E Additional Plots for Chapter 7 163

List of Figures

1 Arrangement of the dissertation. xviii

1.1 A general structure of current supply chains. 7
1.2 A general structure of future supply chains. 8
1.3 Supply chain planning: Decisions across different levels and functional areas. 9
1.4 A 3-echelon retail supply chain. 11
1.5 Industrial implementation plan. 14

2.1 Summary of reviewed literature. 17
2.2 Proposed classification scheme. 29
2.3 Different types of supply chain networks. 29

3.1 Workflow of an inventory management system. Generally IMSs comprise two sub-systems: forecasting system and replenishment planning system. The forecasting systems uses historical sales and exogenous data to predict future demand using machine learning algorithms. The replenishment planning system uses those forecasts and the ordering constraints to suggest optimized replenishment decisions. 38
3.2 The dynamics of an inventory system. Various basic events of an inventory system are depicted. The flow of events are, recording opening inventory, order delivery, demand realization, and recording closing inventory or shortages. 38
3.3 Relative assessment of two inventory management systems. The comparison can be done between two new IMSs or between new and existing systems. . . 39
3.4 Evolution of inventory with time 39
3.5 Profitability model 40
3.6 Components value chain profitability 41
3.7 Shortages are expected where inventory and sales are close. 46
3.8 Flowchart for the proposed simulation method. 47

4.1 A generalized depiction of multi-echelon supply chain. 51
4.2 Proposed inventory management framework. 52
4.3 Single supplier single item replenishment problem. 53
4.4 Multi-supplier single item replenishment problem. 55
4.5 Joint replenishment problem with a single supplier. 56
4.6 Joint replenishment planning with supplier selection. 57

4.7 Illustration of differentiation between single-item pack and multi-item pack during promotional ordering. . . . 58

5.1 Illustration of an ordering mechanism with a rolling horizon length T and planning horizon length $\hat{T}$. During each time period t order quantities are determined considering the opening inventory level and demand over a t to $t+T$ window. . . . 65

5.2 Illustration of demand samples. We present three distinct demand samples, which are random realizations of the demand during time $t = 1$ to $T = 10$. . . . 66

5.3 Illustration of an ordering mechanism with the IECA. . . . 68

5.4 Illustration of changes in inventory quantities for the different forecast samples as depicted in Figure 5.2. . . . 69

5.5 Illustration of cost as a function of order quantities for 100 randomly generated samples of length 10, mean 5, inventory holding cost 1 and backorder cost 5. . . . 71

5.6 Steps of the OH-IECA. For different values of Ω, the optimal order quantity and the optimal cost can be determined using (5.31) and (5.15) respectively. Then the order quantity with the minimum immediate period cost (5.17) is chosen. . . . 74

5.7 Illustration of distribution of percentage excess cost in 100,000 random problem instances. The average error is 0.227%. . . . 76

5.8 Illustration of performances of DP and the OH-IECA for different uncertainty levels. . . . 79

6.1 Figure demonstrating different possible coverage periods depending upon the order time. The rolling horizon length is 8. At the beginning of $t = 1$, coverage period can be chosen between $t = 1, 2, ..., 8$. Similarly, at the beginning of $t = 2$, the coverage period can be chosen between $t = 2, 3, ..., 8$ and so on. . . 90

6.2 The supply chain network under study. . . . 92

6.3 An example of solution using the CSS approach. . . . 94

6.4 An example of solution using the DSS approach. . . . 95

6.5 Comparative performances of OH-IECA and RDP approaches for the SRP with non-stationary demand. The average optimality gap for OH-IECA is 7.3% and that of SDP is 3.2%. . . . 101

7.1 Supply chain network structure under study. . . . 107

7.2 Illustration of the nature of inventory computation. In Case 1, the order quantity q_i for an item i is greater than its forecast F_{kiz}. It results in excess inventory $q_i - F_{kiz}$ as indicated in the shaded porting. This is because, in such cases 100% demand can be fulfilled. However, if the order quantity is less than or equal to the forecast quantity then, the excess inventory is zero as a maximum demand equal to the order quantity can be fulfilled. On the contrary, in Case 2, the shortfall from the forecast indicates the lost sales even if it fulfills the service level constraint. . . . 111

7.3 Illustration of ABC classification. X-axis represents items, and Y-axis represents cumulative volume after the items are sorted in descending order of sales volume. The interval from the origin to point A on the X-axis represents high volume items constituting 50% volume, i.e. class-A items. The interval from point A to point B on the X-axis represents the next high volume items constituting from top 50% to 80% volume, i.e. class-B items. All the remaining items belong to class-C. 112

7.4 Flowchart illustrating the proposed methodology. Inputs are forecast and packaging constraints. ABC-VED classification is used to define minimum service level. The problem is formulated as a mixed integer linear programming (MILP) model, and an ϵ-constrained like method is used to generate Pareto optimal solutions. 113

7.5 Flowchart illustrating the method used to evaluate the quality of solutions. We compare each solution with the corresponding real sales. If the order quantity obtained from the solution is higher than the real sales then it results in leftover inventory and if it is lower than the real sales then it results in lost sales. 114

7.6 Performance of the models. The circles indicate the results from the deterministic approach. The triangles represent the results from the probabilistic approach. Point A indicates performance of the solution currently used by the firm. Point B indicates the ideal solution if the exact demand is known in advance. Results of top 10 retailers presented in Table 7.6 correspond to points x (Prob. Appr.) and y (Det. Appr.). 116

7.7 Representation of a solution of problem M-1 with M prepacks as a vector of size M. 119

7.8 Neighborhood representation with $\mathcal{V} = \{-1\}$. 121

7.9 Neighborhood representation with $\mathcal{V} = \{-1, +1\}$. 121

7.10 Effect of initial solution on final quality on instances A-O. 127

7.11 Effect of initial solution on final quality on instances P-X. 127

7.12 Convergence curves for problem instances A and X. 131

7.13 Histogram of $\overline{\Delta}_k$ for different retailers of the case study 133

1 Illustration of demand samples. We present three distinct demand samples, which are random realizations of the demand during time $t = 1$ to $T = 10$. 156

2 Illustration of cost components when we choose $\zeta_z^* = \lfloor \zeta_z \rfloor$, $\zeta_z^* = \lceil \zeta_z \rceil$ or $\zeta_z^* = \zeta_z$. 157

List of Tables

2.1 State-of-the-art and position of our work. 27
2.2 Proposed notations for different sub-field types used for the classification scheme. 31
2.3 Examples of classification using the proposed scheme. 32

3.1 Inventory classification as per IQR 43
3.2 Summary of KPIs 44

5.1 A comparative representation of application conditions of various replenishment planning methods in single-item inventory systems and their qualitative performance indicators. NS(s, S) Policy: Non-stationary (s, S) policy with policy parameters determined by a heuristic, DP: Dynamic Programming, Offline Time: Computation time for policy parameters, Online Time: Computation time for actual orders from the policy or the method. 60
5.2 Applicability and qualitative performance of the proposed coverage period cost (CPC) approach. 63
5.3 Parameters and variables for the optimization models. 66
5.4 Problem instances for performance comparison. 75
5.5 Comparative performance of different approaches for different problem instances for horizon cost criterion. (Ratio of C^h_π to C^h_{min}). 77
5.6 Comparative performance of different approaches for the problem instances for infinite horizon cost criterion. (Ratio of C^∞_π to C^∞_{min}). 77
5.7 Runtime for different approaches (in Seconds). 78
5.8 Average values of $\frac{C^\infty_\pi}{C^\infty_{min}}$ at different variance levels. σ = Forecast standard deviation. $\hat{\sigma}$ = Realized standard deviation. Detail results are given in Table 2 in Appendix C. 78
5.9 Comparative performance of the OH-IECA for the benchmark problems in Veinott Jr and Wagner (1965). C^* is the optimal cost for the problem instance. $C^\infty_{OH-IECA}$ is the infinite expected per period cost using the OH-IECA. CoV indicates the coefficient of variation for that cost. $\overline{\Delta C(\%)}$ is the average excess cost of using the OH-IECA over the optimal cost. min.$\Delta C(\%)$ is the minimum cost difference during the simulations and max.$\Delta C(\%)$ is the maximum cost difference obtained during the simulations. 79

5.10 Comparative performance of the optimal policy and the OH-IECA on the benchmark problems given in Veinott Jr and Wagner (1965) for different realizations of the standard deviation. σ = Forecast standard deviation and $\hat{\sigma}$ = Realized standard deviation. The percentage excess cost of the OH-IECA over the cost obtained with the optimal policy is presented. 80

5.11 Problem instances with non-stationary demand. 80

5.12 Percentage optimality gap of OH-IECA with respect to dynamic programming in case of non-stationary demand. 81

5.13 Comparative performance of the proposed heuristic for the benchmark problems in the presence of different batch sizes (Percentage excess cost over dynamic programming). 82

5.14 Average performance of the proposed heuristic for the benchmark problems at different realizations of the standard deviation and at different value of robustness factor r. (Percentage excess cost over optimal (s, S) policy) for the benchmark instances. 83

6.1 Parameters and variables for the optimization models. 88

6.2 Different MCPC computations. 90

6.3 Notations for the parameters and variables. 93

6.4 Problem instances for the SRP with non-stationary demand. The table presents the expected demands for 12 periods. For each of the instances P1 to P11, the inventory holding cost $H = \{0.1, 0.5, 1\}$, backorder cost $W = \{5, 1020\}$ and fixed order cost $K = \{0, 25, 50\}$. 98

6.5 Problem instances for the SRPSS. 99

6.6 Percentage optimality gap of RDP with respect to dynamic programming in case of non-stationary demand. 100

6.7 Comparison between CSS and DSS approaches for the SRPSS using DP. . . . 101

6.8 Comparison between CSS and DSS approaches for the SRPSS using RDP. . . 102

7.1 Parameters and variables for the optimization model. 109

7.2 Minimum service level thresholds for different classes of items. In this example, 95% of minimum service level is required for Class-A and Vital items. Differentiated service levels are used for other classes. 112

7.3 Supply chain components. 114

7.4 Problem size and computational time 114

7.5 Service level thresholds for different class of items. 114

7.6 Comparison of solution performances. 115

7.7 Instance sizes and parameters for numerical analysis. N is the number of items, M is the number of prepacks, $\tilde{R}$ is the number of item-prepack combinations, and **SL** is the vector of service levels. The forecast values **F** and pack contents **R**, and **SL** are chosen randomly from the given uniform distributions. 124

7.8 Performances of heuristics used for generating the initial solutions (GH1 and GH2). $\overline{\Delta}$ is the mean percentage gap between the initial solution and the optimal solution. t represents the runtime in seconds. 126

7.9 Effect of initial solution on the quality of final solution. $\overline{\Delta}$ is the mean percentage gap between the initial solution and the optimal solution. t represents the runtime in seconds. σ is the standard deviation of Δ for 20 iterations. The $p-$value is calculated using one way ANOVA between the observations obtained using IGSH-GH1 and IGSH-GH2. 128
7.10 Effect of perturbation size θ and number of perturbations γ. Problem instances A-O. The statistically best results are indicated in bold font. 129
7.11 Effect of perturbation size θ and number of perturbations γ. Problem instance P-X. The statistically best results are indicated in bold font. 130
7.12 Overall performance of different perturbation parameters across problem instances A-P. The best results are indicated in bold font. 130
7.13 Overall performance of different perturbation parameters across problem instances P-X. The best results are indicated in bold font. 131
7.14 Parameters for the case study. 132
7.15 retailer specific results of the case study. 133
7.16 Overall results of the case study. 134

8.1 Possible extensions and industrialization status. The three problems: single-item replenishment planning (SRP), single-item replenishment planning with supplier selection (SRPSS) and the promotional joint replenishment problem (PJRP) are under industrialization. The "Base model" refers to the problem defined in this dissertation without any extension. 138
2 Comparative performance of different approaches for the problem instances when the realized standard deviation is higher than predicted. σ = Forecast standard deviation. $\hat{\sigma}$ = Realized standard deviation. It presents the ratio of C_π^∞ to the minimum C^∞ among the approaches. 160

List of Abbreviations

Abbreviation	Detail
AR	Augmented Reality
ASL	Alpha Service Level
AVL	Availability
B2B	Business to Business
B2C	Business to Consumer
BSL	Beta Service Level
CDC	Central Distribution Centre
CoC	Cluster of Commonality
CoG	Cost of Good
CPC	Coverage Period Cost
CSS	Common Supplier Selection
DoS	Days of Supply
DP	Dynamic Programming
DSS	Dynamic Supplier Selection
ECA	Expected Cost Approach
EH	Enumerative Heuristic
EOQ	Economic Order Quantity
ERP	Enterprise Resource Planning
GDP	Gross Domestic Product
GSH	Guided Search Heurisitc
GSM	Guaranteed Service Model
HS	Hybrid Serice
IEC	Immediate Expected Cost
IECA	Immediated Expected Cost Approach
IGSH	Iterated Guided Search Heuristic
ILS	Iterated Local Search
IMS	Inventory Management System
IoT	Internet of Things
IP	Inventory Position
IPA	Infinitesimal Perturbation Analysis
IQR	Inventory Quality Ratio
IT	Information Technology
ITO	Inventory Turnover
IVL	Inventory Level
IVV	Ineventory Variance
JRP	Joint Replenishment Problem
KPI	Key Performance Indicator
MCPC	Minimum Coverage Period Cost
MDKP	Multi-dimensional Knapsack Problem

Abbreviation	Detail
MILP	Mixed Integer Linear Problem
MIS	Management Information System
MOQ	Minimum Order Quantity
MPNV	Multi-Product Newsvendor
MSSP	Multi-Stage Stochstic Program
OH	Optimized Heuristic
ORV	Order variance
OT	Operational Technology
PJRP	Promotional Joint Replenishment Probelm
RCA	Robust Cost Approach
RDC	Regional Distribution Centre
RDP	Reduced-state Dynamic Program
ReSa	Reduced Sampling
RFID	Radio Frequency Identification
ROI	Return on Investment
RPA	Robotic Process Automation
S&OP	Sales and Operations Planning
SA	Stochastic Approximation
SAA	Sample Average Approximation
SaaS	Software as a Service
SCM	Supply Chain Management
SCN	Supply Chain Network
SCOR	Supply Chain Operation Reference
SCP	Supply Chain Profitability
SCPMS	Supply Chain Performance Management System
SJRP	Stochstic Joint Replenishment Problem
SRP	Single-item Replenishment Problem
SRPSS	Single-item Replenishment Problem with Supplier Selection
SSM	Stochastic Service Model
TIV	Total Inventory Value
VCP	Value Chain Profitability
VED	Vital Essential Desired
VR	Virtual Reality

General Introduction of the Thesis

This doctoral dissertation is a research project under an industrial agreement of training through research (CIFRE) between Vekia SAS, Lille and the University of Lille. At the university the research was carried out with the ORKAD research team of the OPTIMA thematic group in the CRIStAL laboratory.

Vekia is a start up in Lille of around 40 people, founded in 2008. It was built on the idea of using Artificial Intelligence (AI) to automatically and optimally manage the supply chain of large firms and achieve “the right stock in the right place at the right time”. The importance of having an efficient supply chain makes the products of Vekia an important response to a critical problem. This is true for organizations as well as in a global context (ecology, transport-related savings). Historically, major customers of Vekia were from the retail industries (textiles, DIY stores, etc.). Today, Vekia aims to extend into other sectors such as pharmaceutical and spare parts distribution. For all of its customers, Vekia offers SaaS (software as a service) solutions, which make it possible to anticipate the needs at the various entities of the Supply Chain. The major function of the solutions is to suggest optimized order quantities at the downstream entities to the upstream entities. However, the solutions also include numerous alerts, monitoring or piloting tools which assist the users in better decision making. Vekia’s value proposition has several aspects such as: a promise of superior quality of replenishment, AI solutions with a wide variety of data inputs, simplified management and automation of repetitive tasks. Ultimately, Vekia aims to operate the most complex supply chains almost independently.

ORKAD is a research team from the OPTIMA thematic group of the CRIStAL laboratory (Univ. Lille, CNRS, Centrale Lille - UMR 9189). The main goal of the ORKAD team is to simultaneously exploit combinatorial optimization and knowledge extraction to solve optimization problems. Although the two scientific fields have developed more or less independently, the synergy between combinatorial optimization and knowledge extraction offers an opportunity to improve the performance and autonomy of optimization methods thanks to knowledge and, on the other hand, to efficiently solve the problems of knowledge extraction thanks to operational research methods. Approaches adopted are mainly based on combinatorial mono and multi-objective optimization. The ORKAD team carries out its work both academically, as well as in cooperation with hospitals in the region and companies.

In this dissertation, we mainly address inventory control problems from practical point of view. Managing inventory is one of the major activities in supply chains. It has a significant share in the direct costs and it also affects the costs of other activities such as logistics and facility planning. The prime objective of inventory management is to match the demand and supply in a cost effective way. However, there are three key challenges in achieving this. Firstly, we encounter numerous types of inventory optimization problems in practice. Developing new solutions for each one of them is practically not feasible. Secondly, demand, lead time, supply and information records are some of the major sources of uncertainty.

This can lead to excess inventory costs and lower service levels. The third challenge is the complexity of resulting optimization problems. With uncertain parameters, inventory optimization problems usually take the form of multi-stage stochastic optimization problems. They are often intractable.

The main decisions taken by the inventory manager are, when to order?, how much to order? and from where to order? Vekia's cloud-based SaaS solutions suggest optimized decisions in real time to the inventory managers. These solutions must have four properties: to provide best quality solutions, to consume less computational time, to be able to handle large volume of data and to be readily reconfigurable. Therefore, the objective of this thesis can be succinctly put as *"to develop a fast inventory optimization methods that are able to accommodate multiple sources of uncertainty, scalable and extendable to include new problem parameters"*. We study the single-item and multi-item replenishment problems under real-world constraints. Some of the limitations of existing methodologies are: assumptions regarding the underlying distribution of the uncertain parameter, inefficiency in dealing with non-stationary demand, and not being flexible enough to include additional problem specifics. In addition, in the multi-item aspect, we also focus on the promotional ordering problem which has not been studied in the presence of multi-item packings. In this dissertation, we prove that such problems are $\mathcal{NP}$-hard. Apart from developing optimization methods for replenishment problems, we also analyze the performance assessment aspect of real inventory management systems. Based on the above problem statements we list the objectives of this dissertation as follows.

1. To propose a classification scheme for inventory optimization problems.
2. To identify a set of key performance indicators and propose a performance assessment method for inventory management systems.
3. To propose a flexible inventory management framework that encompasses various inventory optimization problems.
4. To propose methodologies for single-item replenishment problems that can be readily industrialized and extended to include additional problem specifics.
5. To examine the impact of multiple suppliers scenario on the decisions of single-item inventory systems.
6. To propose a methodology for the promotional multi-item replenishment planning problem.
7. To test the proposed methodologies on real-world datasets and propose extensions for industrialization.

The above objectives are individually addressed in different chapters. We explain the arrangement of the thesis and contents of each chapter next.

Arrangement of the Dissertation

This dissertation consists of two parts. The first part has three chapters that explain some preliminary concepts, state-of-the-art, propose classification scheme and performance evaluation method. The second part has the remaining five chapters where we propose the solution methods for various problems. The contents of this dissertation are summarized in Figure 1. Chapter 1 gives a general context of supply chains and inventory optimization problems. We also analyze their impact in the real-world. We provide a general overview of Vekia, and its vision and mission. We also discuss the problem statement and research methodology.

An overview of literature on replenishment planning, inventory optimization problem classification and performance assessment is provided in Chapter 2. We determine the state-of-the-art literature and identify the gaps. Then in the later part of Chapter 2, we propose a classification scheme for inventory optimization problems. Our work on classification is built on de Kok *et al.* (2018) and is inspired from Graham's notation (Graham *et al.*, 1979). Additionally, we give due consideration to replenishment planning and problem parameters.

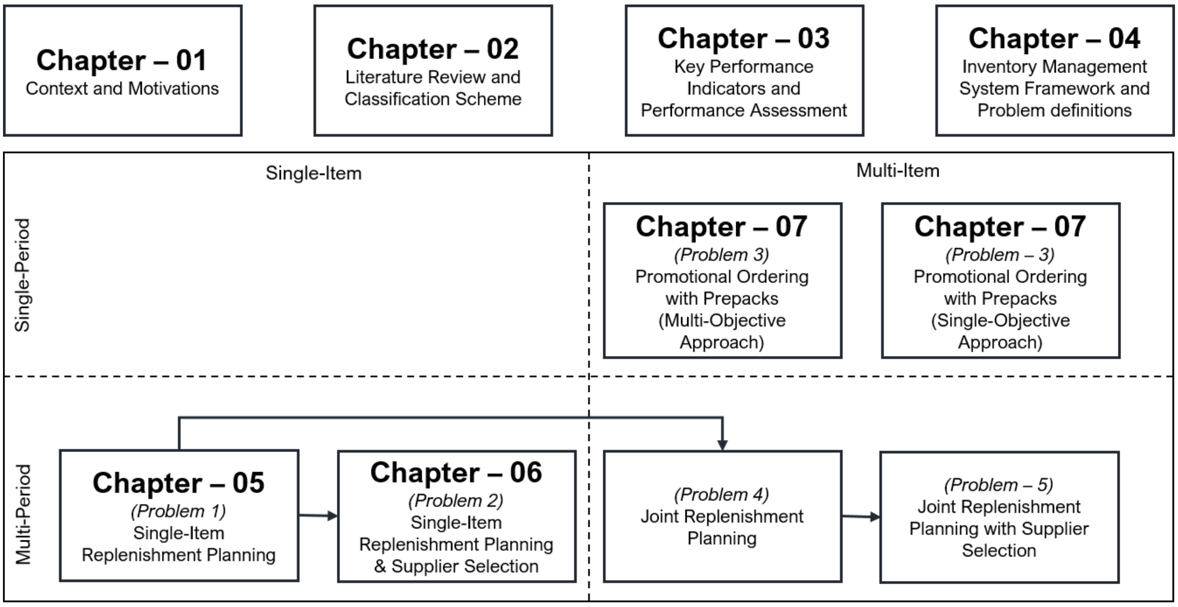


Figure 1: Arrangement of the dissertation.

Chapter 3 lists various key performance indicators (KPIs) to measure the performance of inventory management systems effectively. Those KPIs are derived from various surveys. We also identify challenges while measuring performances of real-world systems and propose a simulation methodology address those. Our proposed performance assessment methodology is novel which enables comparison with existing inventory management systems as well.

In Chapter 4, we propose a modular inventory management framework with each module addressing a particular class of problems. Then, we define the inventory optimization problems addressed in this dissertation in detail. We list their variants/ extensions as well.

Chapter 5 addresses the single-item replenishment problem and we build on Levi *et al.* (2007a), Levi *et al.* (2006), Levi *et al.* (2007b) and Özen *et al.* (2012). We propose a novel sampling-based approach to address the problem. The model is based on sample average approximation of the real cost function. We consider a periodic review inventory policy. We also propose an improved heuristic for faster computation. The proposed model is extended to include batch size constraint.

Chapter 6 considers the extensions to the problem considered in Chapter 5. We propose a dynamic programming approach with fewer states to generate approximated better quality solutions in case of non-stationary demand. Then, we address the supplier selection problem. We analyze two approaches: common supplier selection and dynamic supplier selection, and compare their cost performance. We also develop approximate methods for integrated supplier selection. Parts of this work (Sahu *et al.*, 2020b) has been presented in the international conference on operations research and enterprise systems (ICORES) 2020.

Chapter 7 addresses the replenishment problems with multiple items. In particular, we focus on the promotional ordering problem with multiple items and packings containing multiple items. We first propose a multi-objective approach with lost sale and excess inventory as the two competing objectives. This work (Sahu *et al.*, 2018) has been presented in the international conference on information systems, logistics and supply chain (ILS), 2018. Then, for practical purpose we also propose a single objective approach. In the presence of prepacks the resulting optimization problem is $\mathcal{NP}$-hard. We propose a metaheuristic to obtain near optimal solutions in reasonable time.

Finally in Chapter 8, we provide some industrial extensions to the discussed problems. We conclude and identify some future research areas.

Part I

Preliminaries, State-of-the-Art, Classification and Evaluation

Chapter 1

Context and Motivations

Abstract: *Supply chains have a profound impact on the global economy. With surveys suggesting almost \$460 billion worth of goods moving each year in France only, supply chains also significantly affect the people and the environment. We chose to study the replenishment planning problem in supply chains in this dissertation. This chapter provides a general context of the research and explains the motivations behind selecting the problem. First, we elaborate on the importance of supply chains and the challenges faced during general supply chain planning. Then, we introduce Vekia which funded this research and present its mission and objectives. In the aspect of supply chain planning, we analyze three decision levels: strategic, tactical and operational and classify various decisions accordingly. We emphasize the replenishment planning problems and explain their associated challenges. We also discuss the limitations of existing solution methods. At the end, we explain the general organization of this dissertation and outline our contributions.*

1.1. Introduction

The global economy relies heavily on consumer goods of all kinds: food, textiles, electronics, equipment, automobiles, etc. The supply chain is the business function that drives the physical flow of these finished consumer goods or their components. It comprises two major classes of activities, i.e., demand planning and supply planning. In addition, it includes sub-activities like logistics (warehouses, transportation and packaging, etc.), finance (billing and payment), and information technology (stock status, ordering and tracking, etc.). Its efficient functioning has a direct impact on the finances of the company, as well as globally on performance of the economy, on the environment (transport, storage and destruction of expired stock, etc.) and on the working conditions of the personnel. In France, the supply chain represents 1.8 million direct jobs, and contributes 10% of the gross domestic product (GDP) (MTESFrance, 2019).

As mentioned previously, demand planning and supply planning are the two major activities enabling supply chain functioning. Demand planning deals with prediction of future demands, which would be used in the later phases of supply chain planning. For most accurate forecasts, it uses advanced analytical approaches including machine learning and artificial intelligence to examine historical sales data, current demand signals, customer orders, shipments, macro-economic data as well as other exogenous data such as weather and market indicators. After the predictions are made, alignment of that demand becomes necessary across different functional area. This is accomplished by a managerial process called sales and operation planning (S&OP). Finally, supply planning deals with decisions regarding optimal production quantity, orders, inventory levels and replenishment quantities. Efficient demand and supply planning involves optimization of associated cost parameters and due consideration of practical constraints.

Given the importance of this sector and its cost implication, many private and public actors have carried out work aimed at maximizing their operational efficiency. Some existing challenges are:

1. The development of archiving and data processing systems, which allow the storage of billions of unit of transactions (Big Data), stock positions, orders, sales, etc., now provides astronomical amounts of data that should be used.
2. The increased competitive pressure on companies requires increased efficiency, in particular by limiting the excess inventory and avoiding stock-outs. It is a very complex subject in the context of variable demand that is known in advance only in an uncertain way.
3. The complexities of the supply chain as a result of constantly increasing number of product references (today several hundreds of thousands or even millions), storage sites (several hundred to several tens of thousands) and variability demand (commercial operations in distribution, powerful and fast mode effects), etc.
4. The extension of the supply chain: upstream supply has been globalized for about 15 years and today we are witnessing a globalization of downstream supply to consumers with players like Decathlon or Amazon that cover many countries.

As a consequence, companies with a supply chain business model have an immediate need to improve their performance in this context. The main objectives are

1. Inventory management considering uncertain future demand and practical constraints such as, transport capacity constraints, preparation or manufacturing costs, potential turnover, etc.

2. Computation of production quantities or order quantities in a very limited computation time (from less than a second to a few hours in case of less frequent planning), on a very large number of items.

3. The ability to explain these recommendations to their users.

1.2. About Vekia

Vekia[1] provides cloud-based software as a service (SaaS) forecasting and replenishment solutions to various clients that work in the supply chain domain. Those clients comprise different industry verticals such as retail, pharmaceutical, telecommunication, manufacturing and spare parts service. The primary objective stems from the desire to automate supply chain planning such that the decisions are optimized.

1.2.1. Solutions

Solutions provided by Vekia are broadly of two categories: demand forecasting and replenishment planning. The demand forecasting solutions produce forecasts using historical and

[1]https://www.vekia.fr

exogenous data. The replenishment planning solutions use those forecasts to determine the optimal inventory levels across different nodes (locations) in the supply chain.

Classical supply chain planning consists of three stages, forecasting→ planning→ operations, each having their own objectives. These stages use different sets of algorithms to find the best local solutions (i.e. the optimized solution for the respective stages). The forecasting stage mostly uses statistical or machine learning techniques, the planning stage utilizes optimization techniques and the operation stage evaluates the practical viability of the proposed solutions, and then determines a feasible solution. For practical purposes, only a planning decision has tangible effects on a supply chain. However, stages in classical planning do not consider optimal conditions and practical feasibility together, giving rise to solutions which are locally optimal but globally might not be so.

The objective of the forecasting stage is to produce the best possible estimates of future demand. Traditionally, time series techniques were used to generate forecasts. However, more recently, machine-learning algorithms have replaced these practices with increased accuracy. Forecasts are made at different granularities (product group, time, etc.), and horizons based on needs. They are usually single valued (mean or median forecast) and optionally coming with simple uncertainty measures. Choosing these values essentially amounts to a form of premature decision making, which triggers some information loss when compared to the full underlying probabilities of each possible sale.

Inventory optimization comes next. Depending on the entity (store, warehouse, etc.), an ideal inventory level and ordering frequency are determined based on the forecast. The maximum order quantities are also decided. These define the optimal inventory policy. Yet practical constraints such as packaging, transportation capacity, etc. are not considered at this stage, which may lead to infeasible decisions at the operations stage. Additionally, solution providers usually rely on deterministic forecasting or a simple parametric estimate, as it is computationally simple. This method does not provide optimal solutions due to uncertainty associated with future estimates. In the third stage, practical replenishment decisions are made considering all the practical constraints. The outputs of this stage may contradict the solutions obtained in the previous stages, and therefore can lead to sub-optimality.

1.2.2. Vision

Although the above three stages are practically sound to implement, their inherent sub-optimality must be addressed. The primary objective of any supply chain is to maximize profitability. In the context of inventory optimization, profitability can be improved by increasing the availability of products at the least possible cost. The classical planning perspective that follows forecasting→ planning→ operations has different objectives at different stages. While forecasting focuses on minimizing the deterministic forecast error, the planning stage focuses on maximizing service level or minimizing cost or maximizing profit and the operations stage focuses on practical viability. Separation of the planning and operation stage stems from the complexity of the supply chain and the results can be detrimental to the global profit. Those complexities are discussed below.

Uncertainty: Having a perfect demand forecast is impossible. We may use the most sophisticated machine learning and artificial intelligence algorithms to estimate the future demand, but it still might be different when put into action. The deterministic forecast can be the mean, median or any suitable estimation of the actual demand. However, this value, along with lead time, product returns, future events and recorded values, is uncertain. This makes the classical inventory optimization inherently sub-optimal. The classical approach relies on safety stock to deal with any uncertainty. However, use of safety stock is not always practical as it significantly increases inventory and is a reactive approach.

Practical constraints: Any decision made at the operations stage must satisfy a set of physical and managerial constraints. While some constraints reflect the physical limitations of various supply chain entities like finite capacity of a warehouse or a store or a transportation vehicle, others can arise because of economies of scale and ease of practicality. For example, lot sizes are determined based on the economic viability of the products and by the ease of transportation of products packed in pallets. Purchase orders may also be subject to the MOQs (minimum order quantities) of the supplier. Managerial constraints can include the budget allocation to different departments that limit the purchase order. Sometimes it can be an arbitrary quantity. Temporal constraints, like lead-time and product life expectation, can also affect a decision. However, all of these constraints are usually not considered in the planning stage. This translates either into two independent optimizations or sometimes a completely manual operations stage. In the classical approach, some constraints are incorporated in the planning stage and the practical constraints are incorporated in the operation stages and the forecasting stage does not consider any of these constraints.

Multiple objectives: As supply chains evolve, so do their objectives. Today, supply chains do not have one single objective. While they must generate maximum profit, they should also minimize the environmental impact and maintain an adequate service level, among other things. This transforms the classical single objective optimization problem into a multi-objective one. An optimal value of one objective may lead to an unsatisfactory value of another one.

Forecasting has been given the prime importance in supply chain operations for very long time. However, recent focus on decision and demand driven planning has shifted some of the focus to the real-world decisions. There are various reasons for this shift in focus. The first one can be attributed to the limitations of forecasting methods. Even the best of the forecasting methods cannot be 100% accurate. Performance improvement by higher forecast accuracy is limited by the inventory optimization method. Secondly, any supply chain planning procedure must consider practical constraints so that the plan generated by it is actually feasible. An infeasible plan is of no use.

Depending on the client, the concerned supply chain network can be single or multi-echelon and the flow patterns could also be different. After generating the forecast using advanced machine learning and artificial intelligence techniques, the relevant replenishment quantity is suggested. Hence, the economic importance of the project is tremendous, as any research should not only be academically novel, but also practically applicable. Although, profit remains the prime objective of the supply chain business, nowadays, multiple objectives are simultaneously targeted. This defines the multi-objective requirement. Alongside, the ability to handle large amounts of data is also there.

In this thesis, we develop optimization methods to meet some of the needs stated above by taking into account both the complexity of the problem and the uncertainties on the available information. This dissertation has the following objectives:

1. To propose several models to formalize different inventory optimization problems encountered at Vekia.
2. To propose resolution methods for the problems formalized above.
3. To propose an evaluation framework for replenishment planning systems.
4. To develop an adaptive industrialization framework.

1.3. Context of Supply Chain Planning

Before delving into the details of the concerned problem, we provide some context of the process of supply chain planning. As any inventory optimization problem falls into a broader class of supply chain planning, it must be understood from the strategic point of view. In this section we discuss general supply chain management and its decision levels.

1.3.1. Supply Chain Management

The supply chain drives physical goods from the supplier (or the point of origin) to the end consumer. Those goods pass through a system of suppliers, manufacturing facilities, storage facilities, transportation modes and retailers. Depending upon the industry involved and the domain area, it can either be responsible for transforming raw materials into finished goods, passing them to the end consumer or simply distribute the finished goods.

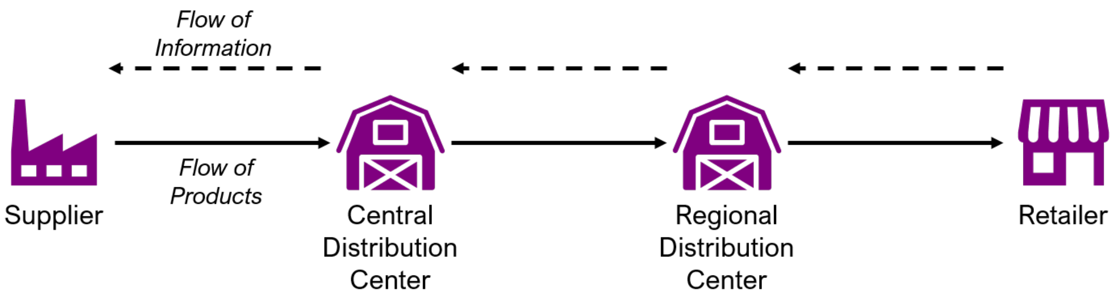


Figure 1.1: A general structure of current supply chains.

The term supply chain network (SCN) is often used to describe the structure of most of the supply chains (Chopra *et al.*, 2013). As previously mentioned, a supply chain consists of entities such as suppliers, manufacturers, warehouses and retailers, etc. These, along with the pattern of flow of physical goods among them, constitute the supply chain network. In a hierarchical setting, each *level* in the supply chain is called an *echelon*. For example, in Figure 1.1, the supplier, the central distribution center, the regional distribution center and the retailer, each constitutes an echelon. A supply chain network also consists of the links that join the facilities together to bring a product from one echelon of the supply chain to the other (Firoozi, 2018).

In addition to the flow of physical goods (products), information also passes from one echelon to another echelon as presented in Figure 1.1. Mapping both information and product flows gives a comprehensive picture of the supply chain network. Physical goods can be in the form of raw materials, semi-finished products or finished products. Similarly, information flow consists of the demand from the end customer to preceding entities and the inventory related information.

Even though this general structure of supply chains is widely adopted, in the future, with the development of new technologies they will look different (as illustrated in Figure 1.2). We expect supply chains of the future to be more connected and enabling centralized decision making. Development of Internet of Things (IoT), barcode, Radio frequency identification (RFID) and additional collaborative planning options is also enabling efficient information sharing.

Supply chain management (SCM) is an effective method to integrate both information and material flows seamlessly across the supply chain. Simchi-Levi *et al.* (2000) define supply chain management as *"the integration of key business processes among a network of interdependent suppliers, manufacturers, distribution centers, and retailers in order to improve the flow of goods, services, and information from original suppliers to final customers, with the objectives of reducing system-wide costs while maintaining required service levels"*. Supply chain management integrates the strategies across suppliers, manufacturers, warehouses, and retailers. It aims to make physical goods available in the right quantities, at the right locations, at the right time at a minimum related costs. Supply chain management is the global approach that is involved with different types of decisions across different entities.

The results of an innovative supply chain management strategy are impressive. Amazon is a great example of an efficient supply chain management on its multi-echelon supply chain network. The revenue of Amazon has reached almost \$136 billion in 2016. In fact, Amazon is the fastest company to reach \$100 billion in sales revenue, taking only 20 years. The combination of sophisticated information technology, an extensive network of warehouses and excellent transportation makes Amazon's supply chain one of the most efficient among in the world (Leblanc, 2017).

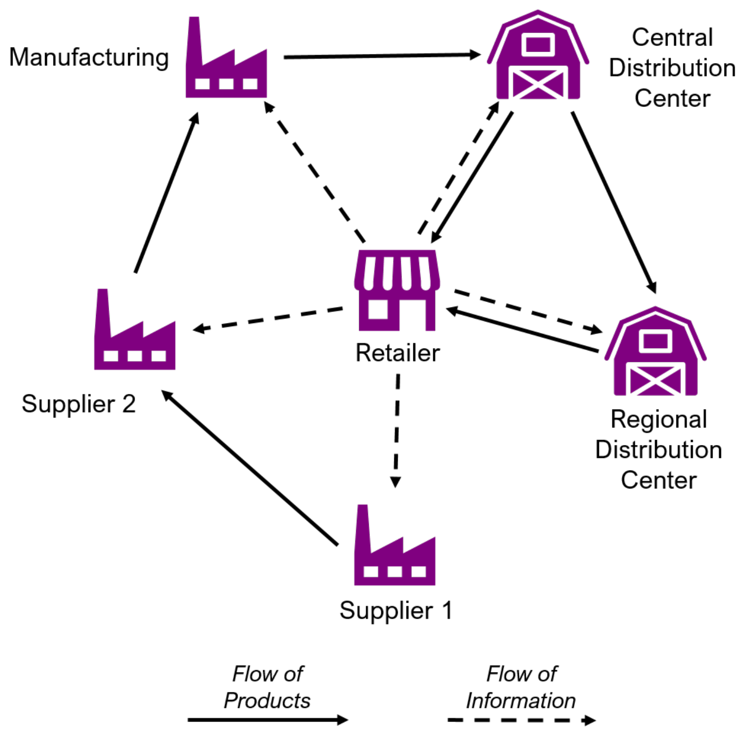


Figure 1.2: A general structure of future supply chains.

1.3.2. Decision Levels

Decisions in supply chain management fall under three categories: strategic decisions, tactical decisions, and operational decisions. In Figure 1.3, various planning decisions are mapped according to the above three categories (Firoozi, 2018). Those decision levels are discussed below in detail.

Strategic Decisions

Strategic decisions are usually taken considering the long term (usually considering a few years) objectives. For example, the decisions regarding what the supply chain's configuration will be, how resources will be allocated, and what processes each stage will perform cannot be taken on a short notice. This is because, those decisions involve physical changes in the supply chain and usually have very high cost implications. Thus, while making these decisions, companies must take into account uncertainty in anticipated market conditions over the next few years. Facility location and supplier selection are some of the examples.

Tactical Decisions

After strategic decisions, tactical decisions are taken considering medium term objectives. Supply chain functions involving market selection, market-location mapping, and inventory policy selection fall under this category. Tactical decisions have a major impact on the operational decisions of the next level. They can be changed on short notice and the financial implications involved with such changes are generally minimal, and sometimes, can be beneficial. However, process change and change management are some of the potential challenges involved with such alteration.

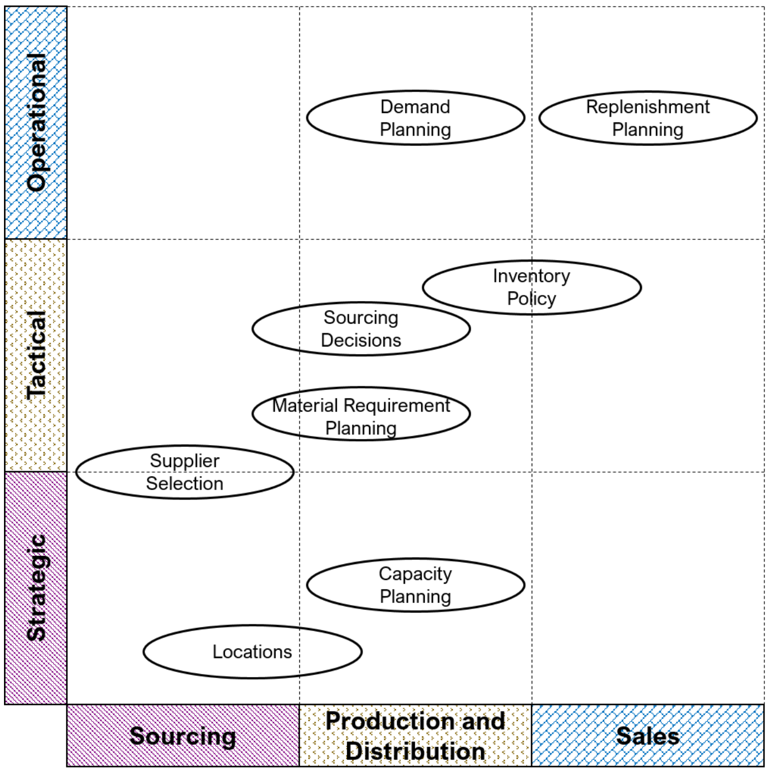


Figure 1.3: Supply chain planning: Decisions across different levels and functional areas.

Operational Decisions

Operational supply chain decisions are short-term, having a time horizon of a week or day. Customer order fulfillment, daily flow management, pick list generation at a warehouse or delivery schedule setting for trucks are some of the examples.

Inventory management is an important problem both at tactical and operational levels. At a tactical level it deals with the problem of positioning inventories across different locations and at different times. At an operational level, however, the problem is more of a replenishment planning type which decides the shipment quantity and in some cases also decides the suppliers. In the upcoming section, we discuss the overall inventory management problem in the context of this dissertation.

1.3.3. Inventory Management

Inventory is kept at various locations of the supply chain to serve as a decoupling point. It can be in the form of raw materials, work in process or finished goods (Silver *et al.*, 1998). Inventory and related operations contribute a significant portion of supply chain cost. The *inventory policies* followed by an organization also decide the effectiveness and responsiveness of its supply chain.

The most challenging matter in inventory management is finding the fundamental trade-off between responsiveness and efficiency while making inventory decisions. An increase in inventory levels generally increases the responsiveness of the supply chain to the customer. It also facilitates a reduction in production and transportation costs per unit, because of improved economies of scale. However, this also increases the inventory holding costs (Chopra *et al.*, 2013).

On a tactical level, the following medium-term inventory related decisions are taken: material requirement planning, sourcing and inventory policy selection, etc. On an operational level, the replenishment quantities for different entities are decided based on the upper level tactical decisions and practical constraints. Supplier selection is sometimes considered in operational level as well. This aspect is more prominent in case of retail supply chains.

Inventory optimization is formally defined as the method of balancing capital investment constraints or objectives and service level goals while taking demand and supply constraints into account. It has been studied extensively over the last century. The earliest model for the economic order quantity (EOQ) was proposed in 1913 by Harris (1913). It is a deterministic model, that finds some applications even today. Harris (1913) considered constant demand rate throughout the year, inventory holding cost and fixed order cost. Also, no shortage was allowed. With the evolution of supply chains, the complexity has also increased. This makes these simpler models less cost effective. In addition, the enormous variety of problems needed separate and sophisticated mathematical formulations and superior optimization methods. We discuss the different problem in detail in upcoming chapters.

Application Sectors

In the context of Vekia, inventory management problems are different depending on the sector of application: such as retail, consumer goods, spare parts, home service, automobile, telecommunication, etc. Even though fundamentally inventory management strives to minimize costs or maximize profits, depending upon the application, the specifics of the problem can change. This has led to development of highly customized mathematical models and optimization algorithms.

Retail industries manage one of the most complex supply chains, primarily due to the sheer number of items they sell. They operate n-echelon supply chains where, n is 2 to 3 in most cases. This indicates the presence of regional distribution centers (RDCs) and central distribution centers (CDCs) respectively. A representative multi-echelon supply chain is depicted in Figure 1.4. Different varieties of items bring separate sets of challenges. Normally, retailers have a larger number of suppliers, and with more suppliers, inventory optimization becomes more difficult. Additionally, most retailers handle items with less shelf life, which need different inventory optimization approaches.

Consumer goods supply chains differ from retail supply chains by number of echelons and involvement of manufacturing facilities. Most consumer good supply chains are B2B, whereas in retail, it is B2C.

In case of spare parts supply chains, service level is usually more important than cost. In case of internal spare parts supply chains, the cost implications of not meeting the desired service level are even higher. Furthermore, most of the unfulfilled demand is backordered.

Automobile and telecommunication industries usually involve bulky items with lower quantities. For example, in the case of telecommunication industries, fiber optic cable is one product which can be ordered as different rolls of predefined cable length. However, those cable lengths can also be customized as per requirement. In such a case, even if there is no concept of unfulfilled demand, any improper inventory management can lead to overstock at certain locations and project delay at others.

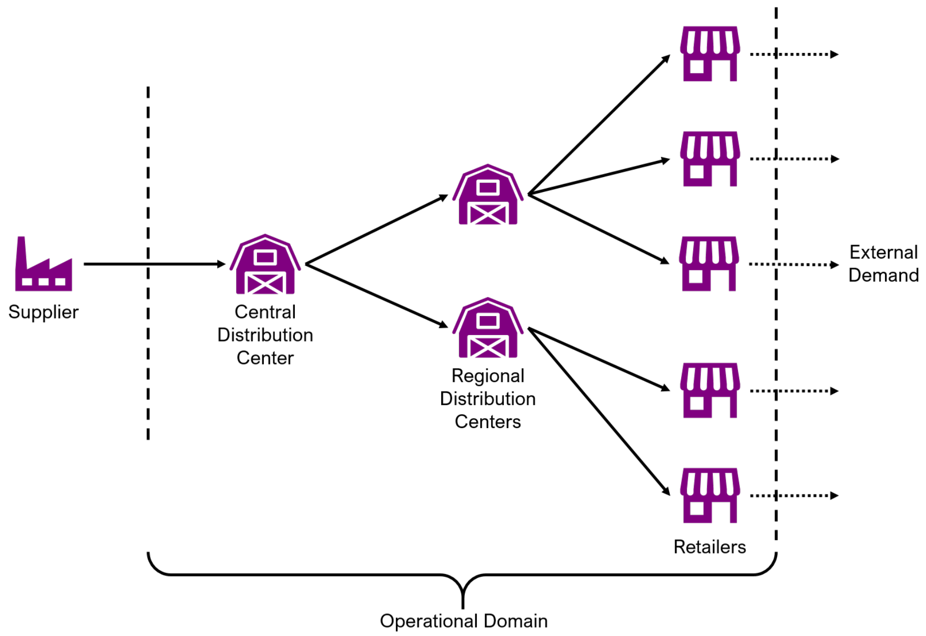


Figure 1.4: A 3-echelon retail supply chain.

1.4. Motivations

Our motivations for choosing inventory management as our research area come from the following: its impacts on supply chain profitability, diversity of such problems and challenges in solving them, limitations of existing methods and the futuristic view of inventory management.

1.4.1. Impacts of Inventory Management

Most businesses understand the need to maximize their working capital. Beyond funding growth and reducing reliance on debt or other forms of external financing, increasing cash availability can help organizations strengthen their balance sheet and enhance operational performance. Yet a gap exists for many businesses between recognizing the imperative for working capital optimization and understanding what steps to take to improve liquidity. A company must balance the needs of the customer with the goals of the company in an integrated inventory management model where the right inventories are in the right place, at the right time. What is needed, is a disciplined process whereby the level of investment in inventory is in line with the expected level of customer service to be provided. To maintain optimal inventory levels, organizations need robust systems to accurately track and maintain control of inventory levels. Internal processes are required to manage vendors and customers, the supply chain, and to maintain control of the inventory. Such processes will enable organizations to track inventory performance, monitor demand patterns, maintain accurate inventory measurements, and ensure suppliers adhere to their commitments. These efforts must be underpinned by the adoption of a cash management culture. When it comes to inventory, this means managers must prevent buyers from over-purchasing for fear of losing a sale. Instead, they must put disciplined processes in place to ensure orders are based on real demand, enabling the company to maintain minimal inventory levels without compromising customer service.

Replenishment planning have two basic objectives: minimize variable and operational costs, and maximize service level. Supply chains incur real costs in purchasing products, storing them, transporting them and disposing them. Moreover, indirect costs are paid for losing sales or not maintaining some minimum stock thresholds. Depending on the sector of application (and the problem type), the objective of a replenishment planning problem can be achieving service level with minimum cost or minimizing overall cost or maximizing profitability or achieving equilibrium while fulfilling operational constraints. Moreover, replenishment planning can also have multiple objectives. Environmental considerations such as carbon emission, energy consumption, and capacity utilization are usually considered as additional objective (s).

Replenishment planning problems are either solved as profit maximization or cost minimization problems. Costs are either in form of real expenses or penalties. Real expenses include purchase costs, inventory holding costs and transportation costs etc. Penalties include backorder penalty, lostsale penalty or penalties for not fulfilling some constraints.

1.4.2. Problem Diversity and Challenges

Inventory management has a vast variety of optimization problems. The specifics depend on the industry, problem type, availability of information, IT infrastructure of the organization, etc.

Different industries encounter different inventory optimization problems. For example, retail industries have different sets of constraints and objectives than that of spare parts industry. Shelf-lives of different products across retailers are also different. While a supermarket has food items with the shortest shelf-lives, the electronics goods or kitchen appliances have some of the longest shelf-lives. Retail fashion industries have a lot of new products coming in at short intervals as well as seasonality of existing products.

The same industry might have *different inventory optimization problems*. For example, a retailer can have regular multi-period inventory optimization problems as well as a single-period inventory optimization problem during promotions. A supplier supplying multiple products can pose a joint replenishment problem (JRP) in addition to single product replenishment problems for other suppliers.

Information availability is another factor affecting adoption of inventory optimization models. Probabilistic consideration of demand is a more cost optimal approach while optimizing inventory. However, if the actual distribution is not known then, either an approximation is used or the single-valued mean demand is considered for planning. Also, even if the probabilistic demand information is available, it is less useful without an effective optimization method that can use such information.

Finally, *IT infrastructure* of an organization also limits what kind of inventory management system can be adopted. However, with recent developments in cloud services and SaaS solutions, this constraint is becoming less relevant. Even though some concerns remain about the storage of data and its ownership by the SaaS provider or the organization itself.

1.4.3. Limitations of Existing Methods

Inventory optimization problems have been studied quite extensively in the literature. But, existing optimization methods are customized for the specific problems. General frameworks in this context are rare or even non-existent. We have previously outlined the variety of inventory optimization problems depending upon the application sector. Furthermore, a general optimization method for all of these problems is computationally expensive. This leads to the need for new methods/ frameworks that are more generalized and computationally less expensive.

Apart from specificity, most methods described in the literature do not reflect practical concerns faced by users and the constraints about information generation. For example, presence of multiple batch sizes (pack sizes), non-parametric demand distribution, mix of lost sale and backorder or inaccuracy in inventory record are usually not taken into account. Other practical limitations include lower scalability and inability to being integrated into the existing system or into a centralized system.

Inventory optimization problems are generally considered hard to solve. In the case of deterministic demand, lot-sizing problems are $\mathcal{NP}$-hard for single item. For stochastic demand, it is often intractable for greater number of periods due to the curse of dimensionality.

Practical constraints are drastically different from those considered in the literature. Hence, very few works have been successfully adopted in practice. Our aim is to develop methods that are not only mathematically rigorous but also practically useful.

1.4.4. Supply Chain Planning in Future

The success of any inventory optimization method is highly dependent on how the inventory is managed from the business point of view. The current state of inventory management in most organizations is decentralized. Moreover, there are three distinct levels of decision making: strategic, tactical and operational as mentioned earlier. However, surveys conclude some sweeping changes to supply chain planning (Deloitte, 2019) in the future. Those changes include *no planning* and *continuous re-planning* scenarios as explained in upcoming paragraphs.

Nowadays, most supply chain plans are not capable of reflecting reality very well. With fast changing scenarios, most plans become outdated at the very moment they are prepared. With new information they are either infeasible or too inaccurate, and therefore, are ignored by many stakeholders. Hence there is a real need to ask whether to plan at all. The difference between strategic and operational plan was described earlier. In the future organizations are expected to move towards more demand driven planning where they would just use the forecast to determine target resource levels and then wait for the actual demand to arise. In such a case, supply chain planning will be replaced by configuration, where the main planning activity is to set up our supply system such that it adapts well to the actual demand. On the contrary, it is also possible to get rid of discrete planning events completely. It will be replaced by plans that are generated every second in the response to new information as it materializes.

1.5. Problem Statement and Research Methodology

In this section, we present the broad research question and outline our research methodology. Previously, we have listed the challenges associated with managing inventory nowadays and also how the inventory management process would evolve in the future. The main challenge we address in this dissertation is developing solution methods for "the sheer variety of inventory optimization problems under uncertainty". Traditionally, customized methods have been developed for different problems. They lack flexibility while dealing with different problems. Moreover, some common shortcomings include assumptions regarding parametric distributions for uncertain parameters and limited number of constraints, etc.

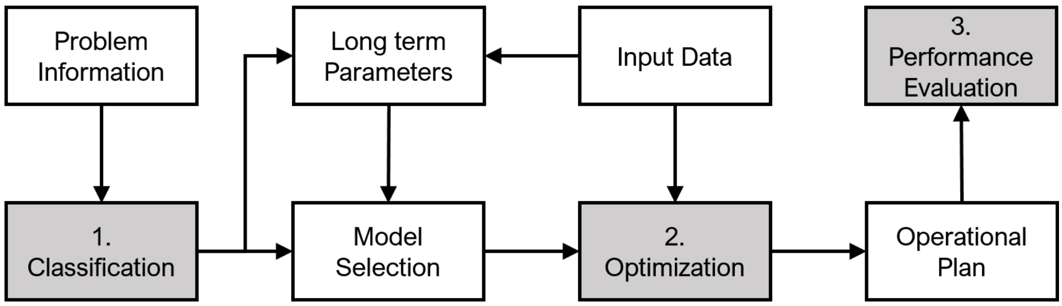


Figure 1.5: Industrial implementation plan.

Our methodology can be summarized as follows. For the large variety of problems, we first propose a classification scheme and a modular inventory management system (IMS) framework. Each module in the framework addresses a particular class of problems, and it is flexible to incorporate changes in problem specifics. We propose sampling-based methods that do not assume any form of distributions or dependencies to address uncertainty. Our initial analysis has been conducted for uncertain demand. However, the proposed methods can readily incorporate uncertainty in lead time, delivered quantities, inventory record and item quality, etc.

We divide the inventory optimization problems into two broad groups: long term and short term problems. The long-term group covers most strategic planning problems and some of the tactical ones. For example, multi-echelon strategic stock placement and supplier selection fall under this category. Under the short-term category day-to-day operational replenishment plans are covered. We focus mainly on problems faced by retail industries. The problems are again divided into two parts: single-item replenishment planning (SRP) problems and multi-item (joint) replenishment planning (JRP) problems. In the case of single-item, two independent problems are identified. The first one deals with the replenishment decisions where there is only one supplier, i.e. the decisions are indifferent to the purchase cost (Chapter 5). The second one deals with the case where multiple suppliers are present and the unit price for the item and supply constraints are different (Chapter 6). Multi-item problems are those in which the cost function for a single item cannot be independently determined for each item. We first analyze the problem of promotional ordering (Chapter 7), which is analogous to single period newsvendor problem but, with some supply constraints. In addition, we propose a classification method for inventory optimization problems (Chapter 2) and also develop a performance evaluation framework (Chapter 3). Various industrial extensions are proposed in Chapter 8.

The industrial implementation plan (see Figure 1.5) consists of three action points. We first classify the replenishment problem using available information. Then, a suitable operational model is chosen as per the problem class. The operational model takes available data and upper level long term planning parameters as inputs.

1.6. Conclusions

In this chapter, we have provided a general context of this thesis. First, we have given an overview of Vekia. It provides cloud based SaaS solutions to supply chain industries. Its vision regarding inventory management is to achieve more demand-driven supply chain planning and automating the process. Then, we have presented a general description of supply chains and supply chain planning process. We have also discussed the three levels of planning: strategic, tactical and operational. This dissertation principally addresses inventory optimization problems. We have introduced inventory management briefly and explained how it varies across industries in this chapter. Our motivations behind selecting the problem were classified under impact, problem diversity and existing limitations. We have also given a brief perspective into supply chain planning in the future. At last, we have presented the problem statement and the research methodology.

In the next chapter, we shall present an overview of literature in the field of inventory optimization, classification of inventory optimization problems and performance assessment. We also propose a classification scheme for inventory optimization problems.

Chapter 2

State-of-the-Art and Classification of Inventory Optimization Problems

Abstract: *The first chapter gave a brief outline of the research. It summarized the context, problem statement and research methodology of this thesis. In this chapter, we aim to provide an elaborate review of literature pertaining to replenishment planning, inventory optimization problem classification and performance assessment. At the end of the literature review, we also analyze the literature. As mentioned in the last chapter, in practice, we encounter various types inventory optimization problems. Industrial customization for each specific problem is difficult. Therefore, we also aim to develop a classification scheme for those problems, and hence, making it possible to incorporate common solution methods. This classification scheme is used throughout the dissertation to present the different problems under study.*

2.1. Introduction

Given the importance of inventory management for supply chains, we aim to provide a practical framework and its associated optimization methods in this dissertation. In order to develop a common approach, different inventory optimization problems need to be classified into groups and a standard nomenclature has to be developed for a formal identification of each problem.

Inventory optimization has been studied extensively over the last century. Various models have been proposed considering a diverse variety of inputs and application areas. As of now, a generalized solution method which can be practically implemented does not exist. Development of such an approach will first require the detailed definition of the problems that it is going to solve. Since the exact definitions are specific to the concerned problems, their number would explode while incorporating every real-world problem. This gives rise to the need of a standard classification scheme for the existing inventory optimization problems.

The rest of this chapter is arranged as follows. We first provide an elaborate review of state-of-the-art methods for the studied problems in Section 2.2. Then, in Section 2.3, we propose a classification method for inventory optimization problems along with the motivations behind the work and some examples. Our contributions in this regard are the following.

1. We propose a classification scheme for inventory optimization problems.
2. We provide a brief literature review of inventory optimization in general.

2.2. State-of-the-Art: Inventory Optimization

We address inventory optimization problems in a very broad sense. Our contributions are in these three areas: classification of inventory optimization problems, performance assessment of inventory management systems and single echelon inventory optimization. A summary of reviewed literature areas is presented in Figure 2.1. Each of these problems is reviewed from the following perspectives.

1. How well the approach addresses the real-world challenges?
2. What is its accuracy and how much runtime does it require?
3. Has it been tested using real data?

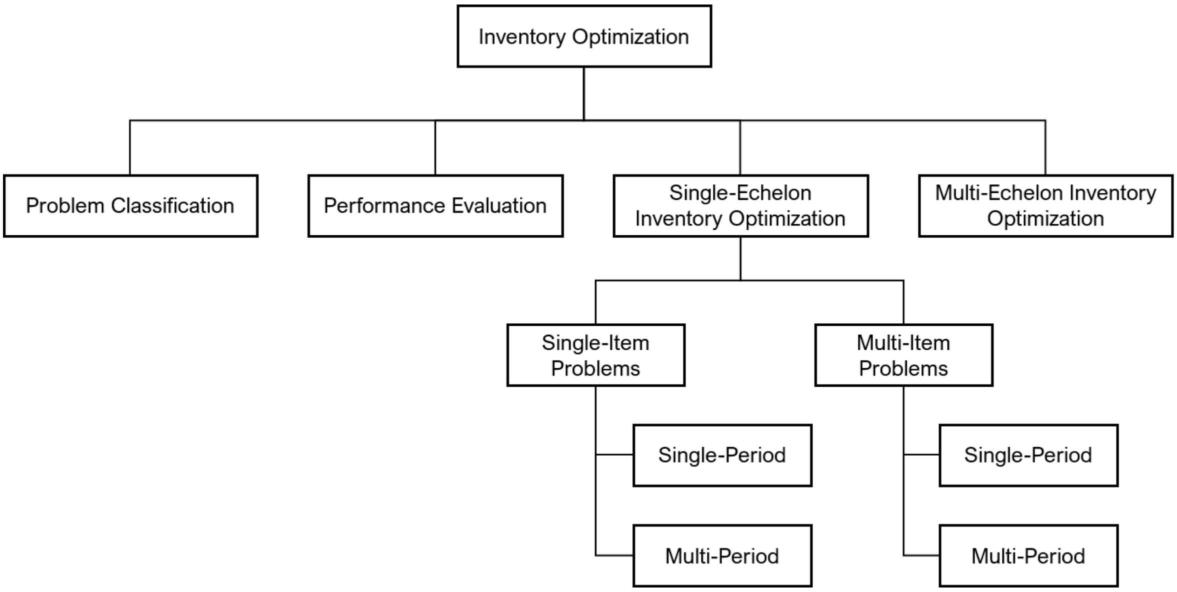


Figure 2.1: Summary of reviewed literature.

The upcoming sections are structured in the following way. First, in Section 2.2.1, we provide a general introduction to inventory optimization problems, their importance and their current state of application in industries. Then, we review the different classification methodology for inventory optimization models and their applicability in Section 2.2.2. Performance evaluation methods are discussed next in Section 2.2.3. Afterwards, a major part of the review, Section 2.2.4, is dedicated to single echelon inventory optimization problems as we address mostly issues pertaining to this in this dissertation. We also provide a brief review of multi-echelon inventory optimization problems in Section 2.2.5. Analysis of this literature is given in Section 2.2.6.

2.2.1. General Inventory Optimization

The costs related to inventory management amount to 8-12% of total sales (CoresightResearch, 2019). Effective inventory management can reduce the inventory levels by up to 25%. Because of such a huge economic impact, inventory optimization has been studied over the last century and more rigorously so over the last five decades. This section presents a preliminary discussion about the problems addressed in this dissertation and inventory optimization in general.

The primary objective of inventory optimization is to minimize the total inventory related cost. However, with recent changes in business trends, the inventory optimization process is viewed from service and environmental perspectives as well. There are three basic cost components: inventory holding cost, shortage cost and fixed order cost. Apart from these, additional costs can be considered depending upon the specific problem. As the name suggests, the inventory holding cost is the cost incurred due to storage of unsold products. The rate may vary and it is largely decided by the product, location and the business. The yearly inventory holding cost is usually in the range of 7-16% of procurement costs (Brown, 2011). Shortage costs are of two types: lost sale cost and backorder cost. Lost sale refers to the situation when the unfulfilled demand is lost completely, i.e., it does not appear as additional demand in the future. On the contrary, backorder refers to the case when the unfulfilled demand is not lost but appears as additional demand in the future. If the unfulfilled demand is fully backordered, then it appears as the additional demand of the next period. The actual cost incurred during a shortage is due to the loss to brand value or the compensation given to the customer or both. The fixed order cost is paid each time an order is placed. Usually, it is independent of the order volume and represents the administrative and/ or order preparation cost.

The cost components above are applicable in most inventory optimization problems, be it single-echelon or multi-echelon. Single-echelon problems allow to consider each location independently. Whereas, in case of multi-echelon problems, interdependency between different locations makes the problems very difficult.

From a mathematical point of view, inventory optimization problems are either deterministic or stochastic. Deterministic models consider the available information to be perfect. Difficulty in these models may arise due to their combinatorial nature (if any) but, they are relatively simple compared to their stochastic counterparts. When demand is constant and deterministic, a static economic order quantity (EOQ) model (Harris, 1913) can be used. An order of quantity equal to EOQ is placed periodically. For time varying deterministic demand, lot-sizing models are used.

In practical situations, however, deterministic information is seldom available. Information related to demand, lead time, item quality, etc. is always uncertain. Hence, stochastic models have gained more importance in recent years as they model more realistic situations. While deterministic models can be viewed as solving a single situation, stochastic models give a more comprehensive optimization considering several possible scenarios. The major sources of uncertainty in a supply chain include demand, lead time, inventory record, received product quality, etc. Even if stochastic inventory models are more realistic, their cost functions are very difficult to evaluate analytically (Zheng, 1992). Therefore, instead of elegant dynamic mathematical models, researchers focused on determining optimal *policy* structures (Scarf, 1959; Clark and Scarf, 1960; Karlin, 1960; Veinott, 1966). The most common is the base stock policy. It can have many forms depending upon problem parameters, such as (s, S), (R, S), etc. Computation methods for such policy parameters are given in Zheng and Federgruen (1991) and Feng and Xiao (2000) for stationary demand. Similarly, for the non-stationary (s, S) policy Askin (1981) and Bollapragada and Morton (1999) provide simple heuristics. More recently, Xiang *et al.* (2018) provide a linear programming approach for non-stationary demand.

Real-world inventory optimization problems have many parameters. Different configurations of those parameters yield different types of problems and, therefore, require different solution methods. Developing customized solutions for such a huge variety of problems is not practical. Hence, method can be developed for a class of problems. Literature in this area is scarce. Those who have addressed it, rarely take solution development into account specifically.

Performance assessment has not received considerable attention in academic literature but, it is one of the fields that has profound impact on decision making during real-world implementing an inventory management system (IMS). This affects the implementation of decisions taken by the IMS as well. Performance assessment consists of two major activities.

1. Selecting the suitable key performance indicators (KPIs) and
2. Defining the computation procedure of those KPIs.

Different businesses require different sets of KPIs for their effective performance evaluation. For example, service level does not have equal importance in retail and pharmaceutical businesses. Inventory turnover also falls in the same category. Once the suitable KPIs are decided, the procedure to compute them must be specified. One of the major challenges faced by retailers while adopting a new IMS is how to compare the existing system to the new one.

There are three categories of performance metrics typically involved in evaluating an inventory system: financial, operational, and service (Petropoulos *et al.*, 2019). Costs and profitability related metrics come under the financial category. Operational metrics consist of inventory quantity, inventory turnover, variance, etc. Finally, service level, fill rate and availability come under the service category.

2.2.2. Classification of Inventory Optimization Problems

Classification of inventory optimization problems can help structuring the problems and it can also influence industrial solution development. However, we found limited literature in this area. In this section, we discuss those articles, their adopted methodology and applicability.

Notation and classification of supply chain related problems are addressed in Gayraud *et al.* (2015) and de Kok *et al.* (2018). The analysis of Gayraud *et al.* (2015) is confined to network design problems. They propose an approach analogous to Graham's notation (Graham *et al.*, 1979) for scheduling problems with three fields $\alpha|\beta|\gamma$. The α, β and γ fields express the general network structures, management rules and performance criteria respectively. Each of these fields has sub-fields. Since, they only consider network design problems, the actual structure of the network becomes a decision variable and therefore, not included in the notation. They only consider one source of uncertainty and do not consider any inventory policy. Also, they do not consider service quality aspects of inventory optimization such as service level, lost sale or backorder.

de Kok *et al.* (2018) focus on multi-echelon inventory optimization only. Their review is more elaborate. They also propose a notation scheme and have more fields compared to Gayraud *et al.* (2015). They consider different network types except a few (separation of general type networks and logistic networks). However, they do not consider number of products, time period or supplier selection aspects. These are necessary to solve an inventory optimization problem. Paterson *et al.* (2011) provide a review of inventory models with transshipment. They also provide a short classification method for such problems.

2.2.3. Performance Assessment of IMSs

Performance assessment is as important as designing the inventory management/ optimization system. It consists of two major activities (see Section 2.2.1). In the context of inventory management, selecting the wrong KPIs can result in choosing the wrong IMS. Even after choosing the right KPIs, in practice, their exact computation faces several challenges. In this section, we present the work already done in these areas.

Shepherd and Günter (2010), Estampe *et al.* (2013) and Kurien and Qureshi (2011) are some of the recent studies on supply chain performance measurement. Estampe *et al.* (2013) presented a characterization of various supply chain evaluation models. They provided a table comprising various models organized by their origin, the type of analysis used, relevant conditions and constraints, the degree of conceptualization and the indicators being proposed. Supply chain operation reference (SCOR) (APICS) planning practices have seen wide adoption in recent days. The linkage between SCOR practices and supply chain performance is studied by Lockamy III and McCormack (2004). More prominently, a supply chain performance metric framework is proposed by Gunasekaran *et al.* (2004). They divided supply chain activities into four main stages: planning, sourcing, manufacturing and delivery. Our study is focused on planning and delivery aspects. Cuthbertson and Piotrowicz (2008) discussed the key dimensions of supply chain performance: sustainability, efficiency and effectiveness, responsiveness, and flexibility. All these dimensions can also be used in the context of inventory management performance. However, efficiency and effectiveness, and responsiveness are, in our view most relevant. Adivar *et al.* (2019) proposed additional seven perspectives: customers, operations, sourcing, finance, information sharing and information technology, transportation, environment for performance assessment in retail supply chains. They also propose some next generation KPIs in the context of retail supply chains.

Shepherd and Günter (2010), Gopal and Thakkar (2012) and Maestrini *et al.* (2017) provide a comprehensive review about supply chain performance management systems (SCPMS). Stangl and Thonemann (2017) analyze some equivalent inventory metrics and how they affect the decision making process of managers. Protopappa-Sieke and Seifert (2010) study the relationships between financial and operational metrics of inventory control.

2.2.4. Single-Echelon Inventory Optimization

Inventory optimization problems and practices vary greatly depending upon the concerned supply chain network. Literature focuses on determining the optimized replenishment quantity and time for single-echelon supply chains. On the other hand, for multi-echelon networks, literature focuses on finding the optimal policy structures or "installation stock" quantities. In this section, we discuss various works on single-echelon inventory optimization.

Single-echelon inventory optimization problems can be further divided into two categories: problems with only one item (product) and joint replenishment problems (problems with multiple items).

Single-Item Inventory Optimization

Inventory optimization problems involving only one item have been studied extensively. In this dissertation, we focus on problems with stochastic demand only. In practice, to make globally optimized replenishment decisions, planners consider various factors, such as number of possible orders during the planning horizon, order batch-size, transportation modes and costs, transportation capacities, discounts, suppliers, etc. Based on all these parameters, two broadly distinct streams of research emerge: single-period models and multi-period models.

As the name suggests, the single-period models have a planning horizon length equal to one period, i.e., only one order is placed for the whole planning horizon. The most basic single-period stochastic model is the classical newsvendor problem (Qin *et al.*, 2011) and the majority of single-period inventory models are its extensions.

Promotional events and new product launches are some of the real-world instances where application of the newsvendor problem is evident. Qin *et al.* (2011) provide an extensive review of newsvendor extensions in the context of supplier costs, buyer risk profile and modeling customer demand. The cases of supplier discount and capacity constraint are addressed in Mohammadivojdan and Geunes (2018). Retailers frequently encounter situations where the suppliers offer the products in fixed batch sizes. More commonly, one product can have multiple batch sizes. While a single batch-size alone does not pose any major challenge as the cost function is convex, problems with multiple constraints or multiple batch-sizes are more complex. Batch-size related problems are better explored in multi-period problems.

The earliest models of multi-period stochastic inventory optimization consider demand to be stationary. It has been studied since the 1950s. The earliest literature in the field are Scarf (1959); Clark and Scarf (1960); Karlin (1960); Veinott (1966). Veinott (1966) study the structure of optimal policy. Interestingly, the structure of the optimal policy is very simple, i.e., of (s, S) type. The policy defines two inventory levels: the reorder level s, and the order up to level S. When the inventory level reaches or goes below the reorder level an order is placed to raise the inventory position to S. The optimal parameters can be found by using policy iteration, but it is time consuming. Simpler methods to compute the stationary (s, S) policy are given by Zheng and Federgruen (1991) and Feng and Xiao (2000). Similarly for the non-stationary (s, S) policy Askin (1981) and Bollapragada and Morton (1999) provide simple heuristics. More recently Xiang *et al.* (2018) proposed a mixed integer linear programming approach to compute the approximate non-stationary (s, S) policy. The heuristic provided therein, although computationally less expensive than dynamic programming, is still not suitable for real-world size problems.

Both the cases of stationary and non-stationary demand rely on the complete knowledge of the underlying probability distribution. Although the non-stationary case is more appealing for practical applications, the methodologies developed in this area can only provide myopic policies at a reasonable computation time. Other practical limitations arise due to the presence of batch-sizes. Widely used policies in the presence of batch-size, such as (R, Q), (R, nQ) and (s, nQ) (Q = Batch size, s, R = Reorder points) are themselves not optimal.

Multi-period inventory models with finite horizon length have also been studied extensively. For a finite planning horizon $\hat{T}$, replenishment and inventory decisions can be taken for time period $t \in \{1, 2, ..., \hat{T}\}$ (Refer Figure 2). However, due to the well known curses of dimensionality (Defourny *et al.*, 2012), multi-stage decision making models for $\hat{T} \geq 3$ are extremely hard to solve. On the contrary, solutions for smaller $\hat{T}$ is too myopic to be implemented in a real-world setting Rahdar *et al.* (2018). Rahdar *et al.* (2018) proposed a tri-level model for multi-stage inventory optimization problem considering a rolling horizon for a single product at a single installation with both demand and lead time as source of uncertainty. Also, they do not consider capacity constraints. In practice however, installations of infinite capacity do not exist.

To summarize, under practical application conditions, the solution methods for the SRP problems face the following challenges: application to non-stationary demand, consideration of practical constraints and the curses of dimensionality. The proposition of Xiang *et al.* (2018) for non-stationary demand requires complete knowledge of demand distribution and it is applicable for planning horizons which are short. Any update in the demand information requires re-evaluation. The static dynamic uncertainty strategies proposed in Rossi *et al.* (2015) and Özen *et al.* (2012) to address the problem can only perform faster computation. They still require the complete demand distribution and lack the flexibility to incorporate batch-size and other constraints. To alleviate the curses of dimensionality, Bertsimas and Thiele (2006) propose a robust optimization approach. It does not require the complete demand distribution and works on a suitable uncertainty budget. However, it yields more cost if the demand information is accurate. The approach can be cumbersome to apply in lost sale scenarios and for longer planning horizon. The approach provided in Rahdar *et al.* (2018) also alleviates the curses of dimensionality but, still they require complete demand information and do not consider practical constraints. Levi *et al.* (2007a) and Neely and Huang (2010) propose sampling-based (distribution free) approaches which are useful for practical applications. However, they still rely on dynamic programming and the approach is computationally expensive.

Integrated inventory optimization supplier selection models has been gaining significance recently. A comprehensive review has been given by Yao and Minner (2017). The recent work of Firouz *et al.* (2017) addressed a supplier selection in a supply chain network with multiple warehouses and multiple suppliers. Replenishment mode also included proactive transshipments. The suppliers varied by price, capacity, quality, and disruption characteristics. They assume stationary, parametric demand. Cheaitou and Van Delft (2013) studied a dual sourcing model where the suppliers were differentiated by their lead times. They provided theoretical bounds and heuristic approximations for the optimal policy in such situations. On the other hand, Fox *et al.* (2006) analyzed a dual sourcing model where the suppliers vary by their fixed and variable costs only. They proved that, under such situations and for log-concave demand density a reduced form of (s, S) type inventory policy is optimal. Zhang *et al.* (2012) relaxed some of the constraints of Fox *et al.* (2006) and partially characterized the optimal inventory policy. They also considered simultaneous sourcing from multiple suppliers for the same item. One of the major shortcomings we found is that, neither of the batch-size or minimum order quantity constraint is considered. Similarly, authors did not take into the service level or the notion of robust solution which is important in case of inaccurate forecasts.

Multi-Item Inventory Optimization

Multi-item inventory optimization are also frequently encountered in practice. Moreover, almost all real-world inventory optimization problems consist of multiple items and the items are assumed to be independent if the interaction is negligible.

The multi-item newsvendor problem has been studied extensively with extensions that are fit for application purposes. Customer demand, supplier pricing policy, and buyer risk profile are some of the key areas where extensions have been made. We will discuss various extensions of the classical newsvendor model. Most of literature assume a parametric distribution of demand and focus on obtaining a closed form expression of the optimal ordering quantity. The multi-item newsvendor problem without any constraint can be solved independently for each item, but with constraints it becomes hard to solve. Problem formulation also depends on the application area. Multi-item newsvendor problem with a single constraint was first considered by Hadley and Whitin (1963). It was extended by Lau and Lau (1995) and Lau and Lau (1996) to address general demand distributions and multiple constraints. These are also applicable when the demand uncertainty can be described by discrete or interval scenarios. Multi-item newsvendor models have been studied in the context of budget constraint (Vairaktarakis, 2000), capacity constraint (Erlebacher, 2000), budget constraint and fixed ordering costs (Moon and Silver, 2000), budget constraint and supplier discount (Zhang, 2010) initial stock of units (Silver and Moon, 2001), random yield scenarios (Abdel-Malek *et al.*, 2008), price discount (Jucker and Rosenblatt, 1985), supplier quantity discounts (Khouja, 1995), (Khouja and Mehrez, 1996), and price elasticity (Khouja, 2000). To our knowledge, only Taleizadeh *et al.* (2008) consider multi-item newsvendor problem with packaging constraints, albeit for most of retail supply chains packaging consideration is prevalent. Recently, Dong (2018) studies the multi-item newsvendor problem with a global budget constraint and propose a robust optimization approach. Chernonog and Goldberg (2018) consider bounded demand distributions. While the assumption regarding parametric distribution of demand is common in the literature, in practical situations, the distribution is not truly known.

Multi-period inventory optimization problems with multiple items is called the joint replenishment problem (JRP) (Balintfy, 1964). With stochastic demand it is called stochastic JRP (SJRP). The periodic JRP is strongly $\mathcal{NP}$-hard for infinite horizon and $\mathcal{NP}$-hard for finite horizon (Cohen-Hillel and Yedidsion, 2018). No such proof is available yet for the SJRP. However, it is commonly regarded as very hard due to its huge state space. Structural results for SJRP show that it does not follow a simple policy even for two item case (Ignall, 1969). For stationary demand Viswanathan (1997), Johansen and Melchiors (2003), Özkaya *et al.* (2006) and Feng *et al.* (2015) propose inventory policies. However, those policies do not consider batch-size or lost sales. Moreover, renewed policy evaluation is required with the evolution of demand. Very recently Yang and Kim (2018) proposed an adaptive JRP policy for non-stationary demand for single buyer.

2.2.5. Multi-Echelon Inventory Optimization

In the multi-echelon setting, inventory control models are studied for both deterministic and stochastic input parameters. Some of the earliest research published in this area are Clark and Scarf (1960) and Simpson Jr (1958). These two papers establish two different streams of stochastic inventory control for multi-echelon supply chains: the stochastic service model (SSM) and the guaranteed service model (GSM). Both approaches solve the problem of safety stock allocation and inventory policy parameters.

Stochastic Service Model

In the SSM approach, the system deals with all demand conditions having no theoretical upper limit. For any demand condition, if sufficient inventory is available at the upstream stage, they are immediately delivered. However, for any shortfall, the downstream stage has to wait for the unavailable items leading to stochastic delay. Echelon order-up-to (s, S) type policies are proven to be optimal in n-echelon serial systems in Clark and Scarf (1960). This holds true for constant lead times, discrete time, and with or without fixed ordering cost. Under stochastic lead times, order-up-to policy is proven optimal by Muharremoglu and Tsitsiklis (2008). For capacitated n-echelon serial systems with identical capacity limits Janakiraman and Muckstadt (2009) derive the optimal policy structure. For n-echelon convergent (assembly type) systems with constant delay and no fixed order costs, Rosling (1989) proves optimality of echelon order-up-to policies. For n-echelon divergent (distribution type) networks Diks and De Kok (1998) prove its optimality. However, for divergent systems the assumption of free transshipment is essential, which was provided by Eppen (1981).

That being said, optimal policy for n-echelon general type networks has not been found yet. Only optimal policies for special structures have been identified by Nadar *et al.* (2014) and Benjaafar *et al.* (2011) for two-echelon systems. All of the above literature are applicable for systems with infinite capacity. However, serial or infinite capacity systems do not exist in practice. Finite capacity adds considerable complexities to multi-echelon inventory systems. As of now there are no analytical results that enable the calculation of optimal policies for finite capacity multi-echelon systems. Nevertheless, for such scenario, Glasserman and Tayur (1994) and Glasserman and Tayur (1995) use Infinitesimal Perturbation Analysis (IPA) to compute the gradient of the cost function as a function of base stock levels. The reason for the complexity of n-echelon inventory analysis is the interdependency between echelon inventory positions and echelon stocks of different items at different points in time. For serial systems, the mutual dependencies can be expressed as recursive equations that link inventory positions across echelons. They are generalized into distribution type systems. However, under finite capacity the inventory position of an item at a point in time depends on the echelon stock of its predecessors over multiple points in time. Hence, we lose the Markov property (de Kok *et al.*, 2018).

Guaranteed Service Models

The GSM approach assumes a bounded demand. The inventory model with GSM was originally proposed by Simpson Jr (1958), but the GSM notion was formally coined by Graves and Willems (2000). The bounded demand assumption enables calculation of optimal base stock policy even in multi-echelon setting. However, finding the right upper bound can be tedious in some practical scenarios. This limits the GSM. Nevertheless, due to their tractability, GSM models have been studied extensively post 2000.

The GSM model deals with strategic safety stock placement at different installations of the supply chain. It assumes a base stock type policy across the network. After the initial work of Graves and Willems (2000) for strategic safety stock deployment in multi-echelon supply chains under relaxed assumptions, Graves and Willems (2005), Graves and Willems (2008), and Graves and Schoenmeyr (2016) extended the work to new products, non-stationary demand, and with capacity constraints respectively. Minner (1997) presented a dynamic programming approach for inventory control in serial, assembly, and distribution type networks. Minner (2012) presented a heuristic for the previous problem, and Minner (2001) provided inventory control solution for reverse logistics networks. Lesnaia (2004) provided optimal safety stock solution for general multi-echelon networks. Humair and Willems (2006) introduced clusters of commonality (CoC) in supply chain networks, and provided solution for networks having CoC. Magnanti *et al.* (2006), Shu and Karimi (2009), Humair and Willems (2011), Li and Jiang (2012), and Grahl *et al.* (2016) presented heuristic approaches for safety stock deployment in GSM framework for general acyclic networks.

Models which are divergent from the classical SSM and GSM approaches have been studied very recently. Ease of practical application and tractability are some of the reasons. We will discuss some of those approaches next. While SSM and GSM approaches are pivotal to inventory optimization literature, in recent years, there has been growing interest on hybrid models, robust optimization models, rolling horizon models, multi-period newsvendor models, and multi-objective models. There are very few papers that discuss the comparative effectiveness of SSM and GSM approaches under different application settings. Also, for general multi-echelon networks there has been no solution under the SSM approach. Klosterhalfen and Minner (2007) presented a comparative study of SSM and GSM approaches. Klosterhalfen and Minner (2010) analyzed SSM and GSM approaches for distribution networks. They show that the cost difference between the approaches can be at most 4%, and the GSM approach is superior for moderate flexibility costs, large warehouse processing times and high retailer service levels.

Klosterhalfen *et al.* (2013) presented an integrated hybrid-service (HS) approach. Instead of every stage of the supply chain operating on SSM or GSM, in HS approach each stage has a choice to choose either SSM or GSM depending upon the flexibility involved, and demand. The HS model mitigated the risk of choosing either of the pure models. Hence, the HS model performs at least as the better between SSM and GSM. Bertsimas and Thiele (2006) introduced robust optimization techniques to supply chain inventory control problems. Interestingly, they yield similar optimal policies as with dynamic programming (SSM, GSM, and HS), but with less computational efforts for large scale problems. However, there is an additional robustness cost.

2.2.6. Position of Our Work

In the previous sections, we have presented some of the relevant works pertaining to our research. In this section, we summarize the contributions of those works and identify the gaps that still exist. We demonstrate where our contributions are able to fill the gaps and position our work with respect to the state-of-the-art. The contributions are relative positioning of our research are presented in Table 2.1) for the single-item replenishment planning and single-item replenishment planning with supplier selection problems.

Our work in this dissertation addresses the following. First, we propose a classification scheme for inventory optimization problem. We identify Gayraud *et al.* (2015) and de Kok *et al.* (2018) as the state-of-the-art in this area. They address network optimization and multi-echelon inventory optimization respectively. Our work addresses general inventory optimization models from a practical replenishment planning viewpoint. We focus on problem parameters that are needed for developing solution instead of classifying existing problems.

Secondly, for performance assessment, we identify Stangl and Thonemann (2017) and Petropoulos *et al.* (2019) as the the state-of-the-art literature. We list the key performance indicators presented in those works. We found one key challenge while assessing the performance of a new IMS with respect to an existing IMS. The two systems cannot be tested under identical conditions in reasonable time. In current market situations, users require a numerical assessment before adopting a new system. Hence, we also propose a novel assessment methodology based on a simulation approach that enables relative comparison under identical conditions.

For the part of solution development, we have addressed three base problems: single-item replenishment planning (SRP), single-item replenishment planning with supplier selection (SRPSS) and promotional joint replenishment planning (PJRP). All of them are under stochastic demand scenario. For the SRP problem, Levi *et al.* (2007a), Özen *et al.* (2012), Rossi *et al.* (2015) and Xiang *et al.* (2018) are identified as the state-of-the-art literature. We analyze them under three parameters: dealing with stochasticity, flexibility of proposed method and ease of industrialization. Levi *et al.* (2007a), Özen *et al.* (2012), Rossi *et al.* (2015) and Xiang *et al.* (2018) all consider only demand to be stochastic. Levi *et al.* (2007a) alone consider a distribution free approach. The remaining assume the demand to follow a fully known parametric distribution. None of them considers additional stochastic parameters. In terms of flexibility, we analyze the ease and possibility of the method being extended to include additional constraints and parameters. Similarly, for the SRPSS problem, we consider Fox *et al.* (2006), Zhang *et al.* (2012), Cheaitou and Van Delft (2013)Firouz *et al.* (2017) as the state-of-the-art. We summarized our contributions and that of the state-of-the-art in Table 2.1 on the SRP and SRPSS problems.

Our works in regards to the SRP and SRPSS problems differ from the identified state-of-the-art literature in all three aspects namely: dealing with stochasticity, flexibility of the proposed methods and ease of industrialization. For the base problems, we consider only the external demand to be stochastic. However, the method is readily adaptable to stochastic lead times and delivery quantities. Our proposed approaches are sampling-based. So, they do not require the complete distributions of the stochastic parameters. In addition, our proposed methods can be easily extended to include practical constraints such as batch-size, minimum order quantity and service level. Moreover, the methods can be used in problems with lost sales and non-stationary demand. Also, they can be extended to have robust solutions in case of higher uncertainty. The proposed methods are computationally less expensive and scalable, hence, suitable for industrialization.

The PJRP problem is novel. We found only one work Taleizadeh *et al.* (2008), that considers batch-size for the multi-item newsvendor problem. The PJRP problem differs due to the presence of prepacks (multi-item packagings). Gao *et al.* (2014a) is the only work to consider prepacks but for multi-period problems. Also they do not consider a mix of prepacks for a single item. Our application condition also differs due to the absence of financial information. Our solution approach is two-fold. We first propose a multi-objective approach for small to medium problem instances to obtain multiple Pareto optimal solutions. Then, for industrial size problems, we propose a single-objective approach along with a metaheuristic. Our approach has been also validated on a real-world case study.

2.3. Classification Scheme for Inventory Optimization Problems

We found two recent articles; Gayraud *et al.* (2015) and de Kok *et al.* (2018), that have discussed classification and typology of different inventory optimization problems in logistics and multi-echelon networks respectively. Both have classified only existing problems in the literature based on some problem dimensions. One major aspect they miss is the length of the planning horizon of the concerned problem which is essential for solving any inventory optimization problem. Beside, they also do not consider the current trends of said problems in the industry. For example, most retailers are now adopting an integrated approach where supplier selection and inventory optimization are accomplished together. This indicates a major shift from discrete planning levels towards a more continuous (fluid) planning. In this dissertation, we focus on the specificities of inventory optimization problems from a practical viewpoint.

In the context of inventory optimization, there are two classes of decisions: *static inventory quantities* at a location in the supply chain and the *flow of inventory* between two locations. While the former can be otherwise stated as the decisions regarding safety stock quantities and/ or target stock quantities, the latter is termed as replenishment planning. Although there is a wide range of inputs that are considered while solving an inventory optimization problem, they can stem from six basic components, which are:

1. the supply chain network in consideration and its dynamics,
2. temporal aspects,
3. customer and external aspects,
4. objective of optimization,
5. internal control measures and
6. problem parameters.

Table 2.1: State-of-the-art and position of our work.

		Uncertainty				Flexibility						Sourcing	
Problem	Work	Demand	Lead time	Others	Distr. free	Lost sales	NS demand	Batch size	MOQ	Service level	Robust solution	Runtime	Scalability
SRP	Levi *et al.* (2007a)	✓			✓		✓					High	Limited
	Özen *et al.* (2012)	✓					✓		✓	✓		High	Limited
	Rossi *et al.* (2015)	✓					✓		✓	✓		Medium	Limited
	Xiang *et al.* (2018)	✓					✓					Medium	Limited
	This work	✓	✓[1]	✓[2]	✓	✓	✓	✓	✓	✓	✓	Low	Sufficient
SRPSS	Fox *et al.* (2006)	✓										Low	Limited
	Zhang *et al.* (2012)	✓				✓						Medium	Limited
	Cheaitou and Van Delft (2013)	✓					✓					Low	Limited
	Firouz *et al.* (2017)	✓											
	This work	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Low	Sufficient

[1] Extendable to incorporate stochastic lead time.
[2] Extendable to incorporate other stochastic parameters as samples.

The rest of this section is arranged as the following. In Section 2.3.1, we discuss the motivations and objectives behind classifying inventory optimization problems. Next, we elaborate on the classification scheme in Section 2.3.2. In Section 2.3.3, we illustrate the scheme with some examples from the recent literature.

2.3.1. Motivations and Objectives

Several review papers have addressed and structured different fields of inventory optimization. However, unlike other fields of operations management and operations research, inventory optimization still lacks a standard classification. de Kok *et al.* (2018) identified the absence of a structured classification of model assumptions in case of multi-echelon stochastic inventory optimization. They also recorded a basic definition of such method: "A classification method that enables a fast and well-accepted terminology for core assumptions of a model and its underlying system". In other areas of operations management and operations research, classification schemes are common. For example, Buzacott and Shanthikumar (1993) and Kendall (1953) for queuing models, Dyckhoff (1990) and Wäscher *et al.* (2007) for cutting and packing problems, Graham *et al.* (1979) for scheduling models, Brucker *et al.* (1999) for project planning, Boysen *et al.* (2007) for assembly line balancing problems, and Copil *et al.* (2017) for lot-sizing and scheduling problems. As mentioned earlier, the aim of this dissertation is two-fold. The first one is to build a common framework for inventory management problem across various domains. A classification scheme is necessary to avoid excessive customization. The objectives of our classification are as follows.

1. To state all important dimensions of modeling assumptions (for example, the structure of the system, the demand processes, the replenishment processes, the objectives, constraints, etc.) so that problems and methods can be compared appropriately.
2. To be able to characterize real-world supply chain problems, therefore, enabling a generalized solution development.

2.3.2. Classification Scheme

We have identified six major dimensions of an inventory optimization problem. They are:

1. Supply chain network,
2. Temporal aspects,
3. Customer and external factors,
4. Optimization objectives
5. Internal control measures and
6. Problem parameters.

Each of these dimensions are presented by a field in the classification scheme. Each of the dimensions has sub-dimensions that are presented by sub-fields. The proposed classification scheme is defined in Figure 2.2. The fields are separated by || and the sub-fields are separated by a semicolon (;). If any of the sub-fields is not applicable or its value is unknown for the concerned problem then, it is denoted by "ϕ". Otherwise, the sub-fields are denoted by the notations given in Table 2.2.

Field 1 Field 2 Field 3

<Echelons>;<Structure>;<Special measure>‖<Lead time>;<Time axis>‖<Demand>;<Shortage>;<Service level>‖
<Objective>‖<Inventory policy>;<Lot size>‖ <Products>;<Suppliers>;<Time horizon>

Field 4 Field 5 Field 6

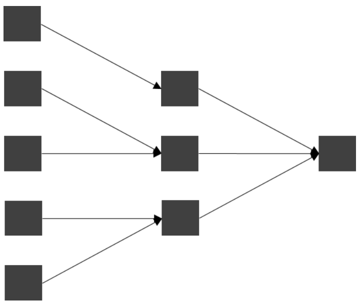
Figure 2.2: Proposed classification scheme.

An inventory optimization problem can be denoted by the above scheme, by specifying the corresponding notation for each sub-dimension. All dimensions along with their sub-dimensions are elaborated next.

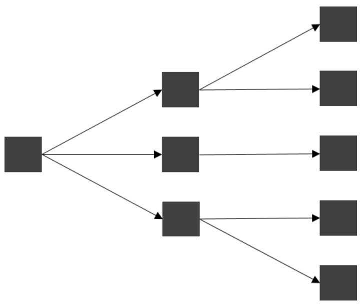
Field 1: Supply Chain Network

The nature of supply chain network greatly influences the solution method that needs to be adopted. For example, solution method for a single echelon supply chain differs from that of a supply chain with multiple echelons or a logistical supply chain. Latest inventory management practices do not keep safety stock in case of single echelon supply chains, however, it is needed in case of multi-echelon supply chain for smooth operation. According to arrangement of echelons, supply chain networks can be broadly divided into three types: single-echelon (1), multi-echelon (N) and logistic (L).

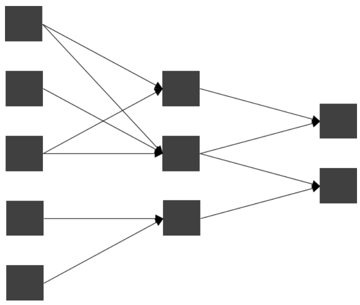
Multi-echelon supply chains also differ by their structure. Some common structures are convergent (assembly type, Figure 2.3a, C), divergent (distribution type, Figure 2.3b, D), general-acyclic (convergent-divergent without cycles, Figure 2.3c, G_a), general-cyclic (convergent-divergent with cycles, Figure 2.3d, G_c), and serial supply chains (Figure 2.3e, S) etc.



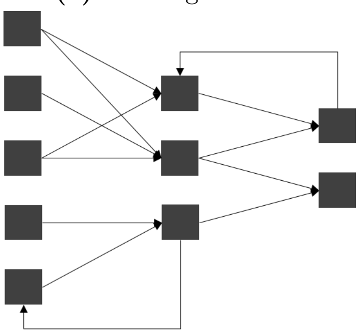
(a) Convergent SCN.



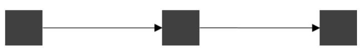
(b) Divergent SCN.



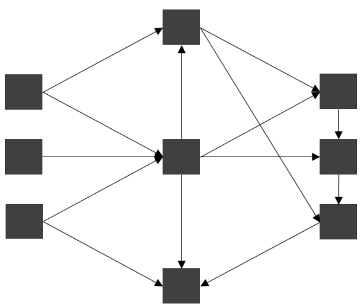
(c) General acyclic SCN.



(d) General cyclic SCN.



(e) Serial SCN.



(f) Logistic SCN.

Figure 2.3: Different types of supply chain networks.

Field 2: Temporal Aspects

The time axis in an inventory optimization problem can be considered continuous (C) or discrete (D). From a practical viewpoint, time is always considered to be discrete. Hence, continuous review inventory policies find little application in practice. That being said, another temporal consideration is delay or lead time. It is considered either deterministic (D), stochastic (S) or zero (0). Most research considers it to be deterministic. However, surveys of logistic providers suggest, in practice, it is more often stochastic and different from what is posted in the management information system (MIS).

Field 3: Customer and External Factors

Customer and external factors pertains to the dimensions that are external to the system. Demand is the first sub-field. It can be considered in three different ways: deterministic (D), stochastic-stationary (S) and stochastic-non-stationary (N). The behavior of an external customer in case of unfulfilled demand constitutes the second factor. The options are lost sale (L), backorder (B) and zero shortage (0). Lost sale refers to the case where the unfulfilled demand is lost completely and it does not appear along with future demand. Backorder refers to the case where the unfulfilled demand is not lost but, appears as the additional demand of the upcoming periods. Full backorder means the unfulfilled demand of the particular period appears as the additional demand of the next period. In some practical situations there can also be the case where the demand is partly lost and partly backordered (M). In case of deterministic demand, all the four options are applicable. However, in case of stochastic demand zero shortage is an impractical approach, and only lost sale, backorder or a mix of lost sale and backorder approach is adopted. In addition to the above, the presence of a service level constraint can be divided into $\alpha-$service level (X) or $\beta-$service level (Y).

Field 4: Optimization Objectives

Most inventory optimization methods aim to minimize total cost (C) or maximize profit (P). Recently, some problems are considered with multiple objectives (M) where, the target is to obtain the efficient frontier (Pareto optimal points).

Field 5: Internal Control Measures

This factor represents the business aspects. It covers the inventory policy followed (if any), lot-sizing, coverage periods, minimum order quantities, etc. Even if these measures do not give the globally optimal solutions, they are implemented to ensure a smooth operation or sometimes they are dictated by external suppliers or customers. Broadly, the inventory policy followed can be divided into periodic review (P), continuous review (C) and dynamic (D). Similarly, presence of lot-size can be of the form unit-size (U), single lot-size of non-unit quantity (S) and multiple lot-sizes of non-unit quantities (M).

Field 6: Problem Parameters

Apart from the above dimensions, problem parameters such as number of products, number of suppliers and time horizon are also needed to formulate the problem. In case of single-item (S), inventory optimization problems are relatively simple than their multi-item (J) counterparts (joint replenishment problems). The optimal inventory policy structure and its efficient evaluation methods are available for most single product cases. However, for the joint replenishment planning, optimal inventory policy structures are available for very few practical cases and its practical evaluation method is still being studied and developed. Similarly, single-supplier (`S`) and multi-suppliers (`M`) problems require very different solution methods. This is also true in case of single-period (`S`) and multi-period (`M`) problems. All fields and their respective sub-fields and their notations are summarized in Table 2.2.

Table 2.2: Proposed notations for different sub-field types used for the classification scheme.

Field (Dimension)	Sub-field (Sub-dimension)	Specifics	Notation
Supply chain network	Echelons	Single	1
		Multiple	N
		Logistic	L
	Structure	Serial	S
		Convergent	C
		Divergent	D
		General acyclic	$\mathtt{G}_a$
		General cyclic	$\mathtt{G}_c$
	Special measure	Transshipment	T
		Sub-contracting	S
Temporal aspects	Lead time	Deterministic	D
		Stochastic	S
		Zero	0
	Time axis	Continuous	C
		Discrete	D
External factor	Demand	Deterministic	D
		Stochastic (Stationary)	S
		Stochastic (Non-stationary)	N
	Shortage	Lost sale	L
		Backorder	B
		Lost sale and backorder	M
		Zero	0
	Service level	Service level (α)	X
		Service level (β)	Y
Optimization objective	–	Cost	C
		Profit	P
		Multiple	M
Control measures	Inventory policy	Periodic review	P
		Continuous review	C
		Dynamic	D
	Lot size	Unit	U
		Single	S
		Multiple	M
Problem parameter	Items	Single	S
		Joint (Multiple)	J
	Suppliers	Single	S
		Multiple	M
	Time horizon	Single period	S
		Multi-period	M

2.3.3. Classification Examples

In this section, we use the classification scheme defined in Figure 2.2 and codify the inventory optimization problems discussed in this dissertation. They are presented in Table 2.3. The columns “Problem”, “Shortage” and “Batch-size” describe the problem and the column “Notation” describes the code as per the proposed classification scheme.

The stochastic single-item replenishment planning problem for one store can be written as `1;`ϕ`||0;D||S;L||C||D; U||U;S;M`. The supply chain is single-echelon and does not have any multi-echelon structure, hence, the first field is denoted by `1;`ϕ. The lead time is assumed to be zero and time axis is considered to be discrete. The second field is denoted by `0;D`. Demand is stochastic and stationary. Unfulfilled demand is lost completely and there is no service level constraint. The third field is therefore noted by `S;L;`ϕ. The objective is cost minimization. Inventory control policy is dynamic and the batch-size is one unit. The fourth and fifth fields are respectively `C` and `D;U`. The problem parameters are as follows. The problems addresses a single item, single supplier and multiple time periods. The sixth field is therefore `S;S;M`.

2.3.4. Recent Practices in Inventory Management

Inventory management practices across industries are evolving, largely driven by the advances in IT capabilities. Development of more connected systems enabling implementation of central planning solutions. As per Pettey (2019), eight emerging technology trends are changing supply chain practices. They are, artificial intelligence, advanced analytics, internet of things, autonomous things, supply chain digital twin, robotic process automation, immersive experience and blockchain. They are discussed in detail hereafter.

Artificial Intelligence

Artificial intelligence (AI) supports the vision for a broader supply chain automation of an organization. There can be three levels of automation: semiautomated, fully automated or a mix of both. It depends on the application conditions. AI solutions are helping in automating various supply chain processes through self-learning and natural language. The most impacted processed are demand forecasting, production planning or predictive maintenance, etc. AI in supply chain consists of technologies that seek to emulate and surpass human performance by improving service levels, deliveries, last mile routing, etc.

Table 2.3: Examples of classification using the proposed scheme.

Problem	Shortage	Lot-size	Notation
Single-item ordering with sample forecasts	Lost sale	Unit	1;ϕ;ϕ||0;D||S;L;ϕ||C||D;U||U;S;M
	Lost sale	Single	1;ϕ;ϕ||0;D||S;L;ϕ||C||D;S||U;S;M
	Backorder	Unit	1;ϕ;ϕ||0;D||S;B;ϕ||C||D;U||U;S;M
	Backorder	Single	1;ϕ;ϕ||0;D||S;B;ϕ||C||D;S||U;S;M
Single-item ordering with sample forecasts and supplier selection	Lost sale	Unit	1;ϕ;ϕ||0;D||S;L;ϕ||C||D;U||U;M;M
	Lost sale	Single	1;ϕ;ϕ||0;D||S;L;ϕ||C||D;S||U;M;M
	Backorder	Unit	1;ϕ;ϕ||0;D||S;B;ϕ||C||D;U||U;M;M
	Backorder	Single	1;ϕ;ϕ||0;D||S;B;ϕ||C||D;S||U;M;M
Multi-item ordering with sample forecasts	Lost sale	Unit	1;ϕ;ϕ||0;D||S;L;ϕ||C||D;U||J;S;M
	Lost sale	Single	1;ϕ;ϕ||0;D||S;L;ϕ||C||D;S||J;S;M
	Backorder	Unit	1;ϕ;ϕ||0;D||S;B;ϕ||C||D;U||J;S;M
	Backorder	Single	1;ϕ;ϕ||0;D||S;B;ϕ||C||D;S||J;S;M
Multi-item ordering with sample forecasts and supplier selection	Lost sale	Unit	1;ϕ;ϕ||0;D||S;L;ϕ||C||D;U||J;M;M
	Lost sale	Single	1;ϕ;ϕ||0;D||S;L;ϕ||C||D;S||J;M;M
	Backorder	Unit	1;ϕ;ϕ||0;D||S;B;ϕ||C||D;U||J;M;M
	Backorder	Single	1;ϕ;ϕ||0;D||S;B;ϕ||C||D;S||J;M;M
Multi-item promotional ordering with probabilistic forecast and prepacks	Lost sale	Unit	1;ϕ;ϕ||0;D||S;L;ϕ||C||D;S||J;S;S
	Lost sale	Single	1;ϕ;ϕ||0;D||S;L;ϕ||C||D;M||J;M;S
	Backorder	Unit	1;ϕ;ϕ||0;D||S;L;ϕ||M||D;S||J;S;S
	Backorder	Single	1;ϕ;ϕ||0;D||S;L;ϕ||M||D;M||J;M;S

Advanced Analytics

Even before AI, the impact of advanced analytics on supply chain decision-making and performance is significant. Related practices are being deployed in real time. Item quality testing and dynamic replenishment are some of the areas. The availability of supply chain data from Internet of Things (IoT), point of sales and external sources helps to understand the current environment to better predict future scenarios. This also helps in making better decisions.

Internet of Things

Incorporation of IoT has improved the end-to-end visibility of supply chains profoundly. With the ability of tracking each item in real time, IoT adoption is seeing a growing trend. The IoT can have a broad impact on asset utilization, uptime, customer service, supply chain performance, product availability and reliability.

Autonomous Things

Autonomous things are often physical devices operating in the real world, such as robots carrying out various jobs and cameras assisting in checking inventory quality. They are enabling new business scenarios and optimizing existing ones. The rapid growth in the number of interconnected, intelligent things has augmented this trend.

Digital Supply Chain Twin

A digital twin is defined as the digital representation of a real-world system. A supply chain digital twin is a digital counterpart of a supply chain presenting the relationships between all its impacting entities. It improves the end-to-end visibility of the actual supply chain. This also improves decision-making speed.

Robotic Process Automation

Robotic Process Automation (RPA) devices eliminate keying errors, enhance process speeds, reduce costs and link applications. RPA has been effective in simple use cases, especially in the absence of APIs or other means for automated data integration.

Immersive Experience

Supply chain leaders can use immersive experience platforms such as virtual reality (VR) and augmented reality (AR) to save time, and make repetitive tasks easy. For an example, using AR to provide renderings of equipment to visualize the footprint in a defined space to compare different configuration options or using voice-controlled personal assistants to remotely check product features or appointments.

Blockchain

Blockchain could potentially fulfill some long-standing challenges across global supply chains. Current blockchain offerings for supply chain include a loose portfolio of technologies and processes that spans database, verification, security, analytics, and contractual and identity management concepts. Blockchain is also being offered as a service or development option across supply chain solutions with objectives such as automation, traceability and security.

2.4. Conclusions

In this chapter, we provided the state-of-the-art literature pertaining to our research. We analyzed the literature in four broad fields: inventory optimization problem classification, inventory management system performance evaluation, single-echelon inventory optimization and multi-echelon inventory optimization. In this dissertation we address the first three areas. We identified the shortcomings in the literature and positioned our work in that respect.

Afterwards, we also presented a classification scheme for general inventory optimization problems. For practical purposes we focused on six major dimensions of an inventory optimization problem. Those dimensions were then incorporated into the classification scheme along with their sub-dimensions as fields and sub-fields respectively.

In the next chapter, we discuss about assessing the performance of real-world inventory management systems along with their key performance indicators.

Chapter 3

Performance Assessment of Inventory Management Systems

Abstract: *This chapter addresses the issue of performance assessment of inventory management systems (IMSs). Performance assessment is essential while selecting a new IMS and also during day-to-day operations. It broadly comprises two activities: selection of key performance indicators (KPIs) and computation of those KPIs. Selection of suitable KPIs is moportant as improper choice may result in selecting a wrong IMS and the associated cost implications can be very high. Further, computation of those selected KPIs can be very challenging for real-world IMSs. This is due to the unavailability of identical testing conditions for the existing and new IMSs, and very high duration for a field testing. In this chapter, we prepare a list of KPIs that are suitable for IMSs in retail, spare parts and industrial supply chains. These industries are chosen from the business perspective of Vekia. We then propose a simulation method called the* $\Delta-$*Method, to compute the previously defined KPIs while assessing a new IMS with respect to an existing one. Part of this chapter has been presented (Sahu et al., 2020a) in the 34th annual conference of the Belgian Operational Research Society (ORBEL-2020) (CW1.).*

3.1. Introduction

The current competitive environment has challenged organizations to make their supply chains balanced between flexibility, speed, quality and responsiveness at the low cost (Martin and others, 2010). Inventory management is one of the key activities in managing a supply chain and its performance is closely linked to that of the latter. The traditional financial metrics can give misleading signals of continuous improvement in innovation in supply chains (Kaplan, 2009). Therefore, several authors have proposed non-financial metrics in addition to the financial ones for supply chain performance assessment (Gunasekaran *et al.*, 2001; Beamon, 1999). For inventory management as well, Petropoulos *et al.* (2019) have suggested metrics under financial, operational and service category. This can give a holistic view of overall performance.

Apart from evaluating the health of the system, performance assessment is also considered as one of the important aspects of inventory management due to the following. Organizations must evaluate the IMS while implementing an IMS or selecting a new IMS. And, a new system can only be adopted if it fares better compared to the other competing systems (when there are multiple competing systems) or it performs better than the existing one. Performance assessment is also necessary for technology vendors such as Vekia as it helps decide an appropriate value proposition.

Performance assessment comprises two major activities. They are:

1. Selection of appropriate KPIs.
2. Computation of selected KPIs.

Each business can have a different set of KPIs that reflects its priorities. For example, service level measures are more suitable performance indicators for a home service business than inventory turnover. On the other hand, retail businesses prefer inventory turnover over service level. After the selection of suitable KPIs, their computation method must also be formalized.

In this chapter, we list the critical KPIs pertaining to inventory management and propose a simulation method to compute them in practical situations. The upcoming sections of this chapter are arranged as follows. In Section 3.2, we briefly discuss the motivations and objectives of this chapter. Section 3.3 reviews some preliminary concepts. We review the existing KPIs across different businesses and provide a comprehensive list of the KPIs in Section 3.4. In Section 3.5, we present the evaluation framework for the KPIs. We provide the conclusions in Section 3.6.

3.2. Motivations and Objectives

From the year 2000, researchers' focus on the domain of supply chain performance measurement approaches and techniques has increased remarkably (Balfaqih *et al.*, 2016). Initial literature focused on developing integrated evaluation frameworks and categorization of performance measures. Over the recent years, however, more attention has been given to the identification of KPIs, implementation of those performance measures and measurement of environmental impacts of supply chains. This, in our view, is due to the shift in business conditions and the resulting changes in the researchers' view of supply chain performance.

Identification of appropriate KPIs for the concerned business model is important as the performance measurement system of an organization has a significant role in managing businesses and supply chains. As per Kaplan and Norton (1996), "No measures, no improvement". This statement is relevant for an in-use IMS as well as when migrating into a new one. It is critical to measure the right thing at the right time so that, decisions can be taken timely. Some drawbacks of performance measurement systems, as pointed out by Balfaqih *et al.* (2016) are the lack of connection between organizations' strategies and the KPIs used, the lack of linking KPIs to customer value, the biased concentration on financial metrics and the existence of several conflicting performance measures (Brewer, 2000).

Lack of standard and comprehensive performance measures for inventory management forces organizations to adhere to legacy systems. Performance evaluation, therefore, appears as a obvious need for designing/modifying/improving the IMS for an organization. But, in practice, assessing an IMS's performance is very situation-specific and complicated. We explain it in the following paragraph.

Historical closing inventory and sales data are two important and basic data points needed to assess the inventory management performance of an organization. This can be obtained from ERP systems and can be used to compute the inventory turnover and other KPIs. However, organizations face difficulties in assessing the service performance where the actual demand/orders are not recorded, for example, in retail stores. We can compute some KPIs for the existing system. However, this is not true in the case where the organization wants to adopt a new system or it wants to compare some competing systems. In order to compute the KPIs for new systems, either the organization can follow AB testing (Gilotte *et al.*, 2018) or it can conduct a simulation study on synthetic dataset. In case of AB testing, the time required to obtain enough data for reliable computation of KPIs is very high (spanning upto a year). This procedure also has the risk of financial losses if the new system does not perform well. On the other hand, simulation yields approximate results and it usually does not take forecast error into account. Additionally, simulations on synthetic data are usually run in a different time frame than the existing system. This causes non-identical test conditions.

The major use case of the research in this chapter is to assist clients in selecting a new IMS. In this context Galasso *et al.* (2016) pose some questions. The efficacy of the process depends on the capacity of the decision maker to assess

1. the current performance, i.e. What happened until today? What is the current progression (in a broader sense)? and
2. the possible IMSs and their related impact, i.e. What will happen and what are the consequences?

In addition, we pose some other questions and address them in this chapter.

1. Are the values of the performance measures computed/simulated during different time frames comparable?
2. How can we reliably compare the existing system and the new system in the same time frame?
3. How can we compare multiple new systems with inaccuracies in forecasts?

Currently, KPIs are subjective to industry practices. As a solution provider, it is important for Vekia to suggest some standard KPIs to its clients that are relevant to the respective industries as well as give clear interpretable performance measures. Computation of those KPIs and assisting the clients in the same are also important. The objectives of this chapter are outlined as such.

1. Propose and standardize KPIs for inventory management in supply chains.
2. Propose a framework to enable *a priori* computation and comparison of those KPIs corresponding to different IMSs (existing and new) under the same application conditions.

3.3. Preliminaries

In this section, we review some preliminary concepts of IMSs. First, we discuss the workflow of an IMS. Afterwards, we discuss the dynamics and parameters of an inventory system.

The primary functions of an IMS are demand forecasting and order proposition (or replenishment planning). Both can be presented as separate systems inside an IMS (see Figure 3.1). The forecasting system takes historical and future organizational data, and exogenous data as inputs, then provides demand forecasts (either in deterministic or in probabilistic form) as output. These forecasts are then used in the replenishment planning system along with other constraints to optimize the order quantities.

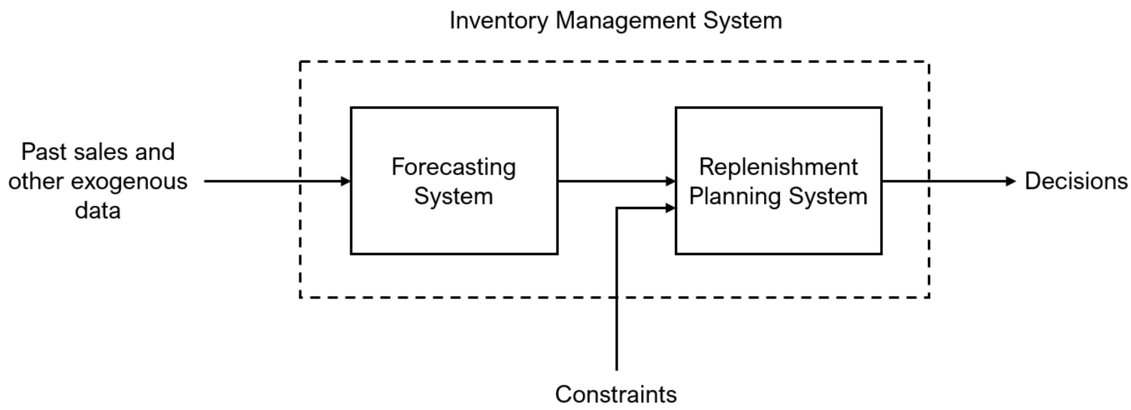


Figure 3.1: Workflow of an inventory management system. Generally IMSs comprise two subsystems: forecasting system and replenishment planning system. The forecasting systems uses historical sales and exogenous data to predict future demand using machine learning algorithms. The replenishment planning system uses those forecasts and the ordering constraints to suggest optimized replenishment decisions.

The dynamics of an inventory system (see Figure 3.2) refers to the detailed step-by-step operations involved. They can be explained as followed. At the beginning of time t, opening inventory is s_{t-1}, which is the closing inventory of the previous time period $t-1$. First, the order supposed to be delivered at time t, Q_t is received and the inventory position is raised to $s_{t-1}+Q_t$. Then demand for the period t, D_t is realized. If the demand D_t is less than the inventory position then, a positive inventory s_t is left at the end of time t. Otherwise, if the demand is more than the inventory position the system encounters lost sales LS_t.

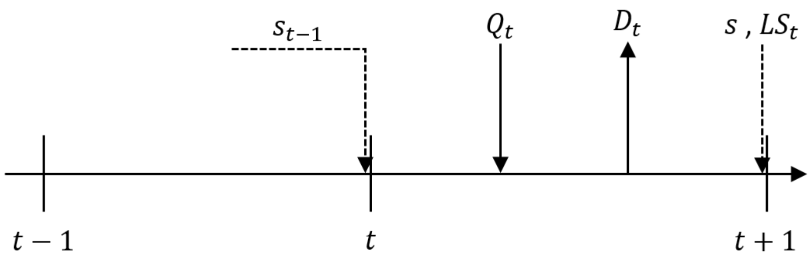


Figure 3.2: The dynamics of an inventory system. Various basic events of an inventory system are depicted. The flow of events are, recording opening inventory, order delivery, demand realization, and recording closing inventory or shortages.

Performance assessment is always relative (see Figure 3.3), i.e. the value of a KPI cannot be interpreted in isolation as good or bad. Either it must be assessed against its previous value or its corresponding values obtained using the new IMSs must be compared depending upon the situation.

Definitions

Some definitions used throughout this chapter are presented in the following paragraph.

In Figure 3.2, at any time t, the planner takes a decision regarding the order quantity considering the demand in the future and the inventory position. Here, the *inventory position* is defined as the on-hand inventory plus the order receivables minus the reserved quantity. *Shortage* happens when the demand received is higher than the inventory position. *Lost sale*

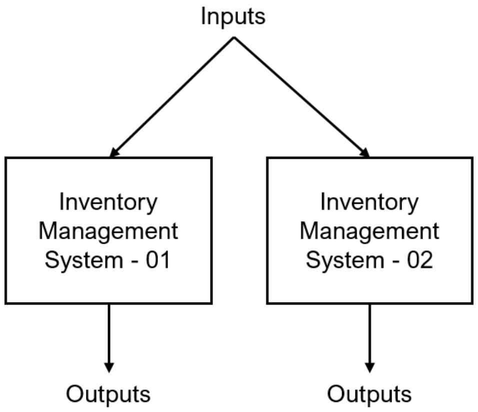


Figure 3.3: Relative assessment of two inventory management systems. The comparison can be done between two new IMSs or between new and existing systems.

is the part of demand that is not fulfilled and lost completely, i.e. the unfulfilled demand does not appear as the additional demand in the future. On the other hand, *backorder* is the part of demand that is not fulfilled but, it appears as the additional demand in the future. The inventory quantity for a item typically follows a pattern as mentioned in the Figure 3.4. For simplicity, a constant demand rate is depicted. It starts with some positive inventory and it decreases with each fulfilled demand. Depending on the inventory policy, orders are placed at specific intervals or at a specific inventory threshold *reorder level* or in a dynamic manner.

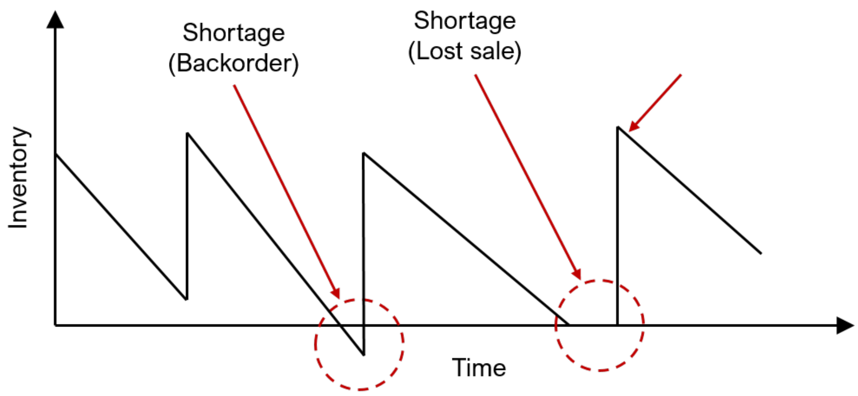


Figure 3.4: Evolution of inventory with time

3.4. KPIs for Inventory Management Systems

The overall performance of any IMS can be expressed as a set of KPIs. Those KPIs can be broadly divided into three categories (Petropoulos *et al.*, 2019). They are:

1. Financial
2. Operational
3. Service

As the name suggests, financial KPIs describe the economic performance of inventory management. Operational KPIs reflect how efficiently the operations are managed related to inventory and service KPIs give a picture about how well the customer demand is fulfilled. Operational and service KPIs are essential to assess the performance from practical point of view. Ideally, the financial KPIs should complement the operational and service ones. Performance of any inventory system can be expressed through any KPI belonging to either of the above categories. However, their computation and relevance depends on the problem specifics. For example, KPI computation methodology for a single period inventory system might not be valid in multi-period case. Next, we specify different KPIs under each category.

3.4.1. Financial KPIs

Under financial KPIs, we list three KPIs: value chain profitability (VCP), total inventory value (TIV) and inventory turnover (ITO) (Rao and Rao, 2009).

Supply chain profitability (SCP) (Gawankar *et al.*, 2013), also known as supply chain surplus, is a common term that represents value addition by the supply chain function of an organization. Its operational concept is, "sharing the profit that remains after subtracting costs incurred in the production and delivery of products or services". The *value chain* is a set of activities that an organization carries out to create value for its customers. Porter (1985) proposed a general-purpose value chain that companies can use to examine all of their activities, and see how they are connected. Common activities can be classified as Inbound logistics, Operations, Outbound logistics, Sales and marketing, and Services. We propose *value chain profitability* (VCP), which is similar to the SCP but, limited in scope. While SCP is the surplus revenue after deducting the whole supply chain cost, VCP is the profitability of part of the system (value chain), say, inbound logistics or outbound logistics. These are two major activity classes, that always comprise inventory control. In the profitability model (See Figure 3.5) VCP does not consider investments. VCP can be mathematically expressed as

$$\text{VCP} = \text{Revenue} - \text{Cost incurred} \tag{3.1}$$

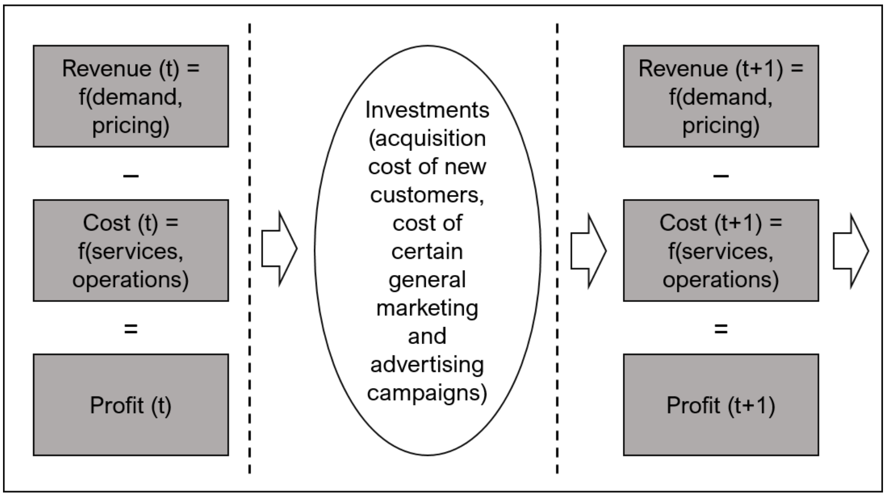


Figure 3.5: Profitability model

Computation of VCP requires the values of total revenue and total cost. Total revenue is equal to the number of units sold times the unit selling price of the respective items. Total cost comprises various components. They are purchase cost of the goods, transportation cost, fixed ordering cost, penalties, inventory holding cost and obsolescence cost, etc. VCP can be calculated at the end of each review period. However, in the presence of fixed costs the frequency can be different for different costs. For example, inventory cost can be computed on daily basis but transportation cost cannot be correctly estimated. Therefore, VCP calculation over a longer horizon gives a better picture. VCP can be computed as follows.

$$\text{VCP} = \text{R} - (\text{CoG}+\text{TC}+\text{FC}+\text{L}+\text{OC}+\text{HC}) \tag{3.2}$$

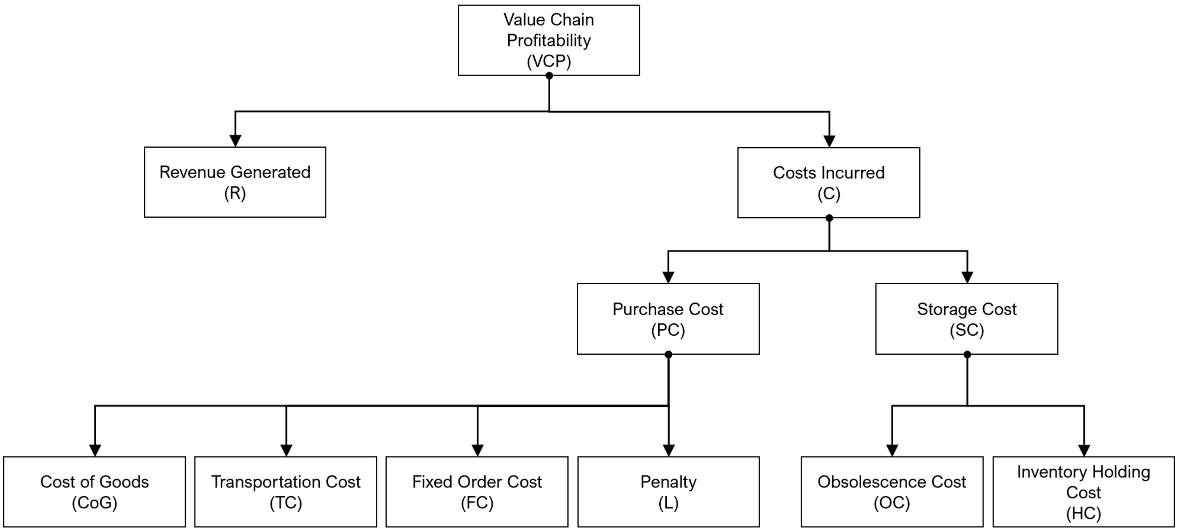


Figure 3.6: Components value chain profitability

The notations are presented in Figure 3.6. The cost of shortage is not added as it is not a real cost that is paid by an organization. However, when VCP is calculated the effect of shortage is reflected in the revenue generated. Interpretation of VCP is quite straightforward. It represents profit. Therefore, when two IMSs are compared, the system having more VCP can be regarded as better.

The *total inventory value* (TIV) is defined as the monetary value of total closing inventory of a period. It can be computed for the whole supply chain or for a location (warehouse, retailer, etc.). It represents the value of total capital blocked. In the case of non-stationary demand (eg: promotions, seasonal demand, etc.) TIV can vary as per demand. However, in case of stationary demand, it remains more or less the same. This also reflects the effectiveness of inventory classification methods. For example, in the common ABC classification, products are classified as per their annual consumption value. Class-A products are subjected to tighter inventory controls. In practice, they should represent the least of the inventory items but comparatively high inventory value.

Computation of TIV requires the closing inventory record and unit purchase price for each item. However, the average TIV over a longer horizon should be computed for a valid assessment. For comparison, two IMSs should start from same TIV. The system that generates lower TIV can be regarded as better if and only if it does not generate significantly lesser sales than the other. The TIVs at two different times of the same IMS cannot be compared as it may be the result of non-stationary demand or inventory policy.

Inventory turnover (ITO) (Rao and Rao, 2009) is defined as the ratio of total cost of goods sold to the average value of the stock held during that period. ITO gives an indication about the pace of inventory movement. For example, during one month 100€ worth goods are sold, and during that time 10€ worth of inventory is held on an average. This results in IT= 10. This means that during a month the inventory is rotated ten times or an item spends

$\frac{1}{10}$ month in inventory. Calculation of ITO requires closing inventory record and sales record of each item along with their buying cost. When comparing two replenishment planning systems, the system that generates higher ITO is better. This is because a higher ITO means items spend less time in inventory and hence, incurring lower inventory costs.

ITO should be calculated on a higher granular level because, depending upon the inventory holding cost or logistical availability the inventory quantity is decided. Therefore it can be misleading to compare performances of locations based on their ITO. Generally a warehouse will have less ITO that a store. Hence, overall ITO calculation of the whole network can give a better picture about the health of the organization. Similarly, ITO calculation for a group of items makes sense if inventory is suitably classified (e.g. ABC, VED, etc.).

3.4.2. Operational KPIs

Under operational KPIs, we list four KPIs: inventory level (IVL), inventory variance (IVV), order quantity variance (ORV) and inventory quality ratio (IQR).

The *total inventory level* (IVL) is defined as the closing inventory quantity for a period in number of units. It can be measured for the whole supply chain or at a location similar to TIV. It follows a similar profile to TIV in time. IVL is useful to assess the inventory performance of a supply chain, particularly when dealing with perishable products.

Computation of IVL requires closing inventory record of each item at each installation at the end of each time period. For a valid interpretation, we can consider the average IVL over a horizon. While comparing two IMSs, they should start from the same IVL. The system that generates lower IVL can be regarded as better if and only if it does not generate significantly lesser sales than the other. However, it is possible that individual IVLs of different items are different for the two systems while having the same total IVL. In this case the different classes or groups of the items must be taken into consideration.

Inventory variance (IVV) (Petropoulos *et al.*, 2019) can be interpreted as an indicator of operational efficiency. Organizations prefer to have gradual and minimal changes in the quantity of inventory held in the system. Erratic changes in inventory can result in over or under utilization of resources associated with inventory management. The closing inventory level stays nearly constant for stationary demand barring any special case such as very large packaging or large lead time, etc. In case of non-stationary demand also, gradual change is desired in order to avoid difficulties in facility planning. For example, a plan for promotion should be done in advance. A steep increase in inventory is not desirable as it may pose challenge towards warehouse operations planning. Also, sometimes, a huge order quantity may not be fulfilled by the supplier due to capacity and transportation constraints. As IVV represents the fluctuation in the inventory levels, the IMS that has lower IVV is better. Even though IVV is not considered as part of the objective during optimization of orders, it is expected of a good IMS.

If a steep rise in inventory with large order quantities in a short time horizon is feasible, the cost performance of such system will be better than that with anticipatory planning. In such scenarios, IVV may not be a suitable KPI. But this needs to be agreed upon by the stakeholders. IVV at any location (store or warehouse) gives a performance indication of that location. Overall variance can ignore the variances at individual locations and, therefore, cannot be properly interpreted. IVV is more applicable in retail supply chains because of their fast moving nature. In case of slow moving supply chains, it is of lesser significance.

Order quantity variance (ORV) (Cannella and Ciancimino, 2010) is of similar significance as inventory variance. Fluctuating order quantities are undesirable. Resource utilization

Table 3.1: Inventory classification as per IQR

Overall classification	Sub category	Description
Active (Inventory $\leq$ Target)	Deficit (Inventory <Min. Target)	Inventory that is below the targeted minimum level. Deficit inventory should not be considered healthy, because it poses a stock-out risk. From an IQR metric point of view, deficit inventories are considered active.
	Healthy (Inventory $\geq$ Min. Target)	Inventory is actively moving, and on-hand inventory value and level are within defined target levels.
Excess (Inventory $>$ Target)	Surplus (In Fast-moving products)	Surplus inventory is the portion of actively moving inventory that exceeds the defined target level.
	Slow-moving	There is forecast demand or usage for the product, but the on-hand level exceeds the six-month supply (six months is an example, because it is commonly used).
	Non-moving	There is no forecast demand over the next 12 months (12 months is an example, because it is commonly used).

and supplier constraints can limit the order quantity. Calculation of order variance over the horizon length requires record of order quantities in each period. ORV at any location (store or warehouse) gives a performance indication of that location. Overall variance can ignore the variances at individual locations, and therefore, cannot be properly interpreted. ORV is better suited to retail supply chains because of its fast moving nature.

Inventory quality ratio (IQR) (Pettey, 2019) is a KPI that reflects the inventory performance with respect to the pre-defined standards of the organization. It is the ratio of total value of "active" inventory to the value of total inventory held at any time. Some inventory terms are defined next. *Active inventory* is defined as the inventory whose days-of-supply (DOS) is under the defined target. Anything above the target is called *excess inventory*. IQR is an inventory performance measure that evaluates current inventory value against future demand in terms of usage value. Under IQR, the categorization of inventory is presented in Table 3.1.

Computation of IQR requires the pre-defined desired DOS levels for each item. Two different levels are widely used: Target DOS and minimum DOS. After these two elements are defined, IQR can be calculated based on in-hand inventory at any time. Apart from on hand inventory record, standard unit cost of each item, forecast horizon and demand forecast information are also required.

IQR is a performance measure with respect to own a business' standards. Hence, the standards should not be chosen arbitrarily, which may render the IQR calculation useless. A higher IQR is considered better during comparison. IQR can be computed for a whole supply chain if such inventory targets are defined. However, it may not give a clear idea about where to focus for improvement. A granular computation at warehouse or store levels can give a better actionable picture. As IQR considers a group of items together, it cannot be computed for a single item. As IQR needs clear inventory targets, it is usually difficult to compute in retail. However, for manufacturing and spare part supply chains it can be well defined.

3.4.3. Service KPIs

Under service KPIs, we list three KPIs: α-service level (ASL), β-service level (BSL) and availability (AVL).

α-Service level (ASL) (also called Type-1 or Cycle service level) (Beyer *et al.*, 2016) is the probability of satisfying all demand during the concerned period. This service level represents the frequency of out-of-stock situations without any regard to the shortage quantities. In backorder environments $\alpha-$service level can be referred as "on time in full (OTIF)" indicating proportion of instances where demand is completely fulfilled.

In practice, α-service level can be computed approximately from the number of shortages. The ratio of the number of shortages to the total number of days gives the α-service level. When two IMSs are compared, the system that generates higher α is considered better. However, a service level measure cannot be analyzed independently as the marginal cost of higher service level increases with the increase in service level. α-service level is applicable for all supply chains.

Table 3.2: Summary of KPIs

Caregoty	KPI	Reference	Usefulness
Financial	Value chain profitability (VCP)	Proposed	Retail
	Inventory turnover (ITO)	Rao and Rao (2009)	Retail, spare parts, industrials
	Tolal inventory value (TIV)	General	Retail, spare parts, industrials
Operational	Inventory level (IVL)	General	Retail, spare parts, industrials
	Inventory variance (IVV)	Petropoulos *et al.* (2019)	Retail, spare parts, industrials
	Order quantity variance (ORV)	Cannella and Ciancimino (2010)	Retail, spare parts, industrials
	Inventory quality rario (IQR)	Pettey (2019)	Retail, spare parts, industrials
Service	α-service level (ASL)	Beyer *et al.* (2016)	Retail
	β-service level (BSL)	Beyer *et al.* (2016)	Spare parts, industrials
	Availability (AVL)	Van Houtum and Kranenburg (2015)	Spare parts

β-Service level (BSL) (also called Type-2 service level or fill rate) (Beyer *et al.*, 2016) can be defined as the proportion of demand fulfilled. Contrary to α-service level, it considers the actual shortage quantities.

In practice, computation of β-service level requires information about the real demand. However, real demand is accurately known only if it is recorded before fulfillment. Otherwise obtaining such information is hard. During comparison of two IMSs, the system generating higher β is considered better. As β-service level uses real demand information, it is more suited for manufacturing and spare parts supply chains.

Availability (AVL) (Van Houtum and Kranenburg, 2015) is another measure of service quality that defines the proportion of all items having non-zero inventory during a particular period. Generally, availability is determined for fast-moving items in spare parts supply chains. This measure, however, does not concern the actual quantity of inventory. Mathematically availability

Computation of availability needs precise definition of the concerned set of items. Without that, it may be counter productive. That is, an increase in availability can be detrimental to overall performance (specifically the total cost). After defining the set of items the calculation of availability becomes simpler. It requires record of inventory at the beginning of each period. During comparison, a higher availability is considered better. This KPI is more suited for spare parts supply chains.

Strategically, sometimes, the storage locations for different items may vary. For example, in case of a very low demand (slow-moving) items, it is better to store them at a central location than at each retailer to increase availability. Although the retailer availability might be low but, the overall availability can be equivalent at a lower inventory cost. It is called inventory pooling. Availability is widely applicable in spare parts supply chains. In retail it can be defined for class-A items. In manufacturing supply chains it is generally not suited.

We summarize the KPIs discussed in the previous section in Table 3.2. Next, we discuss the challenges involved in computing those KPIs in practice and propose a simulation methodology.

3.5. Computation Methodology

In this section, we discuss the second aspect of performance assessment, i.e., computation methods for the selected KPIs. Computation methods also hold key importance as only accurate and relevant computations would result in an appropriate action plan.

Comparison of performance can be of two types depending on the systems to be compared. It can be between an existing IMS and a new IMS, or between two new IMSs. The former comparison can be made using AB testing (Gilotte *et al.*, 2018) and regular simulations. AB testing refers to a small scale implementation of the new IMS at selected locations and then monitoring the performance for a certain duration (can span upto a year). Regular simulations are not suitable in such situations because of difficulties in exact replication of past events. While in the second case both systems can be compared using regular simulation techniques. The previously defined KPIs can be used to assess performance. We propose a simulation method that we call the Δ-method for the former case.

Δ-Method

The proposed method is applicable for situations where there is already an existing IMS and the organization wants to adopt a new one. In such situations AB testing can take longer and regular simulation cannot ensure an identical testbed since the existing system cannot be simulated. One of the major data points missed while simulating the new system with past data is the real demand. Without it, the shortages cannot be estimated.

Significance of Δ: In the absence of information about real demand regular simulations are not well-suited. Here, Δ comes into the picture, and that enables us to perform an accurate "relative comparison" of two systems without knowing the real absolute performances. This also mitigates the possibility of any selection bias in favor of an existing system while estimating the real demand.

Let us first define ι as the assumed proportion of shortage during a specific period. The idea here is to find out "if the existing system is unable to fulfill ι proportion of the demand,

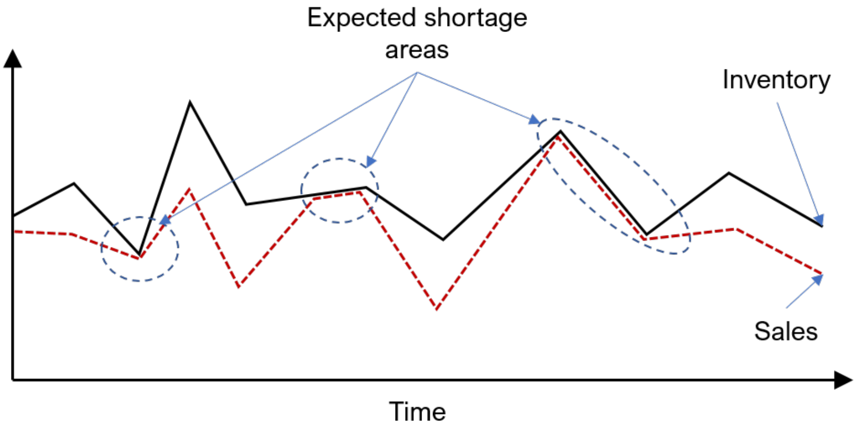


Figure 3.7: Shortages are expected where inventory and sales are close.

then what proportion of the demand will be left unfulfilled by the new system?" Let V_T be the total sales during a horizon of length T. If the existing system has lost ι portion, then the demand will be $\frac{V_T}{1-\iota}$. Our assumption is if the inventory position is comparatively higher than the real sales quantity then there is no shortage (see Figure 3.7). But, if the inventory position is very close or equal to the real sales quantity then some demand is lost. Now, we do not know exactly how much demand is not satisfied during that time, but we know the time and we know over a horizon how much demand is not satisfied.

Let Δ_t be a random positive value for time t. If the inventory position is comparatively higher than the sales at time t then there is no lost sale and we set $\Delta_t = 0$. When the inventory position is close or equal the sales quantity there are some shortages, i.e., $\Delta_t > 0$. Δ_t is chosen such that the sum of Δ_t over the horizon is approximately equal to the total shortages. As the values of Δ_t are generated randomly, we require enough samples of the demand curve for a conclusion. The same process can be repeated for different values of ι and the relative performance can be assessed. The step-by-step simulation is described next and illustrated in Figure 3.8.

1. Choose a suitable starting period in the past and a horizon length (should not be less that 60 days).
2. Collect the past sales and other required features for generating forecast for the selected horizon.
3. Collect the values of parameters which are necessary for running the replenishment planning system.
4. Run the forecasting system and replenishment planning system to get the optimal order quantities.
5. Once the order quantity is known, calculate the total delivery quantity.
6. Update the inventory position.
7. Initialize the demand process based on the Δ-method. Draw a demand sample from the underlying distribution. If the demand received is higher that the inventory position then the excess demand unfulfilled. Otherwise complete demand is satisfied.
8. Compute the closing inventory. For a higher initial inventory position we end up with positive inventory. On the other hand, a lower that demand initial inventory position

will lead to zero closing inventory position for lost sales case and negative inventory position for backorder case.

9. This closing inventory becomes the opening inventory for the next period. Repeat steps 2-8 until the end of the horizon.
10. Compute the KPIs based on the recorded data points.

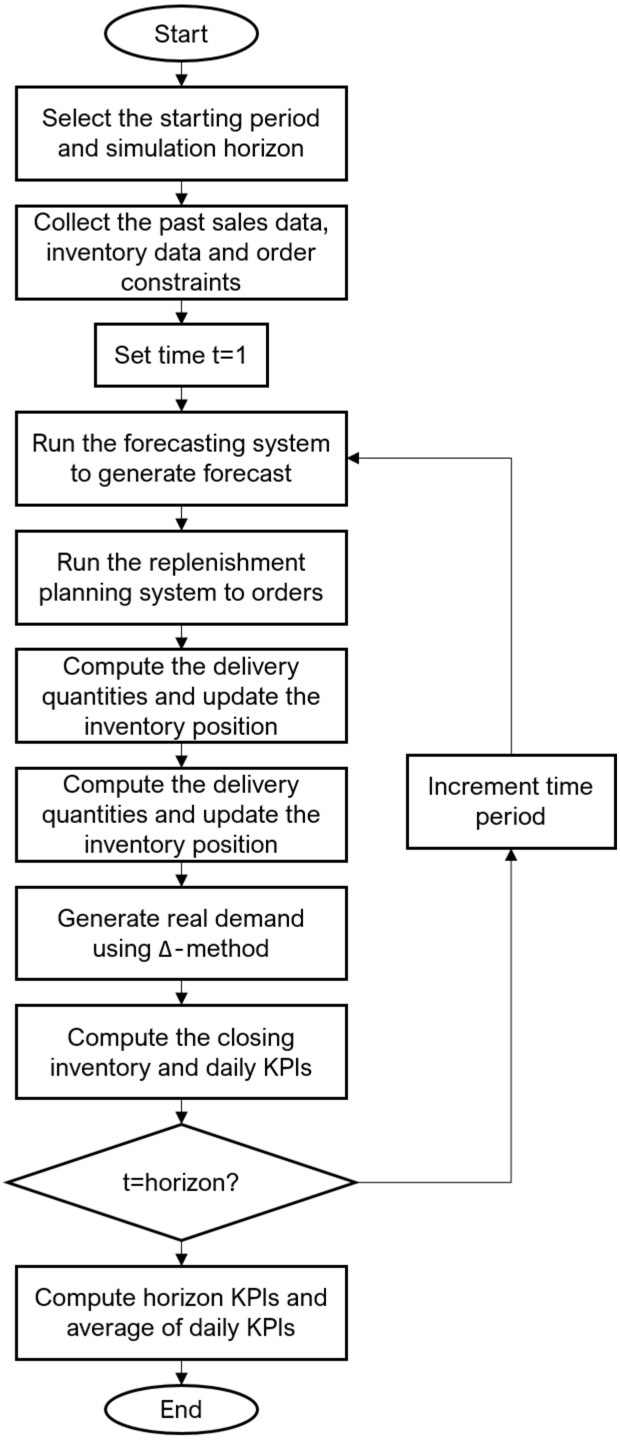


Figure 3.8: Flowchart for the proposed simulation method.

3.6. Conclusions

This chapter has two major aspects. First, we reviewed and proposed KPIs for inventory management systems. Those KPIs reflect the performance of both forecasting and replenishment planning systems together. Those KPIs are classified under three major categories: financial, operational and service. Under financial KPIs, we described the value chain profitability (VCP), total inventory value (TIV) and inventory turnover (ITO). The inventory level (IVL), inventory variance (IVV), order variance (ORV) and inventory quality ratio (IQR) were described under operational KPIs. Service KPIs included α-service level (ASL), β-service level (BSL) and availability (AVL).

Secondly, we also analyzed the challenges while computing the above KPIs in real-world scenarios. While two new IMSs can be compared using regular simulation techniques, assessing a new system with respect to an existing one comes with two major limitations. An accurate method for such situations, AB testing requires longer duration to obtain results. In most cases it can span upto a year. Moreover, this method also has financial risks since it is actually implemented. On the other hand regular simulations do not ensure an identical testbed. We proposed a simulation method that we call the $\Delta-$method. It is capable of evaluating KPIs of existing and new systems under identical conditions in the past. This method does not require lengthy computation horizon as AB testing. Also, this method does not have a risk of financial losses.

This chapter concludes the first part of this dissertation. In the second part we focus on developing optimization methods for replenishment planning problems. In the next chapter, we propose a sampling-based method for the single-item replenishment planning problem.

Part II

Problem Definitions and Solution Methods

Chapter 4

Problem Definitions and Modular Framework

Abstract: *In this chapter, we define various inventory optimization problems that are to be addressed by Vekia. Their solution methods are proposed later in the dissertation. Tactical and operational problems are broadly divided into two categories: single-item replenishment planning problems and multi-item (joint) replenishment planning problems. They can be further sub-divided based on the length of planning horizon (and the corresponding number of replenishment decisions), i.e. single period, multi-period (definite periods) and long-term. We also discuss the replenishment problems during promotional events and supplier selection aspect. Extensions of the basic problems are also proposed to address practical situations. All the above problems are arranged as a modular inventory management framework, with each module addressing a particular variety of problems. The framework also includes strategic multi-echelon inventory optimization and other long-term planning problems.*

4.1. Introduction

Inventory optimization problems have been extensively studied in the past. In Chapter 2, we proposed a classification scheme for various inventory optimization problems. Using that scheme, we also defined the problems addressed in this dissertation (see Table 2.3). Those problems are of two broad classes based on the number of items involved in planning: problems with only one item and problems with multiple items. Depending upon the situation, the planning horizon can span just one period or multiple periods. Therefore, formally, the broad categories of the problems are as followed.

1. Optimization problems involving a single item.
   - (A) Multi-period single-item replenishment planning (SRP).
   - (B) Multi-period single-item replenishment planning with supplier selection (SRPSS).
2. Optimization problems involving multiple items.
   - (A) Single period promotional joint replenishment planning (PJRP).
   - (B) Multi-period joint replenishment problems (JRP).
   - (C) Multi-period joint replenishment problems with supplier selection (JRPSS).

In this dissertation, we consider a general multi-echelon supply chain. We assume a decentralized approach of managing the inventory, where each location in the supply chain can be controlled independently. Each location encounters one or more of the previously mentioned problems. They are the decisions in operational level and require to be solved frequently. In order to be efficient on global cost parameter as well, a central multi-echelon optimization and a long term optimal parameter setting problem are also proposed. The framework is presented in Figure 4.2.

The rest of this chapter is arranged as follows. In Section 4.2, we arrange various types of relevant inventory optimization problems in a modular inventory management framework, explain its different modules and their interrelationship. The problems corresponding to the modules are defined next in Section 4.3 and Section 4.4. Then we analyze possible industrial extensions in Section 4.5. We provide the conclusions in Section 4.6.

4.2. Solution Design Framework

Depending on the organization, the type of supply chain and the type of inventory optimization problem can vary. From a solution design point of view, a modular approach can be configured to address most of the problems with each module addressing a sub-problem. In this section, first we will explain the general nature of those underlying problems then arrange them into a modular framework.

4.2.1. The Supply Chain Network and Planning Problems

A general supply chain can have multiple levels of storage. In Figure 4.1, a typical supply chain network is depicted where each location (manufacturing, warehouse or retail store) can serve as a point to store inventory.

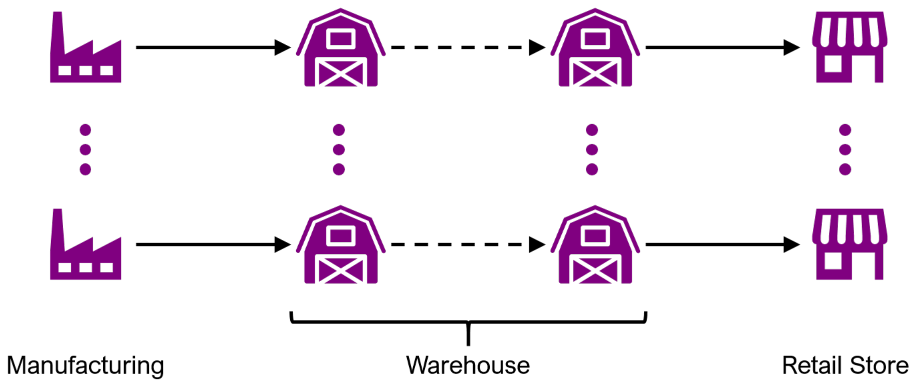


Figure 4.1: A generalized depiction of multi-echelon supply chain.

During day-to-day operations, replenishment process can be defined as the operation of each location to order items from its predecessors (suppliers). Generally those decisions are taken considering the future demand, in-hand inventory level and the operational constraints. The standardized planning involved in this process is called replenishment planning. The common method is as follows. In the beginning of each time period, the manager takes note of the inventory in hand. He generates the demand forecasts for future periods (definite). He then places an order at the supplier(s) considering the operational constraints and the costs. Those orders are delivered immediately. At the end of each time period, inventory holding cost is paid on the closing inventory and for any shortage, the shortage penalty is paid. For every order places a fixed cost is also paid.

4.2.2. Modular Inventory Management Framework

The process described before is a generalized replenishment planning process. The actual planning requires the details of the underlying problems. The proposed inventory management framework has three categories of problems: strategic planning, operational planning and emergency planning. Figure 4.2 illustrates the framework with each module represented by a rectangle. The shaded rectangles are the problems that are addressed during this research. The work on the problem in the hashed rectangle is ongoing and remaining problems are planned for future research.

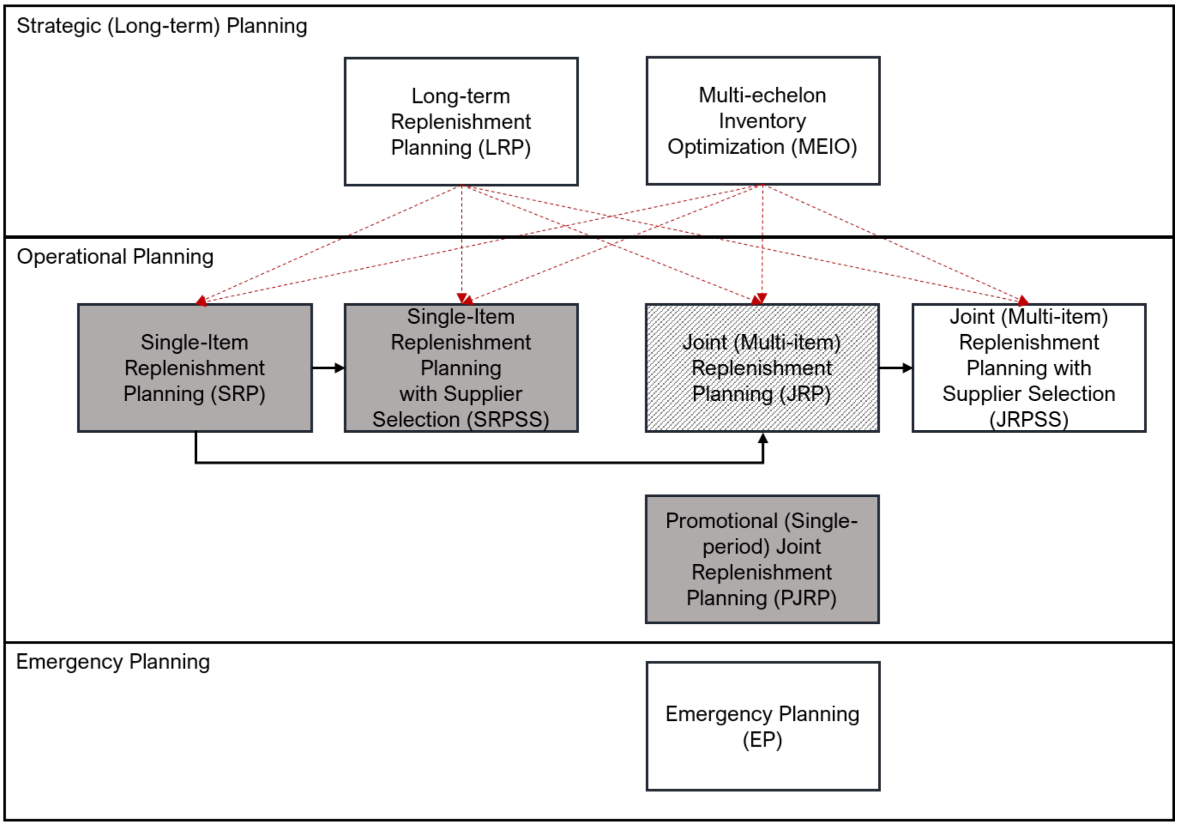


Figure 4.2: Proposed inventory management framework.

Out of the three categories of problems, we focused only on the operational planning part. Those problems may be divided into two categories:

1. Single-item problems and
2. Multi-item problems.

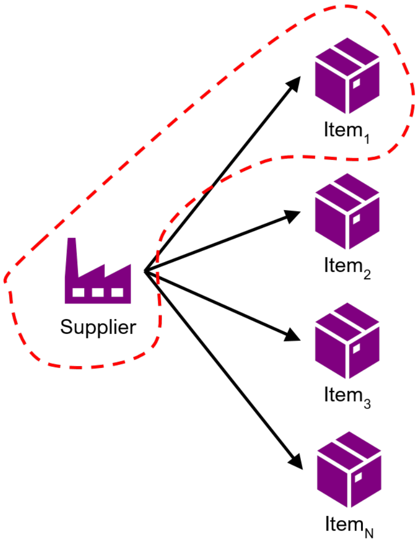
Two types of single-item problems are addressed during this research. The first one is the basic version, single-item replenishment planning (SRP) is the well known inventory control problem where orders are optimized for one item without taking any supplier aspects into account. Cost components include the inventory holding cost, shortage cost and fixed order cost. When supplier selection is also a part of the problem, we define it as single-item replenishment planning with supplier selection (SRPSS). In addition to previous cost parameters, supplier selection also requires supplier specific costs such as unit purchase price, supplier specific fixed costs, etc.

Multi-item problems differ from the single-item ones due to the presence of first degree interaction between the items. We address three types of multi-item problems. The basic version with multiple items and a common fixed cost called the joint replenishment planning (JRP). The second one called the joint replenishment planning with supplier selection (JRPSS). It is an extension of JRP to include supplier selection. The third one is a single-period problem, that is encountered frequently by retailers during promotions called the promotional joint replenishment problem (PJRP).

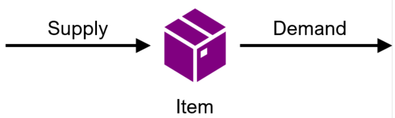
In Figure 4.2, arrows suggest a generalized relationship from solution development point of view. The SRP is a sub-problem of the SRPSS and the JRP. The PJRP is a sub-problem of the JRP and the JRP is a sub-problem of JRPSS. In the upcoming section we explain each of the above problems in more detail.

4.3. Single-Item Inventory Optimization

The inventory control problem for the single-item case is one of the most studied problems in inventory optimization literature. As the name suggests, the ordering decisions for just one product are optimized. Even though very few practical situations deal with only one item, in most multi-item cases, the whole problem can be assumed to be an aggregation of independent problems concerning each item.



(a) Depiction of the complete problem where multiple items are ordered from a single supplier.



(b) The problem in Figure 4.3a is separable for each item in case of SRPs.

Figure 4.3: Single supplier single item replenishment problem.

4.3.1. Single-item Replenishment Problem (SRP)

In Figure 4.3a, a common case of multiple items ordered from a single supplier is depicted. The replenishment problems for each item can be considered independent and can be formulated and optimized separately. In Figure 4.3b, independent control of each item is depicted.

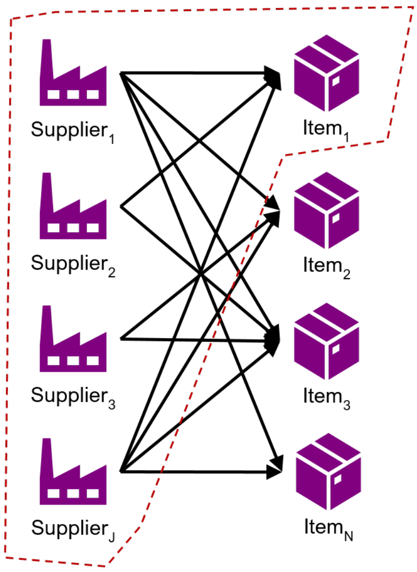
The basic problem can be defined in the following way. Each location plans for several items (in some cases that can exceed one million) to be ordered. For the sake of simplicity, we consider that the global constraints such as total budget and supplier capacity are not violated and we are able to plan independently for each entity. Without any common constraint linking the items, the replenishment problems can be separable for each item. At the beginning of a time period, the inventory manager decides the replenishment quantity and the time. For an item at a given store, the forecast is updated at the beginning of each period. The forecast information is available up to the end of the planning horizon. The ordering decisions are made considering the forecasts during the rolling horizon, in-hand inventory level and the associated cost parameters. When an order is placed, the quantity ordered is delivered immediately. Any excess demand is fully backordered i.e. any unsatisfied demand appears as additional demand of the next period. A fixed order cost is paid for each non-zero order placed. Inventory holding cost is paid for each unit of inventory carried from one period to another. For any backorder, a backorder penalty per unit is also paid.

The basic problem has stationary demand and unit order batch size without any additional supplier constraints. It can be denoted by the notations introduced in Chapter 2 as $\mathtt{1};\phi;\phi||\mathtt{O};\mathtt{D}||\mathtt{S};\mathtt{B};\phi||\mathtt{C}||\mathtt{D};\mathtt{U}||\mathtt{S};\mathtt{S};\mathtt{M}$. However, the supplier can also impose a batch size constraint. Additional sources of uncertainty may include lead time and inventory record accuracy. This problem can be extended to include other practical constraints, such as, minimum order quantity, minimum inventory quantity and service level, etc. During shortage, the associated demand could be lost completely (as in case of retail supply chains). There can also be a mix of backorder and lost sale in case of shortage. Non-stationary demand is another practical scenario. In this dissertation, we also differentiate the SRP problem based on number of suppliers as it requires additional cost parameters. Further details about the problem are discussed in the upcoming section.

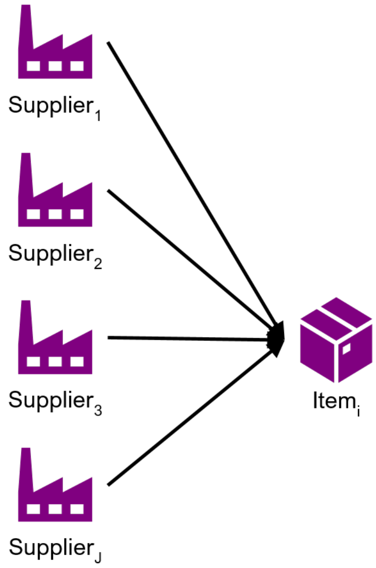
4.3.2. Single-item Replenishment Problem with Supplier Selection (SRPSS)

Previous single supplier model is simple with only one set of cost parameters. However, usually buyers adopt multiple-supplier strategy. This is mainly to mitigate uncertainty or to respect existing service level agreements. An inventory model with multiple suppliers is depicted in Figure 4.4. Retailers plan their inventory with a short review period. Availability of multiple suppliers for the same item poses greater challenges for cost effective operation. Since the total cost (and thereby the profitability) is closely related to the purchase price of an item, an integrated planning method becomes essential. The decision regarding supplier selection must be taken with due consideration to its cost implications. Capacity limitations and service level agreements (SLA) are some additional factors that make multi-supplier planning models essential.

Multi-supplier models in retail can have two basic supply chain network structures. First, a single echelon supply chain with multiple retailers and multiple suppliers. The second type of network structure has multiple retailers at different places along with a central warehouse. The central warehouse in turn, orders from multiple suppliers. We can generalize it to a single entity (retailer or warehouse) having multiple suppliers. Each entity has independent ordering strategy. The process for one entity is presented in Figure 4.4a. The ordering decisions for different items can also be taken independently. This results in each item being able to be ordered from multiple suppliers (see Figure 4.4b). The problem is called a single-item replenishment problem with supplier selection (SRPSS). For the basic problem, suppliers differ by the price they charge per unit item and fixed cost charged per order. Similar to the single-item case, the costs incurred are fixed ordering cost, inventory cost and shortage



(a) Depiction of the complete problem where multiple items are ordered from multiple suppliers.



(b) The problem in Figure 4.4a is separable for each item, and each item can be ordered from multiple suppliers.

Figure 4.4: Multi-supplier single item replenishment problem.

cost, with the additional purchase costs. Any order placed by the entity to any supplier is delivered immediately.

The basic SRPSS can be denoted as $1;\phi;\phi||O;D||S;B;\phi||C||D;U||S;M;M$. Extensions to the basic SRPSS are similar to that of the basic SRP. Different shortage consideration and non-stationary demand are item specific extensions. However, supplier specific extensions such as different batch sizes and different minimum order quantities can also be adopted.

4.4. Multi-Item Inventory Optimization

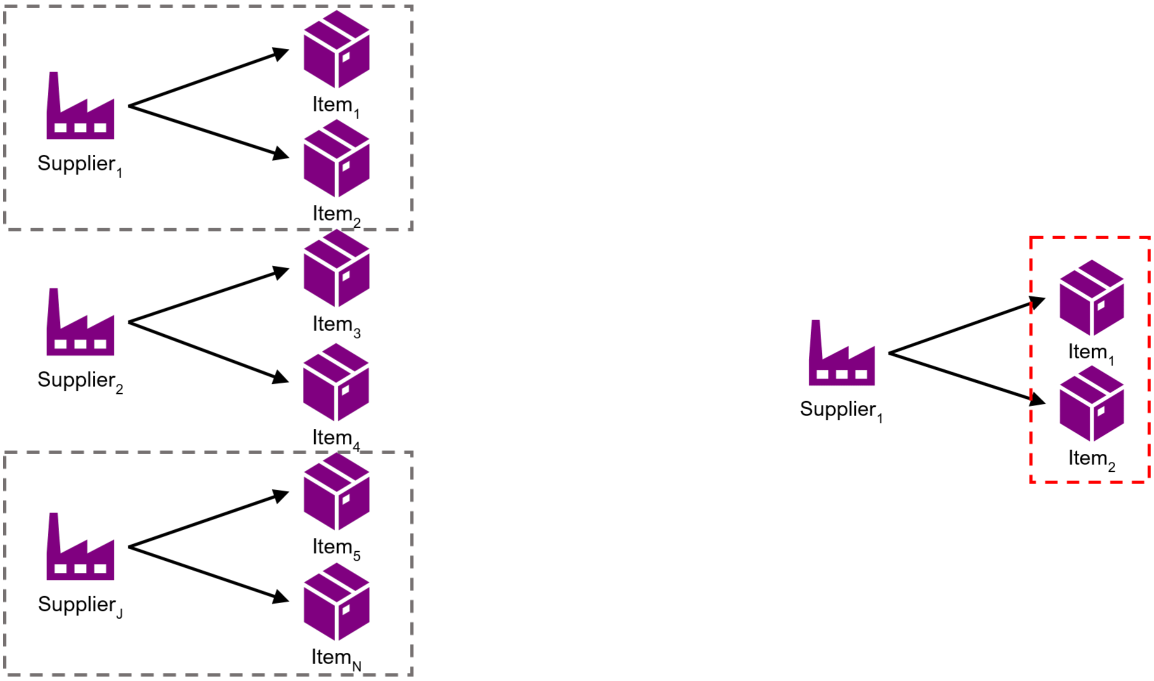
In practice, most inventory optimization problems concern multiple items. However, unlike the previously discussed problems, if the concerned items have a first order interaction between themselves, their control should not be performed independently. First order interaction takes place between items if one or more than one of the following conditions exist(s): common fixed ordering cost (when the fixed order cost is independent of the quantities or the variety of items in an order), common transportation means, common transportation cost, discounts on total purchase amount of all items or on the total quantity (this is also applicable for a class of items), etc.

We first define the joint replenishment problem (JRP) and present its extensions. Then, we define the joint replenishment problem with supplier selection (JRPSS). At last, we propose a promotional ordering problem involving multiple items, i.e. promotional joint replenishment problem (PJRP).

4.4.1. Joint Replenishment Problem (JRP) with Single Supplier

The joint replenishment problem (JRP) consists of integrated inventory control of multiple items which are connected through a joint fixed cost or a joint transportation mode. In general, a location orders multiple items from multiple suppliers. The classical JRP is separable for each supplier when the items do not overlap (see Fig 4.5a). The problem for each supplier is depicted in Fig 4.5b. The supplier charges a fixed cost for each order irrespective of its contents. Inventory is reviewed periodically and order is placed considering the future demands and in-hand inventory positions. Inventory holding costs are charged on the closing inventory of each period and any shortages incur shortage penalties. For the basic problem, the demand for different items are stochastic and stationary. Batch size is one and all the cost parameters are linear.

The basic JRP can be denoted as $1;\phi;\phi||O;D||S;B;\phi||C||D;U||J;S;M$. Extensions include uncertain lead times, inventory quantities and supplier delivery quantities, non-stationary demand, non-unit batch size, presence of minimum order quantities and different shortage consideration (backorder and lost sale).



(a) Depiction of the complete problem where multiple items are ordered from multiple suppliers.

(b) The problem in Figure 4.5a is separable for suppliers, and each supplier provides a group items.

Figure 4.5: Joint replenishment problem with a single supplier.

4.4.2. Joint Replenishment Problem with Supplier Selection (JRPSS)

The joint replenishment problem can be generalized to have multiple suppliers, defined as joint replenishment problem with supplier selection (JRPSS). For a location, the JRPSS is the most complex form of inventory control. The problem arises when the entity has multiple suppliers for multiple items and the items across the supplier overlap. In other words one item has multiple suppliers and the supplier can supply multiple items.

Figure 4.6 depicts the general structure of a joint replenishment problem with multiple suppliers at a location. The location encounters stochastic demands for multiple items. The inventory manager reviews the inventory level of each item at the end of each period, then, looking at the demands ahead, he places order at the suppliers. The orders are delivered immediately. Similar to the JRP the inventory holding costs and the shortage costs are incurred. Other cost parameters are supplier specific: fixed order costs and unit price per item. Also, supplier specific constraints include batch size and minimum order quantity. The basic JRPSS can be denoted as $\texttt{1;}\phi\texttt{;}\phi||\texttt{O;D}||\texttt{S;B;}\phi||\texttt{C}||\texttt{D;U}||\texttt{J;M;M}$.

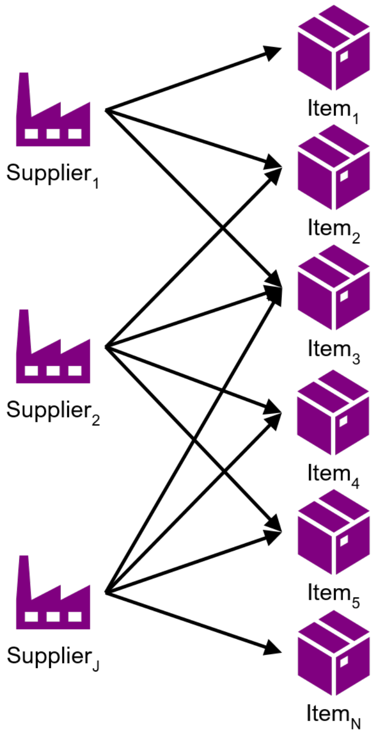


Figure 4.6: Joint replenishment planning with supplier selection.

4.4.3. Promotional Joint Replenishment Problem (PJRP)

Nowadays, promotional events to boost sales are becoming common. Particularly in retail, its effect is more evident. Inventory planning during a promotional event involves a new set of items that are to be sold in the promotion and usually only one order is placed for them. Therefore, the problem is similar to the well known newsvendor problem (single-period ordering) with multiple items. In this dissertation we do not consider a joint constraint such as net budget or volume, however, our major focus is on dealing with the presence of multi-item pack (*a prepack*). The problem is briefly defined below.

We study a single echelon supply chain network with multiple entities (retailers). The inventory control processes at those entities are considered independent since no global budget constraint is present. The organization follows a decentralized planning approach, where, each entity plans its orders. In practice the entities seldom has the central financial data. This makes planning with financial parameters inviable. There are multiple suppliers, however, none has any capacity constraint. So, we consider them as one supplier supplying all the items. Each item can only be ordered as the multiple of its batch size (*a pack*) as defined by the supplier. There are two distinct type of packs (See Fig 4.7). First, single-item packs, where, the pack contains only one type of item. Second, multi-item packs, where, the pack contains multiple items of supplier defined quantities (equal or unequal). Order for an item can contain either single-item packs or multi-item pack or a mix. The basic PJRP can be denoted as `1;`ϕ`;`ϕ`||0;D||S;L;`ϕ`||C||D;S||J;S;S`.

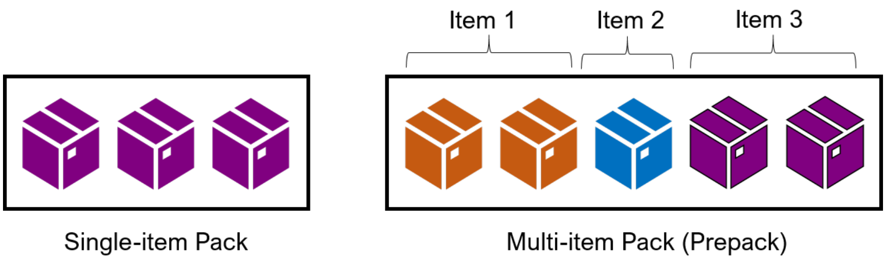


Figure 4.7: Illustration of differentiation between single-item pack and multi-item pack during promotional ordering.

4.5. Industrial Extensions

In reality, inventory optimization problems are very different from those solved in the literature and briefly presented above. They vary by the type of input parameters, type of constraints and application condition. There are some subjective aspects as well. For example, a managers intuition about demand instead of machine generated forecast and the ease of implementation of an outcome also impact the solution methods.

For both single-item and multi-item inventory optimization problems, common industrial extensions are presence of lost sale, a mix of lost sale and backorder during shortage, service level constraint, tackling forecast inaccuracy, multiple batch sizes, non-stationary demand, supplier discounts, capacity constraints, etc.

4.6. Conclusions

In this chapter, we discussed various problems encountered at Vekia, and most of them are addressed in this dissertation. The problems were arranged to form a modular inventory management framework. It can fit to many problem classes during solution development. Broadly, two classes of problems were discussed: single-item inventory optimization problems and multi-item inventory optimization problems (joint replenishment problems). For single-item problems, we focused only on multi-period problems. Further, they were classified into single-supplier and multi-supplier problems. Similarly, multi-item problems are classified into single-period (promotional), multi-period with single supplier and multi-period with multiple suppliers. All the above problem classes were defined for a single location (or retailer) as we consider them to be independent. We also discussed their possible extensions.

In the next chapter, we propose a sampling-based optimization method for the basic SRP problem. The method can be extendable to include additional problem parameters and practical constraints.

Chapter 5

Sampling-based Replenishment Planning in Single-Item Inventory Systems

Abstract: *In this chapter, we propose an innovative and practical replenishment planning method for the single item inventory system. The method is distribution-free and uses samples of future demand. The corresponding replenishment decisions are implemented in a rolling horizon manner. We first investigate two different mathematical models for the replenishment planning problem with backorder: expected cost approach, robust cost approach and propose a novel approach, namely, coverage period cost approach. Then, we formulate an exact algorithm and also propose an approximation heuristic for the immediate expected cost approach. Numerical analyses on synthetic and benchmark datasets show that the proposed heuristic performs well, both in terms of solution quality and runtime. It can be used for both stationary and non-stationary demand distributions. Various extensions are also proposed to reflect its readiness for practical application. Those extensions include batch-size constraint, higher realized demand uncertainty and minimum order quantity, etc.*

5.1. Introduction

The single-item replenishment planning (SRP) problem is one of the most widely studied inventory optimization problems in the literature. This has great strategic significance for both distribution and retail firms. While the initial studies focused on deriving solutions for the *standard* problem, recent literature gives importance to real-world problem parameters and execution time. As discussed in Chapter 2, this type of problem can have many variants due to changes in the network structure, temporal aspects, consideration of demand, shortage, etc. This large variety of problems can be present within one organization. An inventory management system must be able to cope with such situation. Dedicated frameworks for each of these problems might yield better result individually, however, this approach may pose challenges during solution development and servicing.

Since in a real-world setting, situations having deterministic demand are scarce, we focus solely on the problems assuming the demand to be stochastic. Replenishment planning with stochastic parameters falls under the class of multi-stage stochastic problems (MSSP) (Homem-de Mello and Bayraksan, 2014). For stationary demand and full backorder, the seminal paper by Scarf (1959) proves that the optimal inventory policy is of (s, S) type where S denotes the order up to level and s denotes the reorder level. However, the practical applicability of a (s, S) type inventory policy is severely limited by its simple application conditions. Real-world conditions are different from what is regularly assumed in the literature: non-stationary demand, non-parametric demand distribution, order batch size and incomplete demand information to name a few. The non-stationary (s, S) policy (Xiang *et al.*, 2018) is computationally expensive as it requires repeated computations of policy parameters. The comparative measures of various well-known methods for stochastic replenishment planning are given in Table 5.1. Both optimal and approximate methods suffer from either very limited application scope or high computation time.

Table 5.1: A comparative representation of application conditions of various replenishment planning methods in single-item inventory systems and their qualitative performance indicators. NS(s, S) Policy: Non-stationary (s, S) policy with policy parameters determined by a heuristic, DP: Dynamic Programming, Offline Time: Computation time for policy parameters, Online Time: Computation time for actual orders from the policy or the method.

Method	Demand Distribution Type			Batch Size	Offline Time (Online Time)	Optimality
	Stationary/ Non-stationary	Parametric/ Non-parametric	Independent/ Non-independent			
(s,S)Policy	Stationary	Parametric	Independent	Unit	Medium(Low)	Optimal
(R,Q)Policy	Stationary	Parametric	Independent	Any	Medium(Low)	Approximate
(s,Q)Policy	Stationary	Parametric	Independent	Any	Medium(Low)	Approximate
(s,nQ)Policy	Stationary	Parametric	Independent	Any	Medium(Low)	Approximate
NS(s,S)Policy	Both	Parametric	Independent	Unit	Medium(Low)	Approximate
DP	Both	Both	Both	Any	-(Very High)	Optimal

The solution methods for any problem under the class of MSSP depend on the availability of information. They can be classified into two broad categories: probabilistic approach and scenario-based approach (Pan and Nagi, 2010). The probabilistic approach can be used when the complete probability distributions of the underlying uncertain parameters are available. An example of probabilistic approach is stochastic programming. Even in the presence of multiple sources of uncertainty, stochastic programming is an effective tool to find the optimal solution. If the complete probability distribution is not known, the scenario-based approach can be adopted. The scenario-based approach is a distribution-free method that involves characterizing the uncertainty by a set of scenarios that represent a number of potential future states. The sampling-based approach (Levi *et al.*, 2006) can be considered as a variant of the scenario-based approach. During the course of this dissertation we propose such approaches for replenishment planning.

In this chapter, we study the multi-period stochastic inventory system involving a single product and propose a near-optimal replenishment planning methodology. Afterwards, we also propose various extensions for different application conditions. Our contributions in this regard are the following:

1. We examine two mathematical models: expected cost approach and robust cost approach, and propose a novel approach, namely, coverage period cost (CPC) approach for the replenishment planning problem in a stochastic inventory system using samples.

2. We propose a heuristic for *online* computation of order quantities with much less computational effort than dynamic programming or policy evaluation.

3. We propose various extensions to the above methodology to address batch-size, higher realized uncertainty and minimum order quantity.

The rest of this chapter is organized in the following way. In Section 5.2, we briefly discuss the context of the problem and our motivation behind the problem and the methodology.

Then, we provide a brief review of related literature in Section 5.3. Some preliminary concepts are discussed in Section 5.4. The mathematical models are presented in Section 5.5 and the main heuristics are given in Section 5.6. While Section 5.7 presents the experimental protocol, in Section 5.8, we provide the results of our experiments. Extensions to the main heuristics and their results are presented in Section 5.9. At last, in Section 5.10, we provide the conclusions and discuss some future perspectives.

5.2. Context and Motivations

In this section, we provide a brief description of the problem and discuss our motivations behind using a sampling-based approach.

5.2.1. Context

We study a single-echelon supply chain where each store is controlled independently. This decentralized planning may not be the most cost efficient approach, however, from a practical point of view, it is more manageable. It is also near-optimal if the localized solutions are coherent with the global objectives. For simplicity, we consider that the global constraints such as total budget and supplier capacity are not violated and we are able to plan independently for each store. In each store, there are several items (in some cases that can exceed one million) to be ordered and the ordering process for each item is also independent. This constitutes multiple SSRPs. The specifics of the underlying processes are as follows.

For an item at a given retailer, the forecast is updated at the beginning of each period. The forecast is generated up to the end of the planning horizon. The ordering decisions are made considering the demand forecasts during the rolling horizon (see Section 5.4), current inventory level and the associated cost parameters. When an order is placed, the quantity ordered is delivered immediately. Any excess demand is fully backordered, i.e., any unsatisfied demand appears as additional demand of the next period. A fixed order cost is paid for each non-zero order placed. Inventory holding cost is paid for each unit of inventory carried from one period to another. For any backorder, a backorder penalty per unit is also paid. The supplier can also impose a batch size constraint. We consider the external demand to be the only source of uncertainty. Using the classification scheme from Chapter 2, the base problem can be defined as:
$\texttt{1;}\phi\texttt{;}\phi||\texttt{O;D}||\texttt{S;B;}\phi||\texttt{C}||\texttt{P;U}||\texttt{S;S;M}$. This basic structure can have extensions, such as with lost sales, single batch size, deterministic or stochastic lead time. For example:
$\texttt{1;}\phi\texttt{;}\phi||\texttt{O;D}||\texttt{S;}\boldsymbol{L}\texttt{;}\phi||\texttt{C}||\texttt{P;U}||\texttt{S;S;M}$
$\texttt{1;}\phi\texttt{;}\phi||\texttt{O;D}||\texttt{S;B;}\phi||\texttt{C}||\texttt{P;}\boldsymbol{S}||\texttt{S;S;M}$
$\texttt{1;}\phi\texttt{;}\phi||\boldsymbol{D}\texttt{;D}||\texttt{S;B;}\phi||\texttt{C}||\texttt{P;U}||\texttt{S;S;M}$.

5.2.2. Motivations

Even if the SRP problem has been studied widely in the literature, none of the existing methods can be universally applied or easily adaptable to the real-world conditions. In Table 5.1, we have summarized the limitations of existing well-known methods.

Solving the SRP problem using traditional methods comes with two limitations. First, due to the curse of dimensionality (Defourny *et al.*, 2012) the MSSPs are hard to solve when the number of stages is large (usually for ≥ 3). On the other hand, a smaller number of stages can yield myopic results (Rahdar *et al.*, 2018). Non-stationary (s, S) policies require dynamic computation of the policy parameters (which is computationally expensive), parametric and independent demand distributions. For any MSSP, it often becomes very difficult to enumerate all the possible outcomes with increasing number of stages. This, as a prerequisite for the computation of the expectations in stochastic programming, makes the approach difficult to adopt. In such cases, *sampling techniques* are natural tools to use (Shapiro, 2003). The other motivation for adopting a sampling-based method comes from their ease of use in practical applications. Standard forecasting methods are capable of producing probabilistic forecasts. However, handling multi-dimensional distributions is very hard in parametric forms and some forecasting methods naturally yield samples. Therefore, direct use of multi-dimensional samples addresses both of the above concerns.

5.3. Related Literature

Related literature comprises two major research fields. First is stochastic inventory optimization, and second is sample-based optimization. In this section, we discuss separately the research already done in these two topics that are pertinent to our problem. Since our work focuses on sample-based inventory optimization, we review the literature on this topic as well.

5.3.1. Stochastic Inventory Optimization

Single item inventory optimization under stochastic demand has been studied since the 1950s. The earliest literature in the field (Scarf, 1959; Clark and Scarf, 1960; Karlin, 1960; Veinott, 1966) study the structure of optimal policy. Interestingly, the structure of the optimal policy is very simple, i.e. of (s, S) type. The policy defines two inventory levels: the reorder level s, and the order up to level S. When the inventory level reaches or goes below the reorder level an order is placed to raise the inventory position to S. The optimal parameters can be found by using policy iteration, but it is time consuming. Simpler methods to compute the stationary (s, S) policy are given by Zheng and Federgruen (1991) and Feng and Xiao (2000). Similarly for the non-stationary (s, S) policy Askin (1981) and Bollapragada and Morton (1999) provide simple heuristic. More recently Xiang *et al.* (2018) proposed a mixed integer linear programming approach to compute the approximate non-stationary (s, S) policy. The heuristic provided therein, although computationally less expensive than dynamic programming, is still not suitable for real-world size problems.

Both the cases of stationary and non-stationary demand rely on the complete knowledge of the underlying probability distribution. Although the non-stationary case is more appealing for practical applications, the methodologies developed in this area can only provide myopic policies at a reasonable computation time. Other practical limitation arises due to the presence of batch-size. Widely used policies in the presence of batch-size, such as (R, Q), (R, nQ) and (s, nQ) (Q = Batch-size, s, R = Reorder points) are themselves not optimal.

As discussed in the motivations, under practical assumptions, the MSSP suffers from the curse of dimensionality (Defourny *et al.*, 2012). Bertsimas and Thiele (2006), Levi *et al.* (2007a), Neely and Huang (2010) and Rahdar *et al.* (2018) proposed different methods to avoid the problem and approximate the solution. Also, Rossi *et al.* (2015) and Özen *et al.* (2012) explore static dynamic uncertainty strategy to address the problem. However, all of the methods mentioned above assume the demand to be independent across time. Our proposed method provides an approximate solution to the MSSP while making no assumptions regarding stationarity or independence of underlying demand distributions. It can also be extended to include the batch size constraint and to address more uncertain demand realizations. Moreover, runtime increases linearly with the length of the planing horizon. Applicability of our proposed method is summarized in Table 5.2.

5.3.2. Sampling-Based Stochastic Optimization

As the total number of outcomes increases, sample-based stochastic optimization becomes more practical. A detailed survey of Monte-Carlo sample-based methods for stochastic optimization can be found in Homem-de Mello and Bayraksan (2014). This method has been successfully used in many practical problems, for example vehicle routing (Kenyon and Morton, 2003; Verweij *et al.*, 2003), engineering design (Royset and Polak, 2004), supply chain network design (Santoso *et al.*, 2005), power generation and transmission (Jirutitijaroen and Singh, 2008), asset liability management (Hilli *et al.*, 2007), and machine learning (Byrd *et al.*, 2011).

Approaches for sample-based stochastic optimization include *sample average approximation* (SAA) and *stochastic approximation* (SA). SAA works by replacing the stochastic programming problem by its sample approximation, which is a deterministic optimization problem. The SA approach on the other hand works on sequential approximation strategy, where approximation and sample generation are carried out alternatively. For MSSPs like stochastic inventory optimization, Shapiro (2003) outlined the need for conditional sampling. This makes the number of samples required for the MSSP grow exponentially with the number of periods (stages). Some methods to address this issue are dual dynamic programming (Shapiro, 2011), and ReSa (Reduced Sampling) (Hindsberger and Philpott, 2014).

Table 5.2: Applicability and qualitative performance of the proposed coverage period cost (CPC) approach.

Parameter	Sub-Parameter	Options	Applicability	Remarks
Demand	Temporal change	Stationary	✓	
		Non-stationary	✓	Extension
	Distribution	Parametric	✓	
		Non-parametric	✓	
	Temporal dependency	Independent	✓	
		Dependent	✓	
Batch size	Size	Unit	✓	
		Non-unit	✓	Extension
	Numbers allowed	Single	✓	
		Multiple	✓	Extension
Shortage	–	Backorder	✓	
		Lost sale	✓	Extension
Uncertainty	–	Equal	✓	
		Increased	✓	Extension
MOQ	–	–	✓	Extension
Runtime	–	Online	–	Low
		Offline		NA

5.3.3. Sampling-Based Stochastic Inventory Optimization

In most of real-life scenarios the explicit demand distributions are hardly known or very complex to work with (Levi *et al.*, 2006). Under such assumptions, sampling-based methods are suitable. Sampling-based methods have been used for stochastic inventory optimization in two contexts. First for the single-period (newsvendor) problem and second for the multi-period problem. The SAA approach performs well for the single-period problem (Homem-de Mello and Bayraksan, 2014). Levi *et al.* (2006) and Levi *et al.* (2007b) are some of the initial works on distribution free approaches to address inventory optimization problems. Levi *et al.* (2007b) addressed both the newsvendor problem and its multi-period extension. For the multi-period model they proposed a dynamic programming algorithm that gives approximate base stock policies. They assumed the demand to be independently (not necessarily identically) distributed. For faster runtime they also proposed a myopic method that only evaluates the optimal cost for one period. Even though this method is computationally simple, the solution quality can be far from the optimal.

Halman (2015) provided a fully polynomial time approximation scheme for stochastic dynamic program with samples. Their method can also be applied to approximate the solutions given by Levi *et al.* (2006) and Levi *et al.* (2007b). In order to deal with industrial size problems, our analysis differs in two aspects. First, instead of a multi-stage problem we analyze the cost of the first period only. Second, we provide a close-form approximation and the computational complexity of the final algorithm is of the order equal to the number of stages. Moreover, we also focus on solution flexibility so that it can be easily extended to include additional practical constraints. The applicability of the proposed approach is summarized in Table 5.2. The base problem can be defined as the single-item replenishment problem with full backorder unit batch size. The proposed method can be applied to non-parametric demand as well. For non-stationary demand, lost sales, non-unit batch size, multiple batch size and increased uncertainty (inaccurate forecast), extensions can be used. The detail applicability of our proposed method is given in Table 5.2.

5.4. Preliminaries

Before going into the detailed formulations, we discuss some preliminaries. All of our proposed methods are implemented in a rolling horizon approach, in which demand for the duration of the rolling horizon is considered for ordering decisions instead of the whole planning horizon. We elaborate the rolling horizon method in the upcoming paragraph. In addition, we also present the notations and dynamic programming formulation of the concerned problem.

Rolling Horizon: A rolling horizon approach (Rahdar *et al.*, 2018) is necessary primarily because of two reasons: the scarcity of demand information far in the future and the curse of dimensionality. These are two concerns that must be considered while designing a stochastic inventory optimization solution. While the former affects the solution quality, the latter affects the runtime required to find a solution. When demand upto a longer horizon is used, it yields good quality solutions. However, due to scarcity and inaccuracy of such information as well as the curse of dimensionality, a compromise can be made if the decisions are implemented in a rolling horizon manner. Under this approach, demand information up to T (rolling horizon length) periods is available. The total length of planning horizon $\hat{T}$ can extend up to infinity. The decision is taken in each period considering the demand up to T periods and initial inventory level s_0. Depending upon the adopted method multiple ordering decisions may be evaluated during each period, but only the first one is implemented. The process is then repeated with updated forecast for the next period and so on. This way the rolling horizon approach can be utilized when the demand distributions are dependent. The benefits of this approach include exploitation of the future demand information and reduction in problem size. The rolling horizon approach is depicted in Figure 5.1.

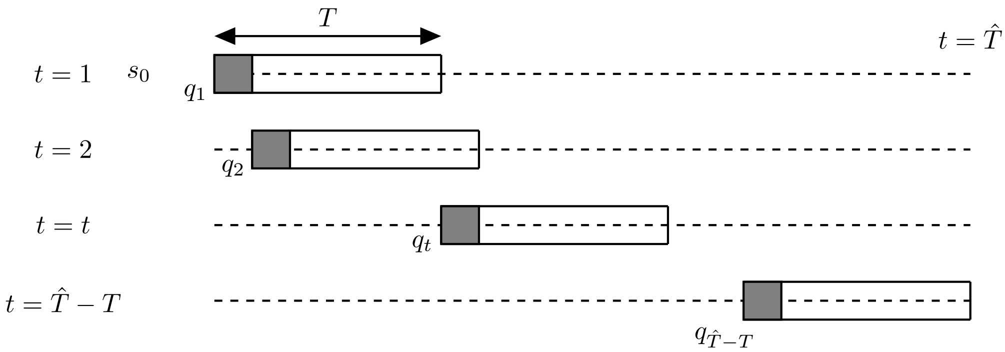


Figure 5.1: Illustration of an ordering mechanism with a rolling horizon length T and planning horizon length $\hat{T}$. During each time period t order quantities are determined considering the opening inventory level and demand over a t to $t+T$ window.

Dynamic Programming: The multi-period stochastic inventory optimization problem with rolling horizon approach is a multi-stage stochastic problem that can be solved optimally with dynamic programming. This also requires the end state to be known and demands across time to be independent. The infinite horizon cost calculation is proposed in Scarf (1959). The structure of the obtained *policy* is simple only if the underlying demand distribution is a parametric one. In case of non-independent distributions it becomes very hard to express them and to find an optimal solution. The dynamic program is as follows (refer to table 5.3 for notations). Let $\mathcal{C}_t(s_{t-1})$ be the expected total cost of an optimal policy over time periods t to T for initial inventory position s_{t-1}. Then the functional equation of the dynamic program will be

$$\begin{aligned}\mathcal{C}_t(s_{t-1}) = \min_{q_t} \Big\{ \mathbb{E}\big[H(s_{t-1}+q_t-d_t)^+ + W[-s_{t-1}-q_t+d_t]^+ + K\alpha_t + \\ \mathcal{C}_{t+1}(s_{t-1}+q_t-d_t)\big] \Big\}\end{aligned} \tag{5.1}$$

5.5. Sampling-based Optimization Models

Although, solving the above dynamic program can give us the desired result, in practice it is computationally prohibitive. In this section, we investigate various optimization models that use probabilistic forecast as samples. The first model minimizes the expected cost for the horizon over the set of samples. The second model minimizes the maximum cost of the horizon over different samples in the sample set. Finally, we propose a mathematical formulation for the expected cost during the first period and propose to minimize this expected cost to obtain near-optimal order quantities. We use a dynamic programming approach to compare the results of the previously stated methods.

Forecasts as Samples: For an item i, we consider N samples and each sample $j \in \{1..N\}$ is a demand *trace* that can be represented by a vector $\mathbf{x}_j$ of size T, where T is the length of the rolling horizon. Figure 5.2 depicts a set of three samples with $T = 10$. The samples drawn can be from any underlying distribution as the optimization method does not assume any specific property.

Table 5.3: Parameters and variables for the optimization models.

Sets	
$\mathcal{Z}$	Set of samples, index for sample $z \in \{1, ..., Z\}$
$\mathcal{T}$	Set of time periods, index for periods $t \in \{1, ..., T\}$
Parameters	
P_z	Occurrence probability of sample z, $\in [0, 1]$
F_{zt}	Demand forecast for sample z at time t, $\in \mathbb{R}^+$
K	Fixed ordering cost per order, $\in \mathbb{R}^+$
H	Inventory holding cost per unit, per unit time period, $\in \mathbb{R}^+$
W	Backorder cost per unit, $\in \mathbb{R}^+$
A	Discount factor, $\in (0, 1]$
$\mathbf{M}$	A suitably large positive number
Decision Variables	
q_t	Order quantity at time t, $\in \mathbb{Z}^+$
Internal Variables	
d_t	Random demand variable at time period t, $\in \mathbb{Z}^+$
s_{zt}	Inventory for sample z at the end of period t, $\in \mathbb{R}$
l_{zt}	Backorder quantity for sample z at time t, $\in \mathbb{R}^+$
c_z	Total cost of sample z, $\in \mathbb{R}^+$
α_t	Binary, 1 if fixed cost is charged at time t and 0 otherwise $\in \mathbb{Z}^+$
m	Worst case cost $\in \mathbb{R}^+$

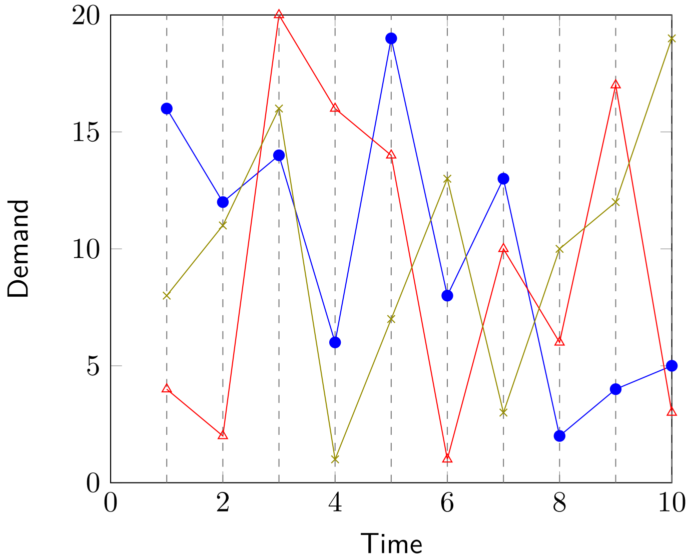


Figure 5.2: Illustration of demand samples. We present three distinct demand samples, which are random realizations of the demand during time $t = 1$ to $T = 10$.

5.5.1. Expected Cost Approach (ECA)

The ECA aims to minimize the expected cost of ordering decision $\mathbf{q}$ across all samples. Here, we consider the problem from a static uncertainty strategy and evaluate a set of decisions along time. The optimization model is as follows.

$$\text{minimize} \sum_{z=1}^{Z} P_z c_z \tag{5.2}$$

subject to

$$s_{z0} = s_0 \qquad \forall z \in \mathcal{Z} \tag{5.3}$$

$$s_{zt} \geq s_{z0} + \sum_{x=1}^{t} q_x - \sum_{x=1}^{t} F_{zx} \qquad \forall z \in \mathcal{Z}, t \in \mathcal{T} \tag{5.4}$$

$$l_{zt} \geq \sum_{x=1}^{t} F_{zx} - \sum_{x=1}^{t} q_x - s_{z0} \qquad \forall z \in \mathcal{Z}, t \in \mathcal{T} \tag{5.5}$$

$$q_t \leq \alpha_t \mathbf{M} \qquad \forall t \in \mathcal{T} \tag{5.6}$$

$$c_z = \sum_{t=1}^{T} (\alpha_t K + s_{zt} H + l_{zt} W) A^t \qquad \forall z \in \mathcal{Z} \tag{5.7}$$

$$s_{zt} \geq 0 \qquad \forall z \in \mathcal{Z}, t \in \mathcal{T} \tag{5.8}$$

$$l_{zt} \geq 0 \qquad \forall z \in \mathcal{Z}, t \in \mathcal{T} \tag{5.9}$$

Constraint (5.3) refers to the initial state of inventory. Constraint (5.4) computes the positive inventory quantity during each period and constraint (5.5) computes the backorder quantity. When the unmet demand is fully backordered, it appears as the additional demand in the following period. Therefore, it is necessary to consider the cumulative demand and orders and not only the ones for the current period. Constraint (5.6) ensures the binary indicator α_t to be equal to one for any non-zero order quantity. The individual sample costs are calculated as per constraint (5.7). Constraints (5.8) and (5.9) are positivity constraints.

5.5.2. Robust Cost Approach (RCA)

The RCA is similar to the ECA except it aims to minimize the maximum cost across samples. Formally,

$$\text{minimize } m \tag{5.10}$$

subject to (5.3)-(5.9) and

$$m \geq P_z c_z \qquad \forall z \in \mathcal{Z} \tag{5.11}$$

An additional constraint (5.11) is used to determine the maximum (worst case) cost m across samples.

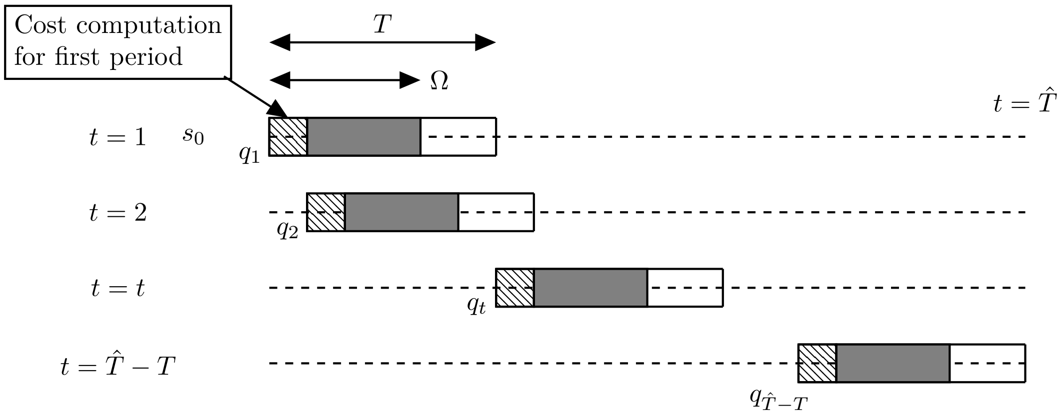


Figure 5.3: Illustration of an ordering mechanism with the IECA.

5.5.3. Immediate Expected Cost Approach (IECA)

We propose an approach based on the notion of coverage period cost (CPC). We define the *coverage period* as the time horizon between two consecutive orders. The CPC is defined as the expected cost incurred during the coverage period. CPC included inventory holding cost, shortage cost (backorder or lost sale cost) and fixed order cost. Let us consider a case with the opening inventory equal to s_0. The future demand from $t = T^1$ to $t = T^2$ is expressed in terms of samples. The choice of length of horizon T depends upon the accuracy level of forecasts as well as the quality of required results. Given that all samples constitute only non-negative values of demand, we can determine the inventory traces for the future starting with an inventory level s_0. The resulting process for $T^1 = 1$ and $T^2 = 10$ is depicted in Figure 5.4. For each trace $z \in \mathcal{Z}$, we can determine the projected inventory values. Therefore, it is possible to calculate the inventory quantities and the backorder quantities for each sample and each order quantity. Furthermore, this enables us to calculate the cost for any coverage period T^1 to T^2. We aim to find out the order quantity that minimizes the expected cost for the first period taking in that coverage period. We denote the expected cost during the coverage period as $\tilde{C}(s_0, q, T^1, T^2)$ name it the coverage period cost (CPC). The minimum coverage period cost (MCPC) of the horizon $t = T^1$ to $t = T^2$ is equal to $\tilde{C}^*(s_0, q, T^1, T^2)$. Mathematically,

$$\begin{aligned}\tilde{C}(s_0, q, T^1, T^2) = \sum_{z=1}^{Z} P_z \Big[\sum_{t=T^1}^{t=T^2} \Big(H\,[s_0 + q - \sum_{\tau=1}^{t} F_{z\tau}]^+ \\ + W\,[\sum_{\tau=1}^{t} F_{z\tau} - s_0 - q]^+ \Big) + K\alpha \Big]\end{aligned} \tag{5.12}$$

$$\alpha = \begin{cases} 1 \text{ , if } q > 0 \\ 0 \text{ , otherwise} \end{cases} \tag{5.13}$$

$$\begin{aligned}\tilde{C}^*(s_0, q, T^1, T^2) = \min_{q} \sum_{z=1}^{Z} P_j \Big[\sum_{t=T^1}^{t=T^2} \Big(H\,[s_0 + q - \sum_{\tau=1}^{t} F_{z\tau}]^+ \\ + W\,[\sum_{\tau=1}^{t} F_{z\tau} - s_0 - q]^+ \Big) + K\alpha \Big]\end{aligned} \tag{5.14}$$

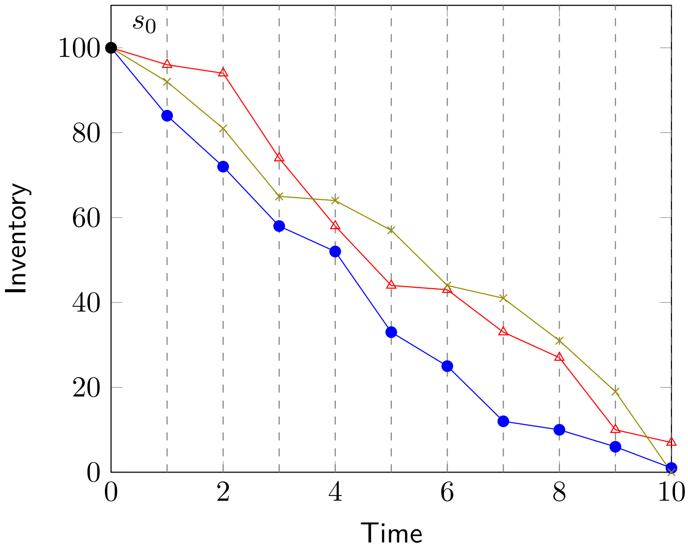


Figure 5.4: Illustration of changes in inventory quantities for the different forecast samples as depicted in Figure 5.2.

We propose an immediate expected cost (IEC) defined as the portion of coverage period cost incurred during the period immediately after the order. Since we follow a static-dynamic uncertainty strategy, we have one ordering epoch at the beginning of each time period. The corresponding order size may be zero during the evaluation, however, there is a possibility of ordering. Therefore, we propose to minimize the IEC as it is the cost that is impacted the most and cannot be changed once the time period has passed. We call this approach as immediate expected cost approach (IECA). In order to evaluate the IEC, let us consider the ordering process at $t = 1$ for a coverage period length Ω. If an order of quantity q is placed and delivered at the beginning of $t = 1$ and initial inventory is s_0 then, the total cost during the coverage period, for $\Omega \in \{1, 2, 3, ..., T\}$, will be equal to

$$\tilde{C}(s_0, q, 1, \Omega) = \sum_{z=1}^{Z} P_z \Big[\sum_{t=1}^{\Omega} \Big(H[s_0 + q - \sum_{\tau=1}^{t} F_{z\tau}]^+ + W[\sum_{\tau=1}^{t} F_{z\tau} - s_0 - q]^+ \Big) + K\alpha \Big] \quad (5.15)$$

$$\alpha = \begin{cases} 1 \text{ , if } q > 0 \\ 0 \text{ , otherwise} \end{cases} \quad (5.16)$$

We will denote $\tilde{C}(s_0, q, 1, \Omega)$ by $\tilde{C}_\Omega(s_0, q)$ for simplicity. As per definition, the IEC is the cost incurred during period $t = 1$. We differentiate two scenarios. Under stationary demand scenario, the demand distribution remains the same through out the rolling horizon. The cost during the period $t = 1$ will be equal to the average per period cost during the rolling horizon. Mathematically, it will be equal to

$$D_\Omega(s_0, q) = \frac{1}{\Omega} \tilde{C}_\Omega(s_0, q) \quad (5.17)$$

When the demand is non-stationary we approximate the expected cost for the first period as

$$D_\Omega(s_0, q) = \sum_{z=1}^{N} P_z \left[\frac{1}{\Omega} \sum_{t=1}^{\Omega} \left(H[s_0 + q - \sum_{\tau=1}^{t} F_{z\tau}]^+ + W[\sum_{\tau=1}^{t} F_{z\tau} - s_0 - q]^+ \right) + \mathcal{R}_\Omega K\alpha \right] \tag{5.18}$$

$$\text{where } \mathcal{R}_\Omega = \frac{\sum_{i=1}^{N} F_{z1}}{\sum_{z=1}^{N} \sum_{t=1}^{\Omega} F_{zt}} \tag{5.19}$$

We propose to minimize $D_\Omega(s_0, q)$ over $\Omega \in \{1...T\}$ and $q \in \{0, 1, 2, ...\}$. Mathematically,

$$q^*, \Omega^* = \arg\min_{q,\Omega} D_\Omega(s_0, q) \tag{5.20}$$

As we adopt a rolling horizon approach, only the first order is placed. A new order can be placed in each time period. Therefore, the immediate time period that follows an order is the most affected one. The central idea is to favor this period by minimizing its expected cost.

Some observations from equations (5.15) and (5.18) are as follows. For any order quantity $q \geq 0$, $K\alpha$ is a constant. For any time period $t \in \{1, ..., \Omega\}$ the sum of inventory holding cost and backorder cost can be either positive or equal to zero. Therefore, for any q the total cost can either increase with increase in T or remain equal. Similarly, the nominator term of $\mathcal{R}_\Omega$ is a constant and the denominator term is non-decreasing in Ω. Therefore, $\mathcal{R}_\Omega$ is non-increasing in Ω for any q.

5.6. Heuristic Approaches

The ECA and RCA optimization models with short rolling horizon can be solved using any commercial solver. With longer rolling horizon, computations can be very time consuming. Since the mathematical formulation of the ECA and RCA follow static uncertainty strategy, their solutions are not cost-optimal. Heuristics for these models are desirable only if they have the potential to provide good quality solutions when compared with the global optimum. From the numerical analysis, however, the performances of ECA and RCA were found to be far from the optimum. Hence, we do not consider the heuristic approach for these models. On the contrary, we propose an enumerative search heuristic to solve (5.20) for the IECA model. In this section we explain the enumerative heuristic as well as propose an improved approximation heuristic for better computation time.

5.6.1. Enumerative Search Heuristic

The proposed enumerative search heuristic EH-IECA evaluates the cost function (5.17) for all possible combinations of q and Ω and selects the pair that gives the minimum cost. The heuristic is formally presented as Algorithm 1.

The complexity of Algorithm 1 is in the order of $\mathcal{O}(T|\mathcal{Q}|)$, where $\mathcal{Q}$ is the set of all possible order quantities i.e. $\{0, 1, ...q_{max}\}$. The choice of $\mathcal{Q}$ depends on the magnitude of the demand distribution. In the next sub-section, we will present an improved version of the algorithm based on the structural property of the cost function.

Algorithm 1: Enumerative Search Heuristic (EH-IECA)

Input : $\mathbf{F}, H, W, K, T, s_0$
Output: q^*
initialization $\Omega = 1, q_{\max} = \max_z \sum_{t=1}^{T} F_{zt}, q^* = 0, v' = \infty, v = \infty$
repeat
    Compute $D_\Omega(s_0, q)$, $\forall q \in \{0, 1, 2..., q_{max}\}$
    $v' \leftarrow \min_q D_\Omega(s_0, q)$
    if $v' < v$ **then**
        $v \leftarrow v'$
        $q^* \leftarrow \arg\min_{q \in \{0,1,2...,q_{max}\}} D_\Omega(s_0, q)$
    end
    $\Omega \leftarrow \Omega + 1$
until $\Omega = T$;

5.6.2. Convexity Results and Optimized Heuristic

Let us consider a sample $\mathbf{F}_z = \{F_{z1}, F_{z2}, ..., F_{zT}\}$ and a coverage period Ω. The total cost during $t = 1$ to $t = \Omega$ for any order quantity $q \geq 1$ is plotted in Figure 5.5. It represents the curves for different samples with initial inventory $s_0 = 0$. Intuitively we believe the cost function for any sample is convex and attains a minimum. Therefore, if the minimizing value of q can be found out for each $\Omega \in \{1, 2, ..., T\}$ then, the requirement of enumeration for each value of q in Algorithm 1 can be substantially reduced. In this subsection we shall prove that the cost function (5.17) is indeed convex for $q \geq 1$ for any $\Omega \in \{1, 2, ..., T\}$ and present an approximation method to find the minimizing q. We also present a modified heuristic based on these results called Optimized Heuristic-Expected Immediate Cost Approach (OH-IECA), whose pseudo-code is presented in Algorithm 2.

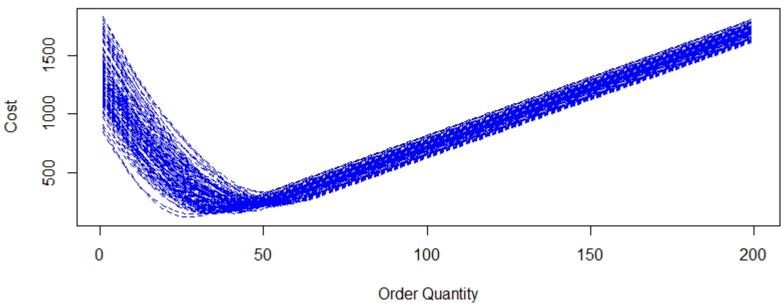


Figure 5.5: Illustration of cost as a function of order quantities for 100 randomly generated samples of length 10, mean 5, inventory holding cost 1 and backorder cost 5.

Proof of Convexity

For any sample z, let us represent the total cost for a coverage period Ω by $\Xi_z(s_0, \Omega, q)$, where s_0 is the initial inventory quantity and q is the order quantity. Recalling the total cost function (5.15), we get

$$\tilde{C}_\Omega(s_0, q) = \sum_{z=1}^{Z} P_z \Xi_z(s_0, \Omega, q) \tag{5.21}$$

where

$$\Xi_z(s_0,q,\Omega) = \sum_{t=1}^{\Omega}\left(H[s_0+q-\sum_{\tau=1}^{t}F_{z\tau}]^+ + W[\sum_{\tau=1}^{t}F_{z\tau}-s_0-q]^+\right)+K\alpha \tag{5.22}$$

$$= \sum_{t=1}^{\zeta_z}H\left(s_0+q-\sum_{\tau=1}^{t}F_{z\tau}\right)+\sum_{t=\zeta_z+1}^{\Omega}W\left(\sum_{\tau=1}^{t}F_{z\tau}-s_0-q\right)+K\alpha \tag{5.23}$$

$$= X(s_0,\Omega,q)+Y(s_0,\Omega,q)+K\alpha \tag{5.24}$$

In equation (5.23), $\zeta_j \in \{1,2,...\Omega\}$ is such that the cumulative demand up to ζ_j is less than or equal to s_0+q (i.e. demand satisfied) and from ζ_j+1 onwards backorder is encountered. In order to prove the convexity of $\Xi_z(s_0,q,\Omega)$ for $q \geq 1$, we consider the different terms of (5.23) separately. Let the first term be X and the second term be Y. X represents the total inventory holding cost and Y represents the total backorder cost. The third term $K\alpha$ is constant. For $q \geq 1$ it is equal to K, otherwise it is equal to 0. Therefore, for computational purposes we consider the cost function in two different zones. First for $q \geq 1$ and second for $q=0$. For the no order case ($q=0$), we will have a unique value of the cost function. In the following parts we prove the convexity of the cost functions for $q \geq 1$.

Since $\sum_{t=1}^{\zeta_z}F_{zt} \leq s_0+q$, X is always positive or zero. Therefore, any increase in q will augment the value of X. Similarly, $\sum_{t=1}^{\hat{\zeta}_j}F_{zt} \geq s_0+q$, for $\hat{\zeta}_z \in \{\zeta_z+1,...,\Omega\}$, Y is always positive. Therefore, any increase is q will either reduce Y or keep it unchanged.

Property 1: *X, Y and Ξ_z are piecewise linear convex functions in q for any s_0, Ω and $q>0$.*

Proof. See Appendix A ■

Corollary 1: *$\tilde{C}_\Omega(s_0,q)$ and $D_\Omega(s_0,q)$ are piecewise linear convex in q for any s_0, Ω and $q>0$.*

Proof. It is straightforward from (5.21). Since $\tilde{C}_\Omega(s_0,q)$ is the weighted sum of Ξ_z, and Ξ_z is convex $\forall z \in \mathcal{Z}$ from Property 1, it is also piecewise linear convex. $D_\Omega(s_0,q)$ is obtained by multiplying a positive real number, hence, it is also piecewise linear convex. ■

Optimized Heuristic

The upcoming analysis is conducted at constant s_0 and Ω unless otherwise stated. The minimizing ζ_z and q for Ξ_z are given in equation (5.25) and (5.26). Here, $\lceil x \rceil$ is the smallest integer greater than or equal to x. See Appendix B for explanation of (5.25).

$$\zeta_z^* = \arg\min_{\zeta}\Xi_z(s_0,\Omega,q) = \left\lceil \frac{W\Omega}{H+W}\right\rceil \tag{5.25}$$

$$q_z^* = \sum_{t=1}^{\zeta_z^*}F_{zt} \tag{5.26}$$

Since $D_\Omega(s_0,q)$ is also convex in $q>0$ for fixed Ω, we can find the minimizing q, i.e.

$$q_\Omega^* = \arg\min_q D_\Omega(s_0,q) \tag{5.27}$$

Algorithm 2: Optimized Heuristic (OH-IECA)

Input : $\mathbf{F}, H, W, K, T, s_0$
Output: q^*
Initialize $\Omega \leftarrow 1, u \leftarrow \infty, u' \leftarrow \infty, q^* \leftarrow 0$
repeat
  Compute q^*_Ω using (5.31)
  Compute Costs $D_\Omega(s_0, q^*_\Omega)$ and $D_\Omega(s_0, 0)$
  $u' \leftarrow \min_{q \in \{q^*_\Omega, 0\}}(D_\Omega(s_0, q))$
  if $u' < u$ **then**
    $u \leftarrow u'$
    $q^* \leftarrow \arg\min_{q \in \{q^*_\Omega, 0\}}(D_\Omega(s_0, q))$
  end
  $\Omega \leftarrow \Omega + 1$
until $\Omega = T$;

To find the value of q that minimizes $D_\Omega(s_0, q)$ for any s_0 and Ω, let us consider the following. From equation (5.25), as ζ^*_z does not depend on the demand for each period, all ζ^*_z are equal $\forall z \in \mathcal{Z}$. Therefore, due to convexity, the overall optimal order quantity q^* must lie between $\min_z q^*_z$ and $\max_z q^*_z$. Let $\mathfrak{q}_x(.)$ denote the x^{th} quantile of a given sample. Using the results for single sample, we propose to use the following approximations.

$$\zeta^* \approx \frac{W\Omega}{H+W} \tag{5.28}$$

$$\zeta^*_l = \left\lfloor \frac{W\Omega}{H+W} \right\rfloor \tag{5.29}$$

$$\zeta^*_h = \left\lceil \frac{W\Omega}{H+W} \right\rceil \tag{5.30}$$

$$q^*_\Omega \approx \mathfrak{q}_{\min(1,\ \zeta^* - \zeta^*_l)}(F_{z\zeta^*_h}) + \sum_{z=1}^{Z} \sum_{t=1}^{\zeta^*_l} P_z F_{zt} \tag{5.31}$$

It is noteworthy that the previous analysis is valid for $q \geq 1$. Since, for $q = 0$ the value of $K\alpha = 0$, convexity does not hold true. To address this, we also evaluate the immediate period cost if the order quantity is zero. The procedure followed is formalized as Algorithm 2 and depicted in Figure 5.6. The complexity of Algorithm 2 is of the order $\mathcal{O}(2T)$, and it does not depend anymore on the magnitude of the demand distribution.

5.7. Experimental Protocol

In this section, we present the experimental protocol for detailed numerical experiments using proposed formulations and heuristics. We propose two performance indicators whose values are determined through simulations during the experiments. The first one is the cost incurred during the rolling horizon, the second one is an approximation of the expected per-period cost in an infinite horizon setting. While assessing the performance of a general stochastic inventory optimization problem, the demand can be considered to be stationary or non-stationary. For a stationary consideration, we can evaluate the total expected cost for a problem with finite horizon and the per period expected cost with infinite horizon (at least approximately) through simulations. However, for non-stationary demand infinite horizon cost computation (which will approach infinity) is impossible due to unavailability of demand information.

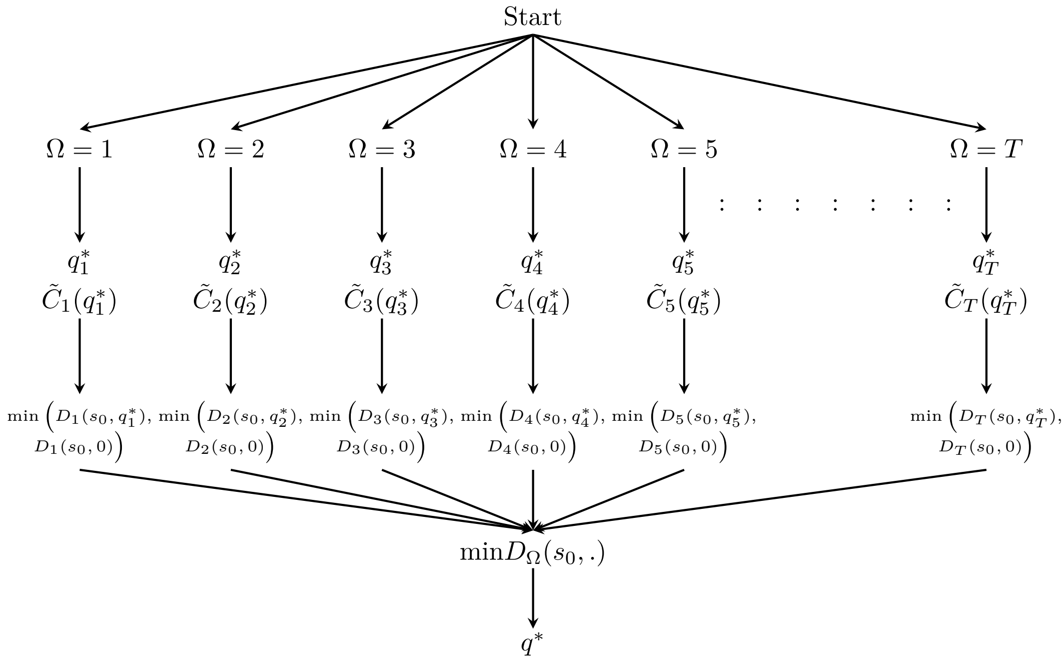


Figure 5.6: Steps of the OH-IECA. For different values of Ω, the optimal order quantity and the optimal cost can be determined using (5.31) and (5.15) respectively. Then the order quantity with the minimum immediate period cost (5.17) is chosen.

.

In Section 5.5, we have presented three different approaches based on demand samples for the SRP problem. The expected cost approach (ECA) is an adaptation of the SAA approach with static uncertainty strategy. The robust cost approach (RCA) is a min-max approach for robust optimization. The proposed immediate expected cost approach with optimized heuristic (OH-IECA) minimizes only the expected cost for the first period.

Our application conditions can be explained in the following way. At an ordering period t, sales information upto period $t-1$ is available. Future demand information up to T periods are then generated in terms of samples. However, the end state after T periods is not known. DP provides the optimal solution on an infinite horizon setting. On a finite horizon setting, it is optimal only if the end state is known or constrained. Therefore, in our analysis, we use its results as a base to compare the OH-IECA with the ECA and RCA. Our test problems also assume independently distributed demand to enable the use of DP approach. Then, we compare the best of them to the optimal policies of some benchmark problems in the literature. Following are the performance indicators used in our analysis.

1. *Horizon cost criterion* C^h: It is the expected total cost up to the end of the planning horizon. This can be obtained both for stationary and non-stationary demand cases.

2. *Infinite horizon cost criterion* C^∞: It is the expected cost per period. This can only be obtained for the stationary demand case in infinite horizon setting.

The procedure to compute C^h is as follows. The order quantity is decided as per the above four procedures (DP, ECA, RCA, OH-IECA). Then the demand is realized as per the forecast distribution (or from a different distribution). The costs of excess inventory or backorders are then computed. The inventory carried forward to the next time period is computed. This becomes the opening inventory for next period. This procedure is repeated until the end of the planning horizon. Although there can be infinitely many realizations, we repeat the whole process 10^6 times to obtain the expected cost. To compute C^∞, a similar procedure is followed but, instead of computing the cost until the end of rolling horizon, the ordering cycle is repeated until 10^6 time periods. Afterwards the average cost is determined. Let q_t^π be the order quantity under procedure $\pi \in \{\text{DP, ECA, RCA, OH-IECA}\}$ at the beginning of time t. The realized demand $\tilde{d}_t$ is drawn at random during each iteration from the underlying distribution.

$$C_\pi^h = \mathbb{E}\Big\{\sum_{t=1}^{T}(H(s_{t-1}^+ + q_t^\pi - \tilde{d}_t)^+ + W(-s_{t-1} - q_t^\pi + \tilde{d}_t)^+ + \alpha_t K)\Big\} \tag{5.32}$$

$$C_\pi^\infty = \lim_{\tau\to\infty}\frac{1}{\tau}\Big\{\sum_{t=1}^{\tau}(H(s_{t-1}^+ + q_t^\pi - \tilde{d}_t)^+ + W(-s_{t-1} - q_t^\pi + \tilde{d}_t)^+ + \alpha_t K)\Big\} \tag{5.33}$$

$$s_t = s_{t-1} + q_t^\pi - \tilde{d}_t \tag{5.34}$$

$$\alpha_t = \begin{cases} 1, \text{ if } q_t^\pi > 0 \\ 0, \text{ otherwise} \end{cases} \tag{5.35}$$

The problem instances are detailed in Table 5.4. There, H, W and K stand for the inventory holding cost, the backorder penalty cost and the fixed order cost respectively. For each problem instance, we assume independent demand distributions (Poisson) with arrival rate λ. We use CPLEX 12.7.1 to compute the optimal solutions in both the ECA and RCA models. All of the experiments are conducted on a Dell Latitude E5470 with 8GB memory.

Table 5.4: Problem instances for performance comparison.

H	W	K	λ
1	10	0	[5, 10, 15, 20, 50]
1	10	10	[5, 10, 15, 20, 50]
0.1	1	10	[5, 10, 15, 20, 50]

5.8. Results of Numerical Experiments

The detailed numerical results about the four proposed approaches are presented in this section. We conduct four different analyses. First, we compare the approaches by horizon cost criterion C^h and then by infinite horizon cost criterion C^∞. We also compare their runtime. Finally, we test the methods under several values of standard deviation of the realized demand.

5.8.1. Approximation Accuracy for the OH-IECA

In Section 5.6, we have presented an approximation Equation (5.31) for the optimal order quantity for the cost function (5.15). Before going into the detailed results, we present the approximation quality of equation (5.31). We use the excess cost over the optimal cost of Equation (5.15) as the evaluation criterion. Figure 5.7 presents the distribution of percentage excess cost in 10^5 random problem instances. On average the excess cost is 0.227% of the optimal cost. For 86.87% of the instances the excess cost is less than or equal to 0.50% of the optimal cost. The maximum percent excess cost was found to be 1.63%.

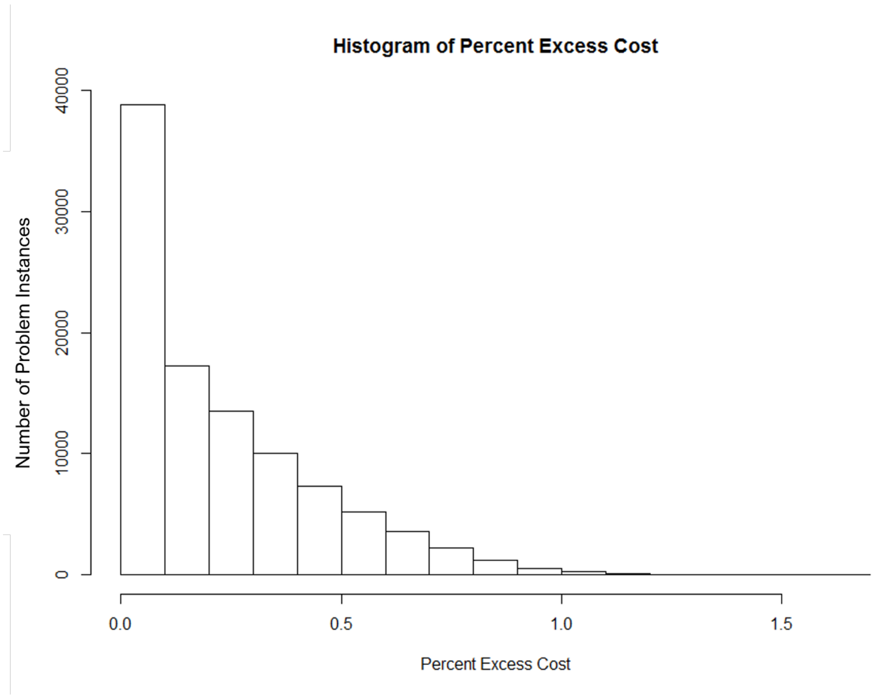


Figure 5.7: Illustration of distribution of percentage excess cost in 100,000 random problem instances. The average error is 0.227%.

5.8.2. Tests on Synthetic Problem Instances

The numerical results are divided into two parts. First, in Tables 5.5, 5.6, and 5.7, results are presented for tests with equal predicted and realized variance. Then we test the methods with higher actual variance, and the results are presented in Table 5.8. In the aforementioned tables, comparison is done by finding the ratio of the performance parameter using the concerned approach to the best performance parameter among all of the approaches for each problem instance, i.e. $\frac{C^h_\pi}{C^h_{min}}$ and $\frac{C^\infty_\pi}{C^\infty_{min}}$, where $C^h_{min} = \min_\pi C^h_\pi$ and $C^\infty_{min} = \min_\pi C^\infty_\pi$.

The horizon cost criterion C^h and the infinite horizon cost criterion C^∞ are presented in Tables 5.5 and 5.6 respectively. For the horizon cost criterion, the performances of DP and OH-IECA are very close. While the average ratio of $C^h_{OH-IECA}$ to the minimum C^h_π among the approaches is 1.014, the average ratio of C^h_{DP} to the minimum C^h_π among the approaches is 1.004. The maximum ratio of $C^h_{OH-IECA}$ to the minimum C^h_π is 1.076. Similarly, for the infinite horizon cost criterion C^∞_π, the solutions obtained using DP and OH-IECA are equivalent. Both of the approaches have average ratio of C^∞_π to the minimum C^∞_π is 1.002. The maximum ratio of $C^\infty_{OH-IECA}$ to the minimum C^∞_π among the approaches is 1.014.

Table 5.5: Comparative performance of different approaches for different problem instances for horizon cost criterion. (Ratio of C_{π}^{h} to C_{min}^{h}).

H	W	K	λ	$\frac{C_{DP}^{h}}{C_{min}^{h}}$	$\frac{C_{ECA}^{h}}{C_{min}^{h}}$	$\frac{C_{RCA}^{h}}{C_{min}^{h}}$	$\frac{C_{OH-IECA}^{h}}{C_{min}^{h}}$
1	10	0	5	**1.000**	1.109	1.109	1.001
			10	**1.000**	1.082	2.274	1.016
			15	**1.000**	1.185	1.052	1.003
			20	1.003	1.118	1.869	**1.000**
			50	**1.000**	1.067	2.049	1.002
1	10	10	5	**1.000**	1.040	1.013	1.003
			10	1.014	1.024	1.020	**1.000**
			15	1.002	1.080	1.022	**1.000**
			20	1.001	1.053	1.397	**1.000**
			50	**1.000**	1.038	1.594	1.001
0.1	1	10	5	**1.000**	1.102	1.020	1.076
			10	**1.000**	1.022	1.083	1.001
			15	**1.000**	1.027	1.031	1.052
			20	1.038	**1.000**	1.062	1.055
			50	1.004	1.010	1.211	**1.000**
			Average	**1.004**	1.063	1.320	1.014

In the above problems, DP does not produce the optimal results because of the absence of end conditions, *nevertheless, the DP results serve as useful baseline for comparison.* Under runtime parameter, the OH-IECA outperforms the other three approaches. While the average runtime for the OH-IECA is only 7 milliseconds, DP takes more than 3 minutes. The ECA and the RCA take about 2 to 3 seconds. Therefore, we obtain more than 99.99% reduction over DP in runtime with equivalent performance using the OH-IECA.

Table 5.6: Comparative performance of different approaches for the problem instances for infinite horizon cost criterion. (Ratio of C_{π}^{∞} to C_{min}^{∞}).

H	W	K	λ	$\frac{C_{DP}^{\infty}}{C_{min}^{\infty}}$	$\frac{C_{ECA}^{\infty}}{C_{min}^{\infty}}$	$\frac{C_{RCA}^{\infty}}{C_{min}^{\infty}}$	$\frac{C_{OH-IECA}^{\infty}}{C_{min}^{\infty}}$
1	10	0	5	1.002	1.109	1.109	**1.000**
			10	**1.000**	1.081	2.273	1.014
			15	**1.000**	1.187	1.052	1.002
			20	1.001	1.114	1.867	**1.000**
			50	**1.000**	1.066	2.052	1.003
1	10	10	5	**1.000**	1.056	1.014	1.009
			10	**1.000**	1.016	1.012	1.006
			15	**1.000**	1.078	1.022	1.001
			20	1.001	1.052	1.395	**1.000**
			50	**1.000**	1.040	1.599	1.002
0.1	1	10	5	1.020	1.046	1.030	**1.000**
			10	1.003	**1.000**	1.262	1.003
			15	1.009	1.001	1.020	**1.000**
			20	**1.000**	1.009	1.027	1.001
			50	1.001	1.003	1.190	**1.000**
			Average	**1.002**	1.057	1.328	**1.002**

Table 5.7: Runtime for different approaches (in Seconds).

				Runtime (in Seconds)			
H	W	K	λ	DP	ECA	RCA	OH-IECA
1	10	0	5	229.920	2.820	4.280	**0.009**
			10	202.724	2.780	2.830	**0.008**
			15	205.850	2.750	3.210	**0.007**
			20	209.174	2.720	2.860	**0.007**
			50	215.153	3.220	3.810	**0.007**
1	10	10	5	216.044	2.960	3.770	**0.007**
			10	209.489	3.030	2.940	**0.007**
			15	198.881	3.660	4.040	**0.007**
			20	198.972	2.990	4.310	**0.007**
			50	195.048	4.480	3.830	**0.007**
0.1	1	10	5	294.668	2.780	3.170	**0.009**
			10	284.111	2.990	3.370	**0.007**
			15	282.327	3.000	3.240	**0.007**
			20	298.658	3.090	3.690	**0.006**
			50	295.803	3.690	3.400	**0.007**
			Average	235.788	3.131	3.517	**0.007**

We also test the performance when the standard deviation of realized demand is higher than that of what has been predicted and considered for inventory decisions. The average values of $\frac{C_\pi^\infty}{C_{min}^\infty}$ are presented in Table 5.8 and Figure 5.8. The detailed results are provided in Table 2 of Appendix C. While the ECA and the RCA do not perform well in C^h and C^∞ when the standard deviation is unchanged, the performance of the RCA shows some improvement with increase in the standard deviation of realized demand. Meanwhile, the performances of DP and the OH-IECA deteriorate marginally with an increase in the standard deviation. The OH-IECA still outperforms the ECA and the RCA. Therefore, we select the OH-IECA for further performance assessment.

Table 5.8: Average values of $\frac{C_\pi^\infty}{C_{min}^\infty}$ at different variance levels. σ = Forecast standard deviation. $\hat{\sigma}$ = Realized standard deviation. Detail results are given in Table 2 in Appendix C.

$\hat{\sigma}/\sigma$	$\frac{C_{DP}^\infty}{C_{min}^\infty}$	$\frac{C_{ECA}^\infty}{C_{min}^\infty}$	$\frac{C_{RCA}^\infty}{C_{min}^\infty}$	$\frac{C_{OH-IECA}^\infty}{C_{min}^\infty}$
1.00	**1.002**	1.057	1.328	**1.002**
1.05	1.009	1.111	1.440	**1.006**
1.25	1.012	1.132	1.423	**1.009**
1.50	1.015	1.137	1.381	**1.013**
2.00	**1.016**	1.130	1.296	1.025

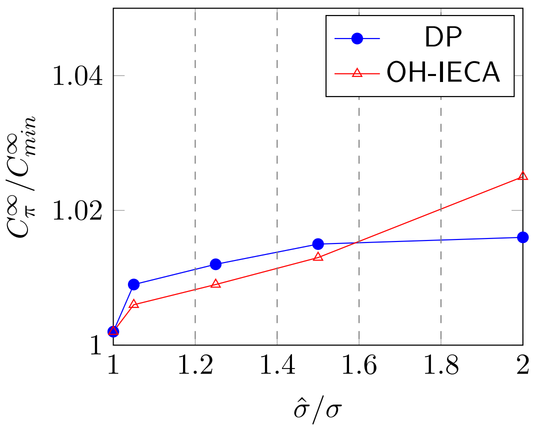


Figure 5.8: Illustration of performances of DP and the OH-IECA for different uncertainty levels.

Table 5.9: Comparative performance of the OH-IECA for the benchmark problems in Veinott Jr and Wagner (1965). C^* is the optimal cost for the problem instance. $C^\infty_{OH-IECA}$ is the infinite expected per period cost using the OH-IECA. CoV indicates the coefficient of variation for that cost. $\overline{\Delta C(\%)}$ is the average excess cost of using the OH-IECA over the optimal cost. min.$\Delta C(\%)$ is the minimum cost difference during the simulations and max.$\Delta C(\%)$ is the maximum cost difference obtained during the simulations.

H	W	K	λ	C^*	$C^\infty_{OH-IECA}$	CoV	$\overline{\Delta C(\%)}$	min.$\Delta C(\%)$	max.$\Delta C(\%)$
1	9	64	21	50.40590	50.77660	0.01170	0.735	0.000	4.304
			22	51.63222	51.81350	0.00720	0.351	0.000	3.181
			23	52.75658	52.90170	0.00415	0.275	0.000	1.753
			24	53.51777	53.74510	0.00520	0.425	0.000	2.721
			51	71.61085	71.76700	0.00436	0.218	0.000	2.169
			52	72.24602	72.38810	0.00370	0.197	0.000	1.447
			55	74.14860	74.38540	0.00999	0.319	0.000	4.771
			59	76.67902	76.89960	0.00687	0.288	0.000	2.181
			61	77.92867	77.92868	0.00383	0.000	0.000	1.718
			63	78.28676	79.21260	0.00799	1.183	0.000	4.550
			64	78.40221	79.26510	0.00737	1.101	0.000	2.921

5.8.3. Tests on Benchmark Problem Instances

In the previous problem instances, the solutions obtained with dynamic programming are not optimal since the end state in not known. We conduct those experiments to be more closer to practical situations. In such situations, end conditions are usually not known, forecast information is reliably available only for short periods, and the decisions are implemented in a longer time horizon.

In order to deeper analyze the performance, we test the OH-IECA on the problems described in Veinott Jr and Wagner (1965). Veinott Jr and Wagner (1965) determined the optimal policies and their infinite horizon expected per period cost. We compute the expected cost per period if the OH-IECA was followed and compare it with the optimal cost. We also compute the coefficient of variation (CoV) of the expected cost, the average difference from the optimal cost, the minimum as well as the worst case cost. All of our results are obtained from running 10^6 simulations. For the heuristic, we use 100 samples with $T = 10$ drawn at random.

Table 5.9 illustrates the results obtained for the benchmark problems if the demand realization follows the same distribution as the forecast. C^* is the optimal per period cost (following the optimal policy). The average optimality gap is 0.463%. As we run multiple iterations ant take the average cost, we also indicate the minimum and maximum cost over the iterations. The coefficient of variation is near 1%. With more number of samples this can be further reduced. In Table 5.10, we present the performance if the realized standard deviation is 0%, 5%, 25%, 50% and 100% higher. The optimal policies from Veinott Jr and Wagner (1965) and the ordering approach using OH-IECA are compared. Table 5.10 presents the average cost difference between the two approaches under different circumstances. A positive difference indicates that the cost obtained using the optimal policy is lower than that obtained using OH-IECA. A negative difference indicates that the OH-IECA performs better.

Table 5.10: Comparative performance of the optimal policy and the OH-IECA on the benchmark problems given in Veinott Jr and Wagner (1965) for different realizations of the standard deviation. σ = Forecast standard deviation and $\hat{\sigma}$ = Realized standard deviation. The percentage excess cost of the OH-IECA over the cost obtained with the optimal policy is presented.

H	W	K	λ	$\Delta C(\%)$ at $\hat{\sigma}/\sigma =$				
				1.000	1.050	1.250	1.500	2.000
1	9	64	21	0.735	0.818	1.322	2.189	2.826
			22	0.351	0.623	1.803	2.685	4.561
			23	0.275	0.931	0.965	1.235	0.412
			24	0.425	1.104	1.396	1.465	0.967
			51	0.218	0.759	0.563	0.863	0.639
			52	0.197	0.867	0.554	0.636	0.605
			55	0.319	0.868	0.896	0.836	0.825
			59	0.288	0.770	1.110	1.596	2.590
			61	0.000	0.387	0.816	1.408	4.215
			63	1.183	0.720	0.563	**-0.038**	**-1.873**
			64	1.101	0.668	0.169	0.069	**-1.631**
			Average	**0.463**	**0.774**	**0.923**	**1.176**	**1.285**

5.8.4. Tests on Non-stationary Demand

We select 11 (P1 to P11) non-stationary demand datasets (see Table 5.11.) from Özen *et al.* (2012). The fixed order costs are set at $K = \{0, 25, 50\}$, the inventory costs are set at $h = \{0.1, 0.5, 1\}$ and the backorder costs are set at $W = \{5, 10, 20\}$. Therefore, we test $3 \times 3 \times 3 = 27$ instances for each of the demand data. The numerical results of OH-IECA for these non-stationary problem instances are given in Table 5.12. The average optimality gap obtained is 7.3% for the 11 problem instances. The proposed methodology is extended in Chapter 6 to give better performance.

Table 5.11: Problem instances with non-stationary demand.

Instance	Time periods											
	1	2	3	4	5	6	7	8	9	10	11	12
P1	50	50	50	50	50	50	50	50	50	50	50	50
P2	68	83	88	83	69	50	31	17	12	17	32	50
P3	10	45	87	91	82	86	75	40	24	34	21	5
P4	3	9	16	28	34	37	43	59	70	91	99	111
P5	111	99	91	70	59	43	37	34	28	16	9	3
P6	27	22	94	27	17	74	120	12	50	28	19	110
P7	5	5	5	5	5	5	5	5	5	5	5	5
P8	7	8	9	8	7	5	3	2	1	2	3	5
P9	1	5	9	9	8	9	8	4	3	2	1	1
P10	1	2	2	3	4	5	5	6	7	8	8	9
P11	9	8	8	7	6	5	5	4	3	2	2	1

Table 5.12: Percentage optimality gap of OH-IECA with respect to dynamic programming in case of non-stationary demand.

H	W	K	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
0.10	5	0	1.98	3.61	3.96	2.11	6.61	2.58	8.11	9.26	24.36	24.62	7.70
0.10	10	0	2.29	6.35	3.66	6.43	9.25	8.39	6.71	18.02	27.76	28.22	13.56
0.10	20	0	10.69	10.30	15.55	10.88	9.85	6.57	13.05	25.43	32.35	40.99	19.75
0.50	5	0	1.69	2.86	0.89	1.87	2.28	0.93	3.88	7.36	18.66	20.06	7.01
0.50	10	0	2.57	2.85	2.18	3.62	2.86	2.41	2.08	13.36	23.12	21.75	8.35
0.50	20	0	3.48	3.58	5.07	2.35	2.75	6.09	8.52	13.44	26.41	22.89	9.65
1.00	5	0	0.84	1.78	1.61	3.77	0.12	2.64	1.94	5.92	16.94	15.92	5.65
1.00	10	0	1.41	0.18	0.43	1.93	1.66	0.47	2.32	7.16	18.14	20.83	6.25
1.00	20	0	0.73	2.24	3.28	2.51	0.31	2.46	3.18	10.15	24.19	19.17	8.85
0.10	5	25	2.45	3.62	2.34	3.59	5.59	4.87	17.81	14.38	14.24	13.69	5.84
0.10	10	25	5.19	5.40	0.10	4.38	9.05	8.25	17.17	13.28	16.20	16.29	1.00
0.10	20	25	8.35	8.53	3.29	3.71	13.6	13.48	15.06	5.26	12.63	20.35	7.92
0.50	5	25	0.27	0.62	0.00	1.15	1.69	4.18	1.48	4.67	11.47	12.80	2.30
0.50	10	25	0.98	1.59	1.36	0.77	1.96	3.17	1.55	6.42	18.68	12.12	3.61
0.50	20	25	1.37	2.38	5.04	0.71	2.38	3.66	2.17	8.76	15.39	6.56	5.78
1.00	5	25	0.00	0.56	0.89	2.09	0.90	0.03	0.76	8.23	17.14	14.84	4.18
1.00	10	25	0.40	1.01	0.93	0.75	1.07	1.38	2.18	6.48	15.10	16.29	2.26
1.00	20	25	0.95	1.05	2.38	0.73	2.03	2.32	1.57	6.42	14.25	12.20	4.47
0.10	5	50	2.51	4.34	4.66	1.01	8.20	5.04	37.94	27.56	10.76	11.52	21.03
0.10	10	50	4.81	3.93	2.63	6.10	11.38	4.60	34.64	25.04	11.86	23.01	11.01
0.10	20	50	7.68	4.22	1.81	9.63	15.45	6.90	31.43	22.99	13.38	27.79	3.57
0.50	5	50	0.69	2.21	2.69	0.68	1.99	4.21	4.48	3.28	11.59	10.42	9.22
0.50	10	50	0.93	1.97	0.94	1.46	1.84	2.21	5.84	3.83	11.72	8.77	3.44
0.50	20	50	1.09	2.33	0.66	2.25	3.66	6.63	4.28	2.55	17.06	4.19	5.18
1.00	5	50	0.25	0.72	0.57	0.54	1.06	2.71	0.94	8.81	17.27	14.78	6.69
1.00	10	50	0.19	1.22	0.63	1.29	1.23	3.89	1.02	4.64	12.40	13.20	2.34
1.00	20	50	0.31	1.48	1.73	0.69	2.21	4.00	1.23	7.27	18.01	12.73	3.31
	Average		2.37	2.99	2.56	2.85	4.48	4.22	8.56	10.73	17.44	17.25	7.03

5.9. Extensions

The proposed method requires minimal modifications to address problems with batch-size constraints, robust solutions or minimum order quantity. It is also flexible to incorporate lost sale and service level constraint. For lost sales, the cost function remains convex and similar closed form approximation can be obtained for samples. We plan to incorporate lost sales and service level constraint during industrialization. We explain the extensions for batch-size constraints, robust solutions, minimum order quantity below.

5.9.1. Addressing Batch Size Constraint

In practice we often encounter replenishment problems with batch-size or pack-size constraints. In this part, the aim is to study how the proposed method can be used in in such situations. In the literature, problems with batch size are dealt by finding the optimal parameters of some standard inventory policies. Those policies can be of (r, Q), (r, nQ) or (s, nQ) type. However, those policies are not optimal themselves. In the presence of batch-size constraint, the changes to our proposed heuristic are presented in Algorithm 3. Since for a coverage period the cost function is convex, we choose the two adjoining points of the minimum that are divisible by the given batch size. Then we follow the method as earlier to obtain the minimum cost over all the possible coverage periods.

Algorithm 3: Optimized Heuristic (OH-IECA$_b$) with Batch Size

Input : $\mathbf{F}, H, W, K, T, s_0, b =$ Batch Size
Output: q^*

Initialize $\Omega \leftarrow 1, u \leftarrow \infty, u' \leftarrow \infty, q^* \leftarrow 0$
repeat
  Compute q^*_Ω using (5.31)
  Compute $q^*_1 = \lfloor \frac{q^*}{b} \rfloor b$ and $q^*_2 = \lceil \frac{q^*}{b} \rceil b$
  Compute Costs $D_\Omega(s_0, q^*_1)$, $D_\Omega(s_0, q^*_2)$ and $D_\Omega(s_0, 0)$
  $u' \leftarrow \min_{q \in \{q^*_1, q^*_2, 0\}}(D_\Omega(s_0, q))$
  if $u' < u$ **then**
    $u \leftarrow u'$
    $q^* \leftarrow \arg\min_{q \in \{q^*_1, q^*_2, 0\}}(D_\Omega(s_0, q))$
  end
  $\Omega \leftarrow \Omega + 1$
until $\Omega = T$;

Table 5.13: Comparative performance of the proposed heuristic for the benchmark problems in the presence of different batch sizes (Percentage excess cost over dynamic programming).

K	H	W	λ	$\overline{\Delta C(\%)}$ at Batch size = 40	25	10	8	5	4	1
1	9	64	21	-1.490	0.062	0.275	0.581	0.437	0.434	0.735
			22	-1.412	0.025	0.794	0.052	0.400	0.408	0.351
			23	-1.346	-0.223	0.480	0.196	0.090	0.124	0.275
			24	-1.107	-0.068	-0.251	0.020	0.080	0.041	0.425
			51	-0.119	-0.674	0.294	-0.055	0.225	0.177	0.218
			52	0.051	-0.981	0.262	0.267	-0.092	0.241	0.197
			55	0.743	0.167	0.303	0.159	0.063	0.058	0.319
			59	0.316	0.668	0.206	0.342	0.390	0.370	0.288
			61	-0.336	0.021	0.573	0.207	0.362	0.869	0.000
			63	-0.378	0.316	0.300	0.409	0.496	0.511	1.183
			64	-0.117	0.250	0.136	0.629	0.529	0.644	1.101
			Average	**-0.473**	**-0.040**	**0.307**	**0.256**	**0.271**	**0.353**	**0.463**

We compared solutions obtained using OH-IECA$_b$ with those obtained using dynamic programming. The cost differences are presented in Table 5.13. The tests are conducted with batch sizes 40, 25, 10, 8, 5 and 4 considering practical packaging situations. In the absence of any knowledge about the end state, dynamic programming does not provide the optimal solution. It is coherent with practical situations as the end state is usually unknown. With higher batch size dynamic programming requires more rolling horizon length to produce better solution. As the rolling horizon length is fixed at 10 periods, it can be observed from Table 5.13 that OH-IECA$_b$ performs better with increase in batch size.

5.9.2. Robust Solution

When the realized demand has higher variance, we proposed to include a **robustness factor** $r \in [0, 1)$ to improve the performance. The robustness factor is part of the approximation given below. This modified approximation is then used in either OH-IECA (Algorithm 2) or OH-IECA$_b$ (Algorithm 3) depending on the presence of batch size.

$$q^*_\Omega \approx \mathfrak{q}_{\min(1,\ \zeta^* - \zeta^*_l + r)}(F_{z\zeta^*_h}) + \sum_{z=1}^{Z} \sum_{t=1}^{\zeta^*_l} P_z F_{zt} \tag{5.36}$$

Table 5.14: Average performance of the proposed heuristic for the benchmark problems at different realizations of the standard deviation and at different value of robustness factor r. (Percentage excess cost over optimal (s, S) policy) for the benchmark instances.

	$\hat{\sigma}/\sigma$				
r	1	1.05	1.25	1.50	2.00
0.00	0.463	0.774	0.923	1.176	1.285
0.05	0.503	0.519	0.440	0.330	-0.140
0.10	0.658	0.408	0.196	-0.084	-0.950
0.20	5.007	3.759	1.853	0.231	-1.865
0.30	5.707	4.319	2.124	0.390	-1.441
0.50	5.601	4.205	2.005	0.335	-1.551

In Table 5.14 the average optimality gaps for the eleven problem instances are presented for various values of robustness factor r at different realization of standard deviation. The results show that a higher robustness factor provides better results if the realized demand is more uncertain.

5.9.3. Addressing Minimum Order Quantity

Suppliers often quote a minimum order quantity to the retailer. The retailer, in such cases, can only place orders whose quantity is higher than or equal to the minimum order quantity. The modified heuristic OH-IECA$_m$ to address this constraint is presented in Algorithm 4.

Algorithm 4: Optimized Heuristic (OH-IECA$_m$) minimum order quantity

Input : $\mathbf{F}, H, W, K, T, s_0, m =$ Minimum Order Qunantity
Output: q^*
Initialize $\Omega \leftarrow 1, u \leftarrow \infty, u' \leftarrow \infty, q^* \leftarrow 0, \tilde{q} \leftarrow 0$
repeat
    Compute q^*_Ω using (5.31)
    if $q^*_\Omega < m$ **then**
        Compute Costs $D_\Omega(s_0, 0)$, $D_\Omega(s_0, m)$
        if $D_\Omega(s_0, q^*) > D_\Omega(s_0, m)$ **then**
            $\tilde{q} \leftarrow m$
            $u' \leftarrow D_\Omega(s_0, m)$
        end
        else
            $\tilde{q} \leftarrow m$
            $u' \leftarrow D_\Omega(s_0, m)$
        end
    end
    else
        $\tilde{q} \leftarrow q^*_\Omega$
        $u' \leftarrow D_\Omega(s_0, q^*_\Omega)$
    end
    if $u' < u$ **then**
        $u \leftarrow u'$
        $q^* \leftarrow \tilde{q}$
    end
    $\Omega \leftarrow \Omega + 1$
until $\Omega = T$;

5.10. Conclusions

In this chapter, we studied the single item multi-period stochastic inventory system and propose a near optimal replenishment planning procedure that works with various practical constraints. The studied problem is a multi-stage stochastic optimization problem. Some limitations of existing solution methods are as follows. Due to the curse of dimensionality, exact methods such as dynamic programming and stochastic programming cannot be used. Optimal policy parameters for the infinite horizon problem can be obtained efficiently with existing methods, however, they work only for parametric and stationary demand distributions. Other practical constraints such as presence of batch size are not considered by such methods. Moreover, when the demand is non stationary, we need to evaluate the policy parameters repeatedly and they can be computationally expensive.

We analyzed two different approaches in which future demand is expressed as random sample traces and ordering decisions are placed as per a rolling horizon approach. The first approach: expected cost approach (ECA) minimizes the expected rolling horizon cost by considering a static uncertainty strategy. The second approach: robust cost approach (RCA) minimizes the maximum cost across samples. It also follows a static uncertainty strategy. We proposed another approach which we call the immediate expected cost approach (IECA), that minimizes the cost for the immediate period after the order.

The first two approaches can be solved using a standard solver. For the IECA, we first proposed an enumerative search heuristic with complexity of the order of $\mathcal{O}(T|\mathcal{Q}|)$. We have provided some convexity properties for the cost function and then proposed an improved approximation heuristic OH-IECA based on it. The complexity of the improved heuristic is of the order $\mathcal{O}(2T)$.

We tested the effectiveness of the proposed ECA, RCA and OH-IECA by conducting a comparative study taking the solution obtained using dynamic programming as a baseline. We proposed two performance evaluation criteria: horizon cost criterion and infinite horizon cost criterion. Apart from those, we also examined the runtime required by each approach. Tests on synthetic data show that the OH-IECA provides equivalent solutions to that of dynamic programming in terms of cost. Moreover, its execution time is almost instantaneous. While DP requires more than 3 minutes, the OH-IECA takes about 7 milliseconds. Therefore, we selected the OH-IECA for further analysis with benchmark problems provided in Veinott Jr and Wagner (1965). The OH-IECA gives 0.463% excess cost over the optimal policy on average. We also analyzed the situation when the variance of realized demand is higher than that of the forecast. In such situation, the OH-IECA still performs close to the optimal policy with the average excess cost equal to 1.285% when the realized standard deviation is twice that of the forecast.

The proposed method requires little modification to include batch sizes. In the presence of batch size the optimal orders can be obtained using dynamic programming. For the benchmark problems, we test our method for batch sizes 40, 25, 10, 8, 5 and 4. The average excess cost is less that 1% in each case.

In this chapter we addressed replenishment planning in the multi-period stochastic inventory system from various practical viewpoints, such as non-parametric, non-independent and non-stationary demand distributions as well as the presence of batch sizes. The proposed method is also computationally much less expensive than dynamic programming and dynamic policy calculation. In the next chapter, we propose to use this approach in a dynamic programming setting to optimize the total cost over the whole rolling horizon without any substantial increase in computation time. Such an approach has the potential to provide even better performance for non-stationary demand and to incorporate supplier selection. Other possible extensions that are relevant during industrialization include lostsale scenario where the cost function remains convex, presence of service level constraint and stochastic joint replenishment planning.

Chapter 6

Sampling-based Replenishment Planning in Non-Stationary Multi-Supplier Inventory Systems

Abstract: *In this chapter, we extend the replenishment planning method proposed in Chapter 5 to obtain better performance with non-stationary demand and to include supplier selection. Existing methods of dealing with non-stationary demand are computationally expensive and less flexible to incorporate additional problem parameters. Supplier selection too, has significant practical implication due to the presence of supplier specific cost parameters and constraints. For example, unit price and fixed cost constitute a significant portion of total inventory cost. We propose a dynamic programming formulation with reduced state space for inventory systems with non-stationary demand. A similar formulation is also proposed for problems with multiple suppliers. We propose and study two solution approaches regarding supplier selection: common supplier selection and dynamic supplier selection. We also conduct numerical experiments to test their efficacy. Parts of this chapter (Sahu* et al.*, 2020b) has been presented in the international conference on operations research and enterprise systems (ICORES) 2020.*

6.1. Introduction

In most practical situations, demand is non-stationary. Solving the single-item replenishment planning (SRP) problem with non-stationary demand falls in the class of multi stage stochastic programming (MSSP) (Homem-de Mello and Bayraksan, 2014). In this chapter, we propose an extension to the method described in Chapter 5 to obtain approximate solution of the SRP problem. Additionally, the proposed method is modified to address the single-item replenishment planning with supplier selection (SRPSS) problem.

Integrated replenishment planning with supplier selection is one of the core problems faced by retailers. With growing competitiveness in the current market, inclusion of purchasing price and thereby supplier selection in inventory optimization becomes very important. The inherent MSSP problem for the multi-period inventory optimization problem is very difficult to solve optimally due to the well known curse of dimensionality Defourny *et al.* (2012). Supplier selection adds additional decision or action states and further increases the complexity. In this chapter, we first analyze the economic benefits of a dynamic supplier selection approach and afterwards develop an approximate method to solve this problem.

Supplier selection has received considerable attention in the inventory optimization literature post 2003. Initially, supplier selection or multiple sourcing options have been seen as a measure of supply chain risk mitigation. However, in addition to that, multiple suppliers can have significant monetary benefits in terms of costs. Most of the replenishment planning models with multiple suppliers optimize the total cost. This cost is the summation of the replenishment costs, inventory holding costs and shortage costs. Replenishment costs consist of purchasing cost of the item, and fixed order/setup costs for placing an order. Typically, the fixed order cost is independent of order quantity and charged for every order placed. The purchasing cost can be different for each supplier and influenced by any discounting schemes. Shortage costs represent the costs paid by the buyer when it is unable to fulfill its demand. It can be either for backordering, lost sales or a mix of both. Additionally, costs such as disposal costs for perishable items, miscellaneous operational costs (i.e. investments, operations, maintenance costs) are also taken into account in some works of the literature. A detailed review of supplier selection problems can be found in Yao and Minner (2017).

Retailers plan their inventory with a short review period. Availability of multiple suppliers for the same item poses greater challenges for cost effective operation. Since the total cost (and thereby the profitability) is closely related to the purchase price of an item, an integrated planning method becomes essential. A single supplier for the planning horizon is more practical than dynamic supplier selection. However, any decision must be taken with due consideration to its cost implications. Our contributions with regard to SRP with non-stationary demand and SRPSS are the following.

1. We propose a dynamic programming approach with smaller state space than the conventional one for the SRP problem with non-stationary demand.
2. We examine the cost impact of a dynamic supplier selection approach over selecting a common supplier for the whole planning horizon.
3. We extend the previous approach for SRP with non-stationary demand and propose a practical framework for dynamic supplier selection and replenishment planning with stochastic demand.

The rest of this chapter is organized as follows. In Section 6.2, we discuss some preliminary concepts regarding the chapter. Then we propose formulations for the SRP problem with non-stationary demand in Section 6.3 as the extension of minimum coverage period cost (MCPC) approach in Chapter 5. In Section 6.4, we present the context of multi-supplier inventory models, the framework for comparison between dynamic supplier selection and selecting a common supplier for the whole planning horizon. We also propose a near-optimal method for supplier selection which can be used for real-world applications. In Section 6.5 we present the results of the experiments and discuss their relevance. We conclude the chapter and propose some future research directions in Section 6.6.

6.2. Preliminaries

For this chapter, two preliminary concepts are required. First, we present various control strategies for multi-period stochastic inventory optimization problems. The second one is rolling horizon framework, for which we refer to Section 5.4.

The control strategy or the uncertainty strategy for the multi-period stochastic inventory optimization problem is defined based on two conditions: when the decisions regarding the order timings (schedule) are taken, and when the corresponding order quantities are decided. Those control strategies are broadly divided into three categories: static, static-dynamic, and dynamic uncertainty (Rossi *et al.*, 2015). When the decision maker determines both the ordering schedule and the order quantities at the very beginning of the planning horizon, it falls under the static uncertainty strategy. In case of the static-dynamic uncertainty strategy, timing of inventory reviews are fixed at the beginning of the planning horizon and the associated order quantities are decided only when orders are issued. The dynamic uncertainty strategy allows the decision maker to decide dynamically at each time period whether or not to place an order and how much to order. This strategy is known to be cost-optimal (Scarf, 1959). Our proposed methodology follows a static-dynamic uncertainty strategy. This is due to difficulty in practical implementation of a dynamic uncertainty strategy.

Table 6.1: Parameters and variables for the optimization models.

Sets	
$\mathcal{Z}$	Set of samples, index for sample $z \in \{1, ..., Z\}$
$\mathcal{T}$	Set of time periods, index for periods $t \in \{1, ..., T\}$
Parameters	
P_z	Occurrence probability of sample z, $\in [0, 1]$
F_{zt}	Demand forecast for sample z at time t, $\in \mathbb{R}^+$
K	Fixed ordering cost per order, $\in \mathbb{R}^+$
H	Inventory holding cost per unit, per unit time period, $\in \mathbb{R}^+$
W	Backorder cost per unit, $\in \mathbb{R}^+$
A	Discount factor, $\in (0, 1]$
$\mathbf{M}$	A suitably large positive number
Decision Variables	
q_t	Order quantity at time t, $\in \mathbb{Z}^+$
Internal Variables	
d_t	Random demand variable at time period t, $\in \mathbb{Z}^+$
s_{zt}	Inventory for sample z at the end of period t, $\in \mathbb{R}$
l_{zt}	Backorder quantity for sample z at time t, $\in \mathbb{R}^+$
c_z	Total cost of sample z, $\in \mathbb{R}^+$
α_t	Binary, 1 if fixed cost is charged at time t and 0 otherwise $\in \mathbb{Z}^+$
m	Worst case cost $\in \mathbb{R}^+$

6.3. SRP with Non-Stationary Demand (Single-Supplier)

In this section, we focus on the solution method for the single-supplier SRP problem with non-stationary demand. For details of the SRP problem we refer to Chapter 5. First, we compute different ordering options in a multi-period inventory system then propose a reduced-state dynamic program (RDP) to obtain the optimized cost for the horizon.

Problem Class

The problem we consider has a single-echelon supply chain network, zero lead time and stochastic non-stationary demand. Excess demand is backordered completely. Also, one item is ordered from a single supplier in a multi-period ordering. This basic problem can be denoted using the nomenclature given in Chapter 2:
`1;`ϕ`;`ϕ`||O;D||N;B;`ϕ`||C||P;U||S;S;M`

This basic structure can have extensions, such as with lost sales, single batch size, deterministic or stochastic lead time. For example:

1;ϕ;ϕ||O;D||N;**L**;ϕ||C||P;U||S;S;M
1;ϕ;ϕ||O;D||N;B;ϕ||C||P;**S**||S;S;M
1;ϕ;ϕ||**D**;D||N;B;ϕ||C||P;U||S;S;M.

6.3.1. Ordering Options

The central idea behind RDP is the use of different "ordering blocks". We explain them next. The notations used are presented in Table 6.1. Let us consider a rolling horizon of length T. An order can be placed at any time $t \in \{1, 2, ..., T\}$. If the order is placed at the beginning of time $t = 1$, the order can have coverage period up to the end of period $t \in \{1, 2, ..., T\}$. If we go further in time, at time $t = 2$, we can have different ordering options based on the state of the inventory.

Recall our initial definition of a coverage period in Chapter 5, that states that the delivered quantity suffices till the end of that coverage period. Our aim is to limit the number of states for which we compute the order quantities. In the spirit of the coverage period, there can be two options. First, we order a positive quantity that lasts till the end of the coverage period. Second, we order zero units, so that we observe backorders at the end of the coverage period. In the first case, we approximate the inventory to be equal to zero. For the second case, we consider the total backorders to be equal to the expected demand unfulfilled. We define $\hat{F}_{T^1}^{T^2} = \sum_{t=T^1}^{T^2} \bar{F}_t$ where $\bar{F}_t = \sum_z P_z F_{zt}$. Then, the possible inventory states at $t = T$ will be $\{0, \hat{F}_{T-1}^{T-1}, \hat{F}_{T-2}^{T-1}, ..., \hat{F}_1^{T-1}\}$. Similarly at $t = T-1$, it will be $\{0, \hat{F}_{T-2}^{T-2}, \hat{F}_{T-3}^{T-2}, ..., \hat{F}_1^{T-2}\}$ and so on. We denote the possible set of states at time t by $\mathcal{B}_t$.

Hence, if an ordering decision is made at time $t = 2$ of the same rolling horizon, the order quantity for any supplier can have coverage period up to end of time period $t \in \{2, 3, ..., T\}$ with an initial inventory in $\mathcal{B}_2$. Similarly, if an ordering decision is made at time t of the same rolling horizon, the order quantity can have coverage period up to $\{t, t+1, ..., T\}$ with initial state in $\mathcal{B}_t$. From the above we have, at $t = 1$, there are T ordering options, at $t = 2$, there are $(T-1)$ ordering options and so on. At $t = T$, there is only one ordering option. Additionally, no order option can also be adopted. Recalling the notations from Chapter 5, the coverage period cost (CPC) is denoted by $\tilde{C}(s_{T^1}, q, T^1, T^2)$ for the coverage period T^1 to T^2 ($T^2 \geq T^1$), order quantity q and opening inventory s_{T^1}. The optimal CPC (i.e. the MCPC) is denoted by $\tilde{C}^*(s_{T^1}, q, T^1, T^2)$. Mathematically, they are defined as follows.

$$\begin{aligned}\tilde{C}(s_{T^1}, q, T^1, T^2) = \sum_{z=1}^{Z} P_z \Big[\sum_{t=T^1}^{T^2} \Big(H\big[s_{T^1} + q - \sum_{\tau=1}^{t} F_{z\tau}\big]^+ \\ + W\big[\sum_{\tau=1}^{t} F_{z\tau} - s_{T^1} - q\big]^+ \Big) + K\alpha \Big]\end{aligned} \tag{6.1}$$

$$\alpha = \begin{cases} 1 \text{ , if } q > 0 \\ 0 \text{ , otherwise} \end{cases} \tag{6.2}$$

$$\begin{aligned}\tilde{C}^*(s_{T^1}, q, T^1, T^2) = \min_q \sum_{z=1}^{Z} P_z \Big[\sum_{t=T^1}^{T^2} \Big(H\big[s_{T^1} + q - \sum_{\tau=1}^{t} F_{z\tau}\big]^+ \\ + W\big[\sum_{\tau=1}^{t} F_{z\tau} - s_{T^1} - q\big]^+ \Big) + K\alpha \Big]\end{aligned} \tag{6.3}$$

From the analysis given in the paragraph above, all possible ordering options during a rolling

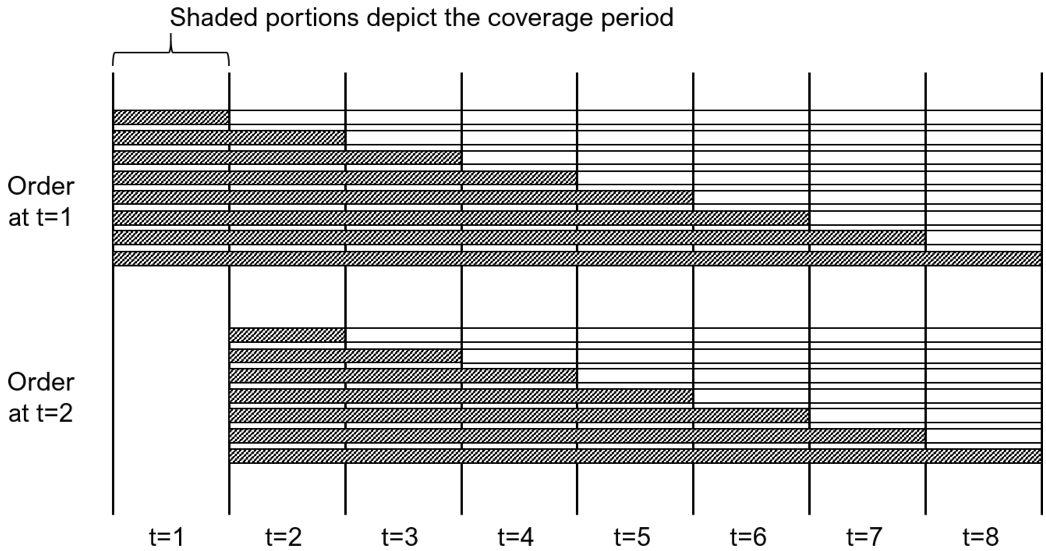


Figure 6.1: Figure demonstrating different possible coverage periods depending upon the order time. The rolling horizon length is 8. At the beginning of $t = 1$, coverage period can be chosen between $t = 1, 2, ..., 8$. Similarly, at the beginning of $t = 2$, the coverage period can be chosen between $t = 2, 3, ..., 8$ and so on.

horizon of length T are given in Table 6.2.

Table 6.2: Different MCPC computations.

$t = 1$	$t = 2$	...	...	$t = T - 1$	$t = T$
$\tilde{C}^*(s_0, q, 1, 1)$	$\tilde{C}^*(\mathcal{B}_2, q, 2, 2)$	...	...	$\tilde{C}^*(\mathcal{B}_{T-1}, q, T-1, T-1)$	$\tilde{C}^*(\mathcal{B}_T, q, T, T)$
$\tilde{C}^*(s_0, q, 1, 2)$	$\tilde{C}^*(\mathcal{B}_2, q, 2, 3)$	...	...	$\tilde{C}^*(\mathcal{B}_{T-1}, q, T-1, T)$	
...	...	...	...		
...	...	...	...		
$\tilde{C}^*(s_0, q, 1, T-1)$	$\tilde{C}^*(\mathcal{B}_2, q, 2, T)$				
$\tilde{C}^*(s_0, q, 1, T)$					

6.3.2. Reduced-state Dynamic Program (RDP)

$\tilde{C}^*(s_{T^1}, q, T^1, T^2)$ denotes the MCPC during the coverage period T^1 and T^2 provided, there is only one order at the beginning of T^1 and none until end of T^2. To find the optimal cost during the rolling horizon, we propose a dynamic programming formulation with reduced state space to solve the multi-stage problem with the cost components as given in Table 6.2. Since we use the CPC, by definition, any order suffices till the end of the coverage period and a new order can be computed for initial inventory in $\mathcal{B}_t$. The proposed dynamic program is a backward recursion method. At the beginning of last time period T of the rolling horizon, there is only one ordering option $\arg\min_q \tilde{C}(s, q, T, T)$ for each inventory state in $\mathcal{B}_T$. The corresponding cost is

$$\mathcal{Y}_T(s_{T-1} \in \mathcal{B}_T) = \tilde{C}^*(s_{T-1}, q, T, T) \tag{6.4}$$

At the beginning of time $t = T - 1$, we have two options for each state. The first one is to have one order having coverage period $T - 1$ to T, and the second one is to have two orders at $T-1$ and T. The best option is the one that gives the least cost.

$$\mathcal{Y}_{T-1}(s_{T-2} \in \mathcal{B}_{T-1}) = \min\Big\{(\tilde{C}^*(s_{T-2}, q, T-1, T-1) + \mathcal{Y}_T(\hat{s}_{T-1})),\\ \tilde{C}^*(s_{T-2}, q, T-1, T)\Big\} \tag{6.5}$$

where

$$\hat{s}_{T-1} = \begin{cases} s_{T-2} - \hat{F}_{T-1}^{T-1} & \text{if } q_{T-1}^* = 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.6}$$

Similarly, at the beginning of $t = T-2$, we have three options to choose from.

$$\mathcal{Y}_{T-2}(s_{T-3} \in \mathcal{B}_{T-2}) = \min\Big\{(\tilde{C}^*(s_{T-3}, q, T-2, T-2) + \mathcal{Y}_{T-1}(\hat{s}_{T-2})),\\ (\tilde{C}^*(s_{T-3}, q, T-2, T-1) + \mathcal{Y}_T(\hat{s}_{T-1})),\\ \tilde{C}^*(s_{T-3}, q, T-2, T)\Big\} \tag{6.7}$$

At $t = 1$,

$$\mathcal{Y}_1 = \min\Big\{(\tilde{C}^*(s_0, q, 1, 1) + \mathcal{Y}_2(\hat{s}_1)),\\ (\tilde{C}^*(s_0, q, 1, 2) + \mathcal{Y}_3(\hat{s}_2)), ..., (\tilde{C}^*(s_0, q, 1, T-1) + \mathcal{Y}_T(\hat{s}_{T-1})),\\ \tilde{C}^*(s_0, q, 1, T)\Big\} \tag{6.8}$$

Therefore, $\mathcal{Y}_1$ corresponds to the minimum cost for the rolling horizon. The replenishment quantity is equal to its corresponding order quantity q^*. The above analysis addresses the situation where we have only one supplier. In the next section, we extend this methodology to include multiple suppliers differentiated by their fixed order cost, unit price and minimum order quantity, etc.

6.4. SRP with Supplier Selection (SRPSS)

In this section, we discuss the case of multi-supplier inventory optimization problems. First, we describe a practical scenario then propose our modeling approaches.

6.4.1. Context and Motivations

Here we elaborate on the SRPSS problem. The problem is motivated by a real-world retailer. It has multiple point of sales at different places along with a central warehouse. We consider one item for replenishment. Future demand is non-stationary. Because of economy of scale, all retailers receive the goods from the central warehouse. The central warehouse in turn orders from the external suppliers (see Figure 6.2). The demand information known at the retailer is stochastic. Therefore, the demand at the central warehouse can also be interpreted as a stochastic process. For each item there are multiple suppliers. Those suppliers differ by the price they charge per unit item, available batch sizes, lead time and fixed cost charged

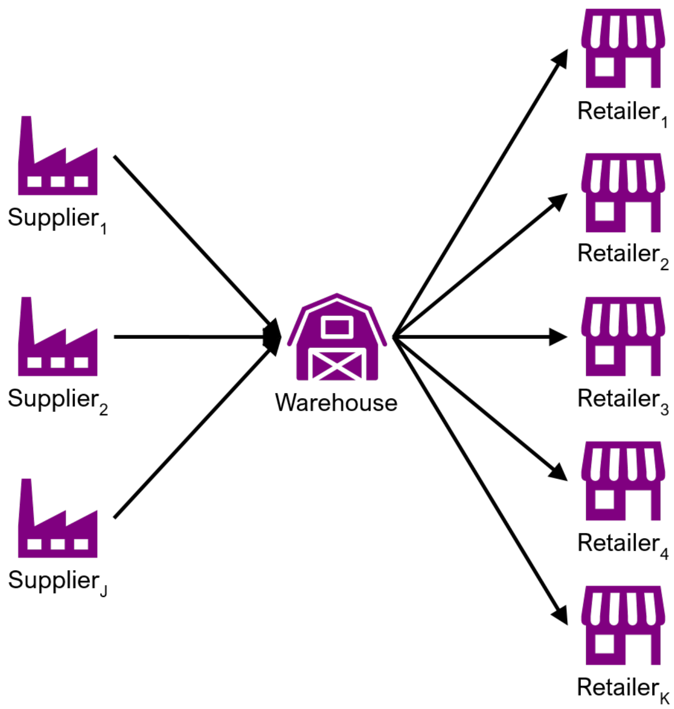


Figure 6.2: The supply chain network under study.

per order. The fixed cost is charged for the transportation and administrative expenses. The retailer aims to minimize total cost incurred during a product life cycle. The usual costs incurred are purchase cost, fixed ordering cost, inventory cost and shortage cost. Any order placed by the warehouse to any supplier is delivered immediately without any lead time. Any product left over after the end of the planning horizon can still be used. Therefore, the salvage value is not taken into consideration. Only inventory cost is charged at the end of the planning horizon.

Two approaches arise from the viewpoint of ease of practical application. First, when the retailer chooses only one supplier for the planning horizon (usually shorter than the product life cycle), and orders from that suppliers only. This approach is easier to implement in practice and the computation process of order quantities is comparatively less expensive than its multi-supplier counterpart. The second approach is to select suppliers dynamically during each ordering decision. This approach is computationally more expensive than the previous one due to the increase in the number of possible decisions in a dynamic programming setting. Beside, this approach is difficult to implement in practice. However, dynamic supplier selection has the potential to be more economical. In this chapter, we aim to first analyze the economic benefits of different supplier selection approaches and propose a cost benefit analysis to the retailer keeping in mind that the practical difficulties can be offset by higher economic gain.

The methods discussed in the previous paragraph give rise to MSSPs. Those MSSPs can become intractable with an increase in number of time periods (stages) and with an increase in number of suppliers. Previous methods given in Cheaitou and Van Delft (2013), Berling and Martínez-de Albéniz (2015) and Berling and Martínez-de Albéniz (2016) consider demand distributions to be independent across time. In practice we often encounter dependent or correlated demand and distributions not following parametric distributions. Such application conditions require new methods. The aim of this chapter is to develop a general framework for the replenishment planning problem with multiple suppliers, that can be implemented in practice.

6.4.2. Problem Formulation

In this section, we propose optimization models for the single-item replenishment problem with supplier selection. The discussed problem has J suppliers. We propose two strategies.

1. Common supplier selection (CSS): When we select only one supplier (best as per cost) for the whole planning horizon.

2. Dynamic supplier selection (DSS): Under this strategy different choices of suppliers can be made during any ordering epoch.

Under the CSS strategy, the multi-stage optimization problem has ordering options concerning only one supplier. On the contrary, under the DSS strategy, an order can be placed at any supplier. We formulate the exact optimization problem for both CSS and DSS approaches. Afterward, we present an approximate optimization framework based on CPC. We consider a planning horizon of length $\hat{T}$, rolling horizon length T and and J suppliers. Suppliers are denoted by $j \in \{1, 2, ..., J\}$. Additional notations are summarized in Table 6.3.

Table 6.3: Notations for the parameters and variables.

Sets	
$\mathcal{J}$	Set of suppliers $j \in \{1, ..., J\}$
Parameters	
d_t	Random demand at time t, $\in \mathbb{R}^+$
U_j	Unit purchase price from supplier j, $\in \mathbb{R}^+$
K_j	Fixed cost of ordering per order from supplier j, $\in \mathbb{R}^+$
Decision Variables	
q_{jt}	Order quantity from supplier j at time t, $\in \mathbb{Z}^+$
α_{jt}	Binary indicator for positive order from supplier j at time t, $\in \{0, 1\}$
s_t	Inventory at the end of time t, $\in \mathbb{Z}^+$

6.4.3. Common Supplier Selection

The multi-period stochastic inventory optimization problem with a common supplier for every period is a finite-stage MSSP. It can be solved optimally with dynamic programming (Özen *et al.*, 2012). This requires the end state to be known and demands across time to be independent. The optimal cost can be found for each supplier independently. Then the supplier having the minimum cost for the planning horizon can be selected. The functional equation of the resulting dynamic program is as follows. With $\mathcal{C}_t^j(s_{t-1})$ being the optimal cost for supplier j with state s_{t-1} at time t, and q_{jt} being the actions

$$\mathcal{C}_t^j(s_{t-1}) = \min_{q_{jt}} \Big\{ \mathbb{E}[H(s_{t-1} + q_{jt} - d_t)^+ + W[-s_{t-1} - q_{jt} + d_t]^+ + K_j \alpha_{jt} + \sum_{j=1}^{J} U_j q_{jt} + \mathcal{C}_{t+1}(s_{t-1} + q_{jt} - d_t)] \Big\} \tag{6.9}$$

$$\alpha_{jt} = \begin{cases} 1 & \text{if } q_{jt} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.10}$$

The first, second, third and fourth terms represent the expected inventory holding costs, expected shortage costs, fixed order cost and purchase costs respectively for period t. The last term represents the minimum expected cost for the next period. The above equation can be solved optimally by value iterations (Puterman, 2014). A solution instance where only one supplier S is selection for the planning horizon, is illustrated in Figure 6.3.

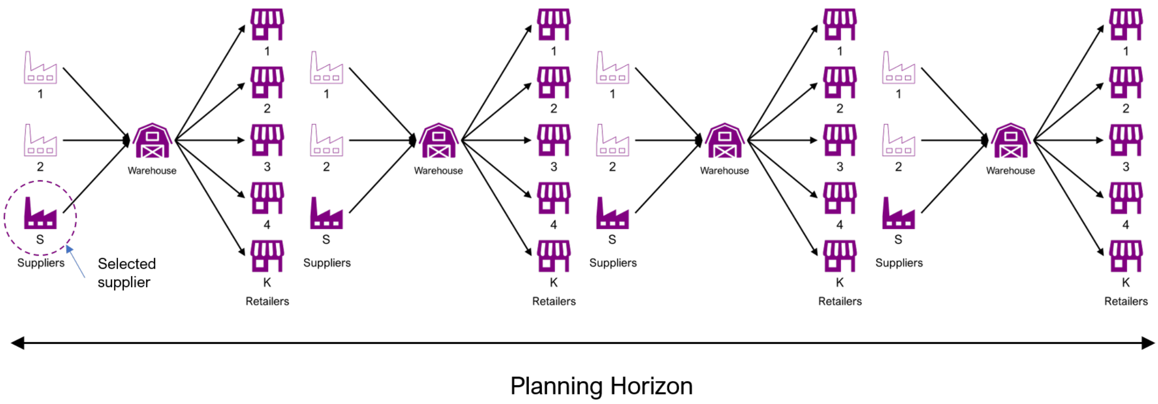


Figure 6.3: An example of solution using the CSS approach.

6.4.4. Dynamic Supplier Selection

The dynamic supplier selection problem is quite similar to the common supplier problem except, it has several sets of possible actions. In the case of common supplier selection, we considered the set of possible actions q_{jt} for each supplier j separately. However, in case of a dynamic supplier selection, we consider all possible actions from all possible suppliers in a single dynamic program. The functional equation is given below. With $\mathcal{C}_t(s_{t-1})$ being the optimal cost with state s_{t-1} at time t, and q_{jt} being the actions

$$\begin{aligned} \mathcal{C}_t(s_{t-1}) = \min_{q_{jt}, j \in \mathcal{J}} \Big\{ & \mathbb{E}[H(s_{t-1} + \sum_{j=1}^{J} q_{jt} - d_t)^+ \\ & + W[-s_{t-1} - \sum_{j=1}^{J} q_{jt} + d_t]^+ + \sum_{j=1}^{J} K_j \alpha_{jt} + \\ & \sum_{j=1}^{J} U_j q_{jt} + \mathcal{C}_{t+1}(s_{t-1} + \sum_{j=1}^{J} q_{jt} - d_t)] \Big\} \end{aligned} \tag{6.11}$$

In the above program, we do not consider the capacity constraint for the suppliers. Inclusion of the capacity constraint can affect the ordering decisions, and we plan to study this in future research. Dynamic supplier selection can be achieved using the above formulation. Choice of supplier at an ordering epoch is affected by the inventory position, unit purchase price, fixed order cost and minimum order quantities, etc. An example of solution using the DSS approach is illustrated in Figure 6.4. During an ordering epoch, the selected supplier is shaded. Different suppliers can be selected at different times.

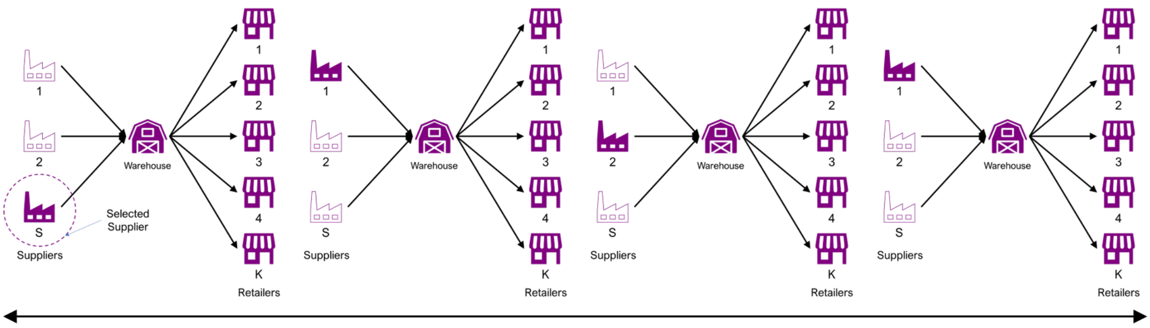


Figure 6.4: An example of solution using the DSS approach.

6.4.5. RDP Framework for the SRPSS

The dynamic programs presented previously can become intractable when the number of periods is high. In this section, we present an approximation framework to alleviate the curse of dimensionality without compromising the solution quality substantially. Recalling the solution procedure from Section 5.5.3 based on MCPC, we add additional decisions based on suppliers.

Let us consider a rolling horizon of length T. An order can be placed at any time $t \in \{1, 2, ..., T\}$. When the decision is taken at time $t = 1$, order can be placed at any one of the suppliers $j \in \{1, 2, ..., J\}$. The order quantity for any supplier can have coverage period up to $\{1, 2, ..., T\}$. If we go further in time at $t = 2$, we can have different ordering options based on the state of inventory. However, our initial definition of the coverage period states that the delivered quantity suffices till the end of that coverage period. Therefore, for each time period $t > 1$, we compute the optimal order quantity and cost assuming the inventory state in $\mathcal{B}_t$. Hence, if an ordering decision is made at time $t = 2$ of the same rolling horizon, the order quantity for any supplier can have coverage period up to $\{2, 3, ..., T\}$. Similarly, if a ordering decision is made at time t of the same rolling horizon, the order quantity for any supplier can have coverage period $\{t, t + 1, ..., T\}$. From the above, at $t = 1$, there are JT

ordering options, at $t = 2$, there are $J(T-1)$ ordering options and so on. At $t = T$, there are J ordering options. Additionally, no order option can also be adopted.

Let $\tilde{C}_j^*(s_{T^1-1}, q, T^1, T^2)$ represent the MCPC for supplier j, for the coverage period T^1 to T^2 ($T^2 \geq T^1$). From the analysis given in the paragraph above, all possible ordering options during a rolling horizon of length T are similar to that given in Table 6.2. Mathematically, $\tilde{C}_j^*(s_{T^1-1}, q, T^1, T^2)$ is defined as follows.

$$\begin{aligned}\tilde{C}_j^*(s_{T^1-1}, q, T^1, T^2) = \min \mathbb{E} \sum_{t=T^1}^{T^2} \Big\{ & H[s_{T^1-1} + q_{jT^1} - \sum_{\tau=T^1}^{t} d_\tau]^+ \\ & + W[-s_{T^1-1} - q_{jT^1} + \sum_{\tau=T^1}^{t} d_\tau]^+ \\ & + U_j q_{jT^1} + \alpha_{jT^1} K_j \Big\}\end{aligned} \tag{6.12}$$

$$\alpha_{jT^1} = \begin{cases} 1 & \text{if } q_{jT^1} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.13}$$

After the computation of all MCPCs, we can solve the multi-stage problem for the whole rolling horizon using a dynamic programming as follows. For CSS,

$$\mathcal{Y}_T^j(s_{T-1} \in \mathcal{B}_T) = \tilde{C}_j^*(s_{T-1}, q, T, T) \tag{6.14}$$

$$\begin{aligned}\mathcal{Y}_{T-1}^j(s_{T-2} \in \mathcal{B}_{T-1}) = \min \Big\{ & (\tilde{C}_j^*(s_{T-2}, q, T-1, T-1) + \mathcal{Y}_T^j(\hat{s}_{T-1}), \\ & \tilde{C}_j^*(s_{T-2}, q, T-1, T) \Big\}\end{aligned} \tag{6.15}$$

$$\begin{aligned}\mathcal{Y}_1^j = \min \Big\{ & (\tilde{C}_j^*(s_0, q, 1, 1) + \mathcal{Y}_2^j(\hat{s}_1)), (\tilde{C}_j^*(s_0, q, 1, 2) + \mathcal{Y}_3^j(\hat{s}_2)), ..., \\ & (\tilde{C}_s^*(s_0, q, 1, T-1) + \mathcal{Y}_T^j(\hat{s}_{T-1})), \tilde{C}_s^*(s_0, q, 1, T) \Big\}\end{aligned} \tag{6.16}$$

The above formulation is a backward recursion. At $t = T$, we have the option of ordering only to cover the period T. Therefore, its expected minimum cost is the MCPC for that supplier with coverage period T. Similar computations are done for all suppliers and the supplier with least $\mathcal{Y}_1^j$ is chosen. Under DSS approach, the following dynamic program is adopted.

$$\mathcal{Y}_T(s_{T-1} \in \mathcal{B}_T) = \min_j \tilde{C}_j^*(s_{T-1}, q, T, T) \tag{6.17}$$

$$\begin{aligned}\mathcal{Y}_{T-1}(s_{T-2} \in \mathcal{B}_{T-1}) = \min \Big\{ & \min_j (\tilde{C}_j^*(s_{T-2}, q, T-1, T-1) + \mathcal{Y}_T(\hat{s}_{T-1}), \\ & \min_j \tilde{C}_j^*(s_{T-2}, q, T-1, T) \Big\}\end{aligned} \tag{6.18}$$

$$\mathcal{Y}_1 = \min \Big\{ \min_j \big(\tilde{C}_j^*(s_0, q, 1, 1) + \mathcal{Y}_2(\hat{s}_1)\big), \min_j \big(\tilde{C}_j^*(s_0, q, 1, 2) + \mathcal{Y}_3(\hat{s}_2)\big), ..., \\ \min_j \big(\tilde{C}_j^*(s_0, q, 1, T-1) + \mathcal{Y}_T(\hat{s}_{T-1})\big), \min_j \tilde{C}_j^*(s_0, q, 1, T) \Big\} \tag{6.19}$$

6.5. Numerical Experiments

In the previous section, we have explained a dynamic programming based approximate solution approach for the SRP with non-stationary demand and for the SRPSS. Experiments are conducted in three phases:

1. Experiments to assess the performance of the proposed method for SRP with non-stationary demand.
2. Experiments to assess the comparative performances of CSS and DSS approaches in SRPSS.
3. Experiments to assess the performance of the proposed method for SRPSS.

6.5.1. Experimental Protocol

Our test-bed for the SRP is as follows. At the beginning of time period $t = 1$, the demand information up to period $\hat{T}$ is available. The decision maker uses this demand information, current inventory and cost parameters to compute the order quantities. We conduct a simulation to compute the expected cost using RDP. The steps are as follows:

S1. Select rolling horizon length T, initial inventory s_0, inventory holding cost H, backorder cost W and fixed order cost K. Set time $t = 1$.

S2. At time t compute the optimized order quantity for the initial inventory, cost parameters and demand samples from t to $t + T$.

S3. The order is received instantly. Update the inventory position as initial inventory plus the order quantity.

S4. Generate random demand d_t from the demand distribution during t. Compute the resulting fixed order cost, inventory holding cost and backorder cost.

S5. Update the inventory level. This is the initial inventory for the next time period $t + 1$.

S6. Repeat S2 to S5 until $t = \hat{T}$. Compute the total cost from $t = 1$ to $t = \hat{T}$. Repeat the simulation 10^3 times and compute the average cost. 10^3 simulations are chosen as we observe less than 0.01% absolute change in the average value.

We conduct tests assuming that the demand follows a Poisson distribution. The mean demands for the problem instances are presented in Table 5.11 in Chapter 5 for the SRP with non-stationary demand. We repeat them again in Table 6.4. For each of the instances P1 to P11, the inventory holding cost $H = \{0.1, 0.5, 1\}$, backorder cost $W = \{5, 1020\}$ and fixed order cost $K = \{0, 25, 50\}$. We compare the expected costs obtained using RDP to the optimal cost obtained using dynamic programming.

Table 6.4: Problem instances for the SRP with non-stationary demand. The table presents the expected demands for 12 periods. For each of the instances P1 to P11, the inventory holding cost $H = \{0.1, 0.5, 1\}$, backorder cost $W = \{5, 1020\}$ and fixed order cost $K = \{0, 25, 50\}$.

Instance	Time periods											
	1	2	3	4	5	6	7	8	9	10	11	12
P1	50	50	50	50	50	50	50	50	50	50	50	50
P2	68	83	88	83	69	50	31	17	12	17	32	50
P3	10	45	87	91	82	86	75	40	24	34	21	5
P4	3	9	16	28	34	37	43	59	70	91	99	111
P5	111	99	91	70	59	43	37	34	28	16	9	3
P6	27	22	94	27	17	74	120	12	50	28	19	110
P7	5	5	5	5	5	5	5	5	5	5	5	5
P8	7	8	9	8	7	5	3	2	1	2	3	5
P9	1	5	9	9	8	9	8	4	3	2	1	1
P10	1	2	2	3	4	5	5	6	7	8	8	9
P11	9	8	8	7	6	5	5	4	3	2	2	1

For the SRPSS, the problem instances are presented in Table 6.5. We take a planning horizon length of $\hat{T} = 20$ periods and rolling horizon length $T = 20$ periods[1]. Demand is assumed to follow a stationary Poisson distribution with $\lambda = 5$. There are 2 suppliers. Their corresponding fixed costs, unit prices and minimum order quantities are given in Table 6.5. Similar to the SRP, for the SRPSS, we conduct simulations to compute the expected costs. For the CSS approach, the procedure is as follows:

C1. Select rolling horizon length T, initial inventory s_0, inventory holding cost H, backorder cost W and fixed order costs $\mathbf{K}$, purchase costs $\mathbf{U}$ and minimum order quantities $\mathbf{MOQ}$.

C2. For each supplier, compute the minimum cost in Equation 6.16. Choose the supplier having minimum cost for the whole planning horizon and select its fixed order cost, purchase cost and minimum order quantity. Repeat the next steps for that supplier.

C3. Set time $t = 1$.

C4. At time t compute the optimized order quantity for the initial inventory, cost parameters and demand samples from t to $t + T$.

C5. The order is received instantly. Update the inventory position as initial inventory plus the order quantity.

C6. Generate random demand d_t from the demand distribution during t. Compute the resulting purchase cost, fixed order cost, inventory holding cost and backorder cost.

C7. Update the inventory level. This is the initial inventory for the next time period $t + 1$.

C8. Repeat C4 to C7 until $t = \hat{T}$. Compute the total cost from $t = 1$ to $t = \hat{T}$. Repeat the simulation 10^3 times and compute the average cost.

For the DSS approach, the procedure is as follows.

D 1. Select rolling horizon length T, initial inventory s_0, inventory holding cost H, backorder cost W and fixed order costs $\mathbf{K}$, purchase costs $\mathbf{U}$ and minimum order quantities $\mathbf{MOQ}$.

[1]We take the length of planning and rolling horizons of same length to have the least cost for the whole planning horizon. A shorter rolling horizon will lead to higher cost for the planning horizon

D 2. Set time $t = 1$.

D 3. At time t compute the optimized order quantity for the initial inventory, cost parameters and demand samples from t to $t = t + T$ using the minimum cost in Equation 6.19.

D 4. The order is received instantly. Update the inventory position as initial inventory plus the order quantity.

D 5. Generate random demand d_t from the demand distribution during t. Compute the resulting purchase cost, fixed order cost, inventory holding cost and backorder cost.

D 6. Update the inventory level. This is the initial inventory for the next time period $t + 1$.

D 7. Repeat C 3 to C 6 until $t = \hat{T}$. Compute the total cost from $t = 1$ to $t = \hat{T}$. Repeat the simulation 10^3 times and compute the average cost.

Table 6.5: Problem instances for the SRPSS.

H	W	K_1	K_2	U_1	U_2	MOQ_1	MOQ_2
1	10	50	50.00	10	10.00	0	0
1	10	50	55.00	10	9.52	0	0
1	10	50	60.50	10	9.07	0	0
1	10	50	66.55	10	8.64	0	0
1	10	50	73.21	10	8.23	0	0
1	10	50	80.53	10	7.84	0	0
1	10	50	88.58	10	7.46	0	0
1	10	50	97.44	10	7.11	0	0
1	10	50	107.18	10	6.77	0	0
1	10	50	117.90	10	6.45	0	0
1	10	50	50.00	10	10	5	5
1	10	50	55.00	10	10	5	10
1	10	50	60.50	10	10	5	15
1	10	50	66.55	10	10	5	20
1	10	50	73.21	10	10	5	25
1	10	50	80.53	10	10	5	30
1	10	50	88.58	10	10	5	35
1	10	50	97.44	10	10	5	40
1	10	50	107.18	10	10	5	45
1	10	50	117.90	10	10	5	50
1	10	100	100	10	10.00	5	5
1	10	100	100	10	9.52	5	10
1	10	100	100	10	9.07	5	15
1	10	100	100	10	8.64	5	20
1	10	100	100	10	8.23	5	25
1	10	100	100	10	7.84	5	30
1	10	100	100	10	7.46	5	35
1	10	100	100	10	7.11	5	40
1	10	100	100	10	6.77	5	45
1	10	100	100	10	6.45	5	50

Some preliminary assessments are as follows. There are three parameters specific to suppliers, fixed order costs K_j, unit prices U_j and minimum order quantities MOQ_j. Varying just one parameter will yield solutions with only one supplier selected. This is because, at any ordering epoch the same supplier will have the least cost. Therefore, we may obtain a difference between the CSS and DSS approaches only when two or more parameters are varied at the same time. In our analysis, we vary two parameters at a time K_j and U_j, K_j and MOQ_j, U_j and MOQ_j. Also, their variations must be coherent, i.e. the supplier with lower fixed order cost would have higher unit price, the supplier with lower fixed order cost would have higher minimum order quantity and the supplier with lower unit price would have higher minimum order quantity.

Table 6.6: Percentage optimality gap of RDP with respect to dynamic programming in case of non-stationary demand.

H	W	K	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
0.1	5	0	4.61	5.56	0.97	4.21	4.11	3.37	0.29	4.44	1.73	0.02	0.8
0.1	10	0	11.68	0.67	1	0.81	3.31	3.79	0.71	0.13	3.57	3.77	3.21
0.1	20	0	8.15	15.1	13.09	11.16	7.55	5.31	14.11	2.59	1.37	1.71	35.64
0.5	5	0	3.71	4.53	2.28	0	1.46	5.05	1.93	1.04	4.05	2.88	4.64
0.5	10	0	0	5.31	2.05	0.94	0.49	4.46	3.93	2.99	4.32	3.95	3.03
0.5	20	0	0.76	1.12	1.22	0.13	1.66	3	3.1	0.3	6.62	3.4	9.49
1	5	0	3.84	0.6	1.35	0.24	7.73	1.49	5.21	4.32	2.69	3.2	0.45
1	10	0	0.94	0.11	0.04	3.73	1.37	1.94	2.98	2.26	5.65	0	5.03
1	20	0	6.72	3.44	1.01	4.14	8.57	1.65	0.33	0	2.3	5.98	2.83
0.1	5	25	1.87	5.81	5.73	0	2.93	2.7	0.97	1.48	3.89	0.44	0.26
0.1	10	25	4.74	3.92	5.21	5.05	4.21	3.47	3.51	3.24	9.16	0.89	5.66
0.1	20	25	3.09	5.13	7.7	3.19	2.81	2.49	3.01	4.2	10.49	3.26	7.85
0.5	5	25	0.89	0.75	2.3	1.23	3.05	2.73	1.07	4.41	7.5	0.13	7.17
0.5	10	25	0.03	0.26	2.5	1.11	5.44	4.51	0.43	4.54	7.49	0.28	2.58
0.5	20	25	3.15	1.42	2.15	0	1.93	4.37	0.7	4.93	5.71	2.29	2.02
1	5	25	0	0.11	0.96	0	1.91	0.32	0.41	4.21	7.23	1.17	5.58
1	10	25	0.11	0.36	0.39	0.56	2.14	0.83	1.58	3.4	5.85	0.09	1.87
1	20	25	0.9	0.93	0.7	1.26	0.72	2.4	1.33	1.27	4.38	1.57	3.75
0.1	5	50	0	0	3.91	0	0	0	1.55	2.74	3.66	0.14	8.61
0.1	10	50	0	1.58	2.27	0	1.79	0	0.64	0.18	4.39	1.41	20.8
0.1	20	50	0	4.71	5.75	0	8.15	8.69	0.81	3.93	1.74	0.91	2.24
0.5	5	50	0.13	2.4	1.84	1.7	4.09	3.6	1.48	6.38	12	0.73	10.01
0.5	10	50	0.26	2.46	3.42	1.1	5.06	1.58	1.79	3.42	7.99	0.93	4.43
0.5	20	50	0.08	2.04	3.15	0.13	4.59	2.86	1.45	6.14	6	0.9	4.69
1	5	50	0.61	0.65	1.42	0.22	3.06	0.91	2.27	8.43	9.97	4.6	10.96
1	10	50	0.26	1.48	1.95	1.05	2.36	3.21	1.59	4.08	7.95	0.19	6.7
1	20	50	0.23	1.55	0.64	0.48	2.86	2.9	0.76	3.65	5.29	0.05	6.23
Average			2.1	2.67	2.78	1.57	3.46	2.88	2.15	3.28	5.67	1.66	6.54

6.5.2. Numerical Results

The results are divided into two parts. First, we present the results of solving the SRP with non-stationary demand using the RDP. Then we focus on the SRPSS. There, we first present the comparative analysis between the CSS and DSS approaches and then we examine the efficacy of the RDP for the problem.

Table 6.6 presents the optimality gap of the RDP on the SRP problem with non-stationary demand. The overall average gap is 3.2%, which outperforms the OH-IECA (in Chapter 5) with overall average gap 7.3%. Detailed comparison is depicted in Figure 6.5.

Next, we examine the results for the SRPSS. The cost performances of the CSS and DSS approaches are given in Table 6.7. $\mathcal{C}_{CSS}$ and $\mathcal{C}_{DSS}$ are the expected optimal costs using dynamic programming. We observe some difference in the optimal costs between those approaches for the chosen problem parameters. In all the cases, the DSS approach either outperforms the CSS approach or performs identically. The average gain is 0.43%.

In the second part, we present the results using our proposed RDP method in Table 6.8 for the same problem instances. We get major gain in terms of runtime. For a planning horizon of 20 periods, DP takes on average 2821 seconds while with RDP, the average runtime 0.006 seconds. The average gaps are 3.93% and 2.92% respectively for the CSS and DSS approaches. In all the instances the DSS approach outperforms the CSS approach, with average gain of 1.41%.

When the suppliers are equivalent in terms of two of the three parameters, both approaches yield equal cost. Only when suppliers are different in at least two parameters, do we observe a difference in costs. The results from Table 6.7 show that the cost difference is higher when the input costs differ more across suppliers.

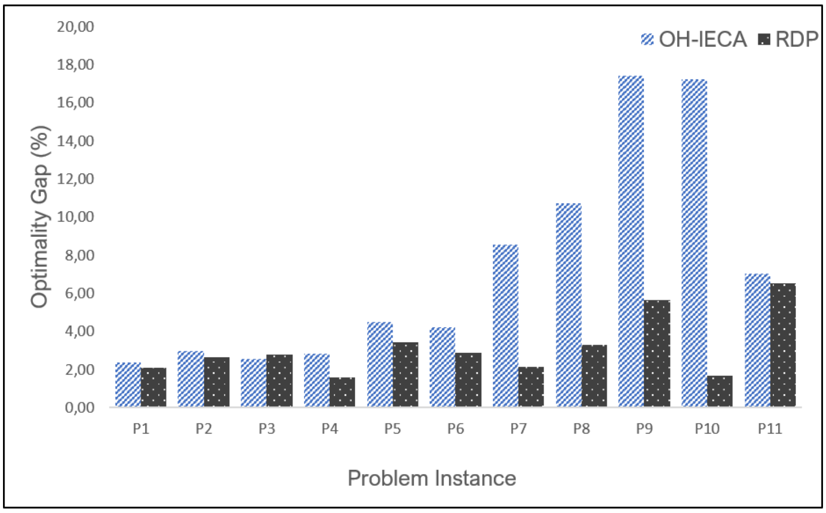


Figure 6.5: Comparative performances of OH-IECA and RDP approaches for the SRP with non-stationary demand. The average optimality gap for OH-IECA is 7.3% and that of SDP is 3.2%.

Table 6.7: Comparison between CSS and DSS approaches for the SRPSS using DP.

H	W	K_1	K_2	R_1	R_2	MOQ_1	MOQ_2	C_{CSS}	C_{DSS}	Gain
1	15	50	50.00	10	10.00	0	0	1469.95	1469.95	0.00 %
1	15	50	55.00	10	9.52	0	0	1444.42	1444.24	0.01 %
1	15	50	60.50	10	9.07	0	0	1422.59	1422.09	0.04 %
1	15	50	66.55	10	8.64	0	0	1404.22	1403.32	0.06 %
1	15	50	73.21	10	8.23	0	0	1388.77	1387.39	0.10 %
1	15	50	80.53	10	7.84	0	0	1375.79	1373.72	0.15 %
1	15	50	88.58	10	7.46	0	0	1365.35	1362.67	0.20 %
1	15	50	97.44	10	7.11	0	0	1358.31	1355.24	0.23 %
1	15	50	107.18	10	6.77	0	0	1355.32	1351.72	0.27 %
1	15	50	117.90	10	6.45	0	0	1356.34	1352.29	0.30 %
1	15	50	50.00	10	10	50	50	1656.53	1656.53	0.00 %
1	15	50	55.00	10	10	50	45	1632.36	1629.16	0.20 %
1	15	50	60.50	10	10	50	40	1637.06	1624.00	0.80 %
1	15	50	66.55	10	10	50	35	1613.70	1609.51	0.26 %
1	15	50	73.21	10	10	50	30	1584.32	1583.88	0.03 %
1	15	50	80.53	10	10	50	25	1598.04	1597.80	0.01 %
1	15	50	88.58	10	10	50	20	1618.77	1615.97	0.17 %
1	15	50	97.44	10	10	50	15	1643.15	1619.69	1.45 %
1	15	50	107.18	10	10	50	10	1656.53	1622.84	2.08 %
1	15	50	117.90	10	10	50	5	1656.53	1628.22	1.74 %
1	15	100	100	10	10.00	5	5	1649.55	1649.55	0.00 %
1	15	100	100	10	9.52	5	10	1603.23	1603.23	0.00 %
1	15	100	100	10	9.07	5	15	1560.39	1559.82	0.04 %
1	15	100	100	10	8.64	5	20	1520.39	1518.46	0.13 %
1	15	100	100	10	8.23	5	25	1483.05	1479.79	0.22 %
1	15	100	100	10	7.84	5	30	1451.28	1446.28	0.35 %
1	15	100	100	10	7.46	5	35	1448.06	1432.43	1.09 %
1	15	100	100	10	7.11	5	40	1434.38	1418.86	1.09 %
1	15	100	100	10	6.77	5	45	1406.27	1390.45	1.14 %
1	15	100	100	10	6.45	5	50	1395.62	1384.97	0.77 %
Average										0.43%

Table 6.8: Comparison between CSS and DSS approaches for the SRPSS using RDP.

K_1	K_2	R_1	R_2	MOQ_1	MOQ_2	$\hat{C}_{CSS}$	Gap	$\hat{C}_{DSS}$	Gap	Gain
50	50	10	10	0	0	1500.16	2.06 %	1500.16	2.06 %	0.00 %
50	55	10	9.52	0	0	1523.55	5.48 %	1504.48	4.17 %	1.27 %
50	60.50	10	9.07	0	0	1510.16	6.16 %	1481.58	4.18 %	1.93 %
50	66.55	10	8.64	0	0	1470.87	4.75 %	1464.41	4.35 %	0.44 %
50	73.21	10	8.23	0	0	1451.49	4.52 %	1447.68	4.35 %	0.26 %
50	80.53	10	7.84	0	0	1452.5	5.58 %	1403.26	2.15 %	3.51 %
50	88.58	10	7.46	0	0	1430.02	4.74 %	1427.28	4.74 %	0.19 %
50	97.44	10	7.11	0	0	1436.12	5.73 %	1408.21	3.91 %	1.98 %
50	107.18	10	6.77	0	0	1404.84	3.65 %	1397.93	3.42 %	0.49 %
50	117.90	10	6.45	0	0	1414.45	4.28 %	1394.39	3.11 %	1.44 %
50	50	10	10	50	50	1669.20	0.76 %	1669.2	0.76 %	0.00 %
50	55	10	10	50	45	1672.75	2.47 %	1666.34	2.28 %	0.38 %
50	60.50	10	10	50	40	1669.29	1.97 %	1644.59	1.27 %	1.50 %
50	66.55	10	10	50	35	1658.44	2.77 %	1649.99	2.52 %	0.51 %
50	73.21	10	10	50	30	1630.89	2.94 %	1612.91	1.83 %	1.11 %
50	80.53	10	10	50	25	1657.62	3.73 %	1638.48	2.55 %	1.17 %
50	88.58	10	10	50	20	1689.76	4.39 %	1689.59	4.56 %	0.01 %
50	97.44	10	10	50	15	1696.26	3.23 %	1649.46	1.84 %	2.84 %
50	107.18	10	10	50	10	1677.77	1.28 %	1644.19	1.32 %	2.04 %
50	117.90	10	10	50	5	1681.25	1.49 %	1660.21	1.96 %	1.27 %
100	100	10	10	5	5	1663.79	0.86 %	1663.79	0.86 %	0.00 %
100	100	10	9.52	5	10	1659.93	3.54 %	1647.90	2.79 %	0.73 %
100	100	10	9.07	5	15	1705.55	9.30 %	1669.93	7.06 %	2.13 %
100	100	10	8.64	5	20	1636.56	7.64 %	1614.79	6.34 %	1.35 %
100	100	10	8.23	5	25	1572.02	6.00 %	1544.63	4.38 %	1.77 %
100	100	10	7.84	5	30	1553.86	7.07 %	1479.70	2.31 %	5.01 %
100	100	10	7.46	5	35	1482.09	2.35 %	1447.04	1.02 %	2.42 %
100	100	10	7.11	5	40	1484.47	3.49 %	1475.31	3.98 %	0.62 %
100	100	10	6.77	5	45	1455.33	3.49 %	1395.35	0.35 %	4.30 %
100	100	10	6.45	5	50	1429.03	2.39 %	1403.10	1.31 %	1.85 %
Average							3.93%		2.92%	1.41%

6.6. Conclusions

In this chapter, we addressed the single-item replenishment problem (SRP) with non stationary demand and the integrated replenishment planning and supplier selection (SRPSS) problem. The first problem can be formulated as a multi-stage stochastic problem. We extended the works on coverage period cost from Chapter 5 to formulate a dynamic program with smaller state space. The proposed method was tested for full backorder scenario with average optimality gap of 3-4%. The proposed method can also be adapted for the extensions of the basic problem to address lost sales and batch sizes.

The second problem is also a multi-stage stochastic program. Due to the curse of dimensionality, it is intractable for medium to large size problems. For its practical importance and complexity, it has received considerable attention in the literature post 2003, however, mostly as a measure of risk mitigation. Nowadays, multi-brand retailers face the problem during their day to day operation. This gives rise to the need of its study as an economic option and the development of faster optimization method. We first conducted the financial benefit analysis of dynamic supplier selection versus selecting a common supplier for the planning horizon. Then we proposed an approximation framework for both approaches.

A common supplier for the whole planning horizon is a practically more appealing feature. However, the dynamic supplier selection results in higher economic benefits. Both of the aforementioned problems are multi-stage stochastic optimization problems. However, the latter one is relatively more complex due to its increased number of possible actions. Numerical analysis suggest that the dynamic supplier supplier selection approach always outperforms the approach with one common supplier, especially when the inventory holding costs and the backorder costs are very different, and when the suppliers impose a minimum order quantity constraints. Finding the optimal solutions of any of the above approaches is time consuming. Hence, we developed an approximation framework based on dynamic programming.

The framework works in two stages. The first stage gives the optimal order quantity and cost for discrete coverage period. We then end up with $\frac{T(T+1)}{2}$ different costs for a rolling horizon of length T. Afterwards, a dynamic programming approach optimizes the total cost for the rolling horizon. We conducted numerical analysis to attest the performance of our proposed method. For the synthetic instances the approach provides near optimal solution at a fraction of the computation time. On average the optimality gap is 3.5%. The average computation time is 6 milliseconds for the RDP.

In industrial context we intend to test the method in lost sale environments and when the suppliers give quantity discounts which are some of the common practices nowadays in retail. Also, deeper analysis can be done to suggest when the economic benefits of a dynamic supplier selection problems outweighs the practical benefits of selecting a common supplier.

In the next chapter, we discuss the replenishment planning problem in the inventory systems with multiple items or joint replenishment problems (JRPs). We focus on the problem during promotions when items are ordered in terms of packs containing multiple items.

Chapter 7

Promotional Replenishment Planning in Multi-item Inventory Systems

Abstract. *In this chapter, we address the replenishment planning problem during retail promotions. The simplest version could be formulated as a single period newsvendor problem. However, inaccurate forecast, limited access to financial information and packaging constraints are some of the major limitations encountered while solving practical problems. We address all of the above three limitations and propose two approaches depending upon the importance and the way to consider those limitations. Thus, first we formulate a multi-objective optimization problem with lost sales and leftover inventory as two major conflicting objectives. We propose a 2-stage method to solve the problem. The first stage defines various service level thresholds through inventory classification. The second stage solves the optimization problem with an ϵ-constraint like method. We propose the use of discrete probabilistic forecast, and compare the results with those obtained using point forecast. The results for a real world problem indicate that both solutions outperform the existing ordering policy and the probabilistic approach outperforms the later. Results from the probabilistic approach show 39% reduction in lost sales and 27% reduction in leftover inventory. We then propose the second approach for practical prudence where, we reformulate the problem to minimize leftover inventory at different service levels. We also propose a metaheuristic to obtain good solutions to real-world size problems in real time. Initial parts of this work (*Sahu et al., *2018*) have been presented in the international conference on information systems, logistics and supply chain (ILS), 2018.*

7.1. Introduction

Nowadays, retailers are conducting frequent promotional events to improve their sales. These events are aimed at increasing footfalls and sales volume, however, they can also pose greater challenges towards inventory management. This can be attributed to multiple factors such as the volume of operation which can be many-fold of the normal level and surge in item variety.

Two factors contribute the most to inventory control performance: forecast accuracy and inventory optimization model. Even though an improved forecast accuracy can contribute substantially towards reducing inventory and logistical costs, accurate point forecast remains a challenge today. Forecasting the demand during promotions comes with even more uncertainty. Predicting the right demand at each retailer is crucial for the success of every retailing company because it helps towards better inventory management, results in better distribution of items across retailers, and minimizes over and under stocking at each retailer. Thereby, it minimizes losses and most importantly maximizes sales and customer satisfaction Linoff and Berry (2011). Literature addressing ordering problems can be divided into two groups depending upon whether they consider demand to be deterministic or stochastic. In case of stochastic demand, it can be either stationary or non-stationary. However, in practice it becomes very difficult to express the demand process as a standard distribution accurately (Ren *et al.*, 2017; Rahdar *et al.*, 2018). We propose to use probabilistic forecast and express the demand as a set of discrete scenarios.

During sales promotions, the replenishment planning can pose challenges to inventory management as it can be constrained by the planning approach, availability of information and distribution constraints. Due to geographical limitations retailers often prefer decentralized planning. However, planners at different locations have limited access to central financial data regarding item costs, storage costs etc. Hence, inventory optimization with cost parameters becomes impractical. In practice, supplier and retailers often deploy single-item and multi-item *prepacks*. A *prepack* can be defined as a collection of items used in retail distribution. The purpose is to facilitate the distribution process by grouping multiple units of one or more items. This also reduces the distribution costs (Gao *et al.*, 2014b). This however makes the replenishment planning more complex and limits flexibility.

With all the challenges described above, the objective of replenishment planning (during sales promotions) is to balance between inventory and service level. While higher inventory can result in higher service level at higher holding cost, lower inventory can lead to lost sales at lower holding cost. All planning decisions are made around 6-12 weeks in advance and orders are placed at the suppliers. Therefore, the retailers can effectively place only one order for one promotion. This problem is similar to the single period newsvendor problem (Arrow *et al.*, 1951). If the retailer orders less than the demand, then it loses sales. On the contrary excess order can lead to higher inventory. In case of perishable items, the salvage value is also lower. Non-perishable items can cause higher holding cost. Specifically during promotions, improper order placement can cause customer dissatisfaction and excess holding cost. In the context of retail sales promotions, our contributions are as follows.

1. We propose a mathematical model for promotional ordering with non-stationary demand, limited financial information, and packaging constraints.

2. We propose a 2-stage method to solve the promotional ordering problem.

3. We propose an additional formulation and a metaheuristic adapted to real-world promotional problems.

The upcoming sections are arranged as follows. In the Section 7.2, we describe the problem and introduce some business rules for practical usability. A brief literature review is given in Section 7.3. Then, in Section 7.4, we formulate the multi-objective optimization problem. We describe our methodology to solve the ordering problem in Section 7.5. We present a real-world case study in Section 7.6 to validate our approach. The reformulated problem and the corresponding metaheuristic are proposed in Section 7.7. Its numerical analysis and another case study are presented in Section 7.9 and Section 7.10. At last in Section 7.11, we summarize the main conclusions.

7.2. Problem Description

The promotional ordering problem may be seen as an extension of the classical newsvendor problem. The major differences can be expressed in three groups that will be detailed afterwards. They are: 1. Demand consideration, 2. Availability of financial information and 3. Presence of multiple prepack constraints.

The classical newsvendor problem assumes a parametric distribution of demand. In practice however, the exact probability distribution of the demand process remains unknown. Inventory planning is done with the information about forecast. It is inherently assumed that the probability density function of the demand and the forecast are the same. This leaves room for the forecast error to be a major challenge to inventory optimization. A second fact is that it is often very difficult to express the actual demand distribution as a standard one. In this chapter, we express the forecast as a set of scenarios having definite probability. We propose to use the service level in quantity context (β-service level) (Beyer *et al.*, 2016; Bowersox *et al.*, 2002), i.e. service level is the percentage of demand satisfied.

The second difference is about availability of relevant financial information. In general, during inventory planning financial information such as buying prices, profitability of items, inventory holding costs, shortage costs and salvage values etc. are considered. But in practice obtaining these information for a complete set of new items having customized packaging is often impossible or very tedious. In case of decentralized planning, financial information is often not available to the planner.

The third difference is regarding packaging types for the items. During retail sales promotions orders are placed as multiple of specific prepacks due to economies of scale and packaging constraints at the suppliers. This also decreases the distribution costs. Selection of best prepacks can be formulated as a mixed integer linear problem (MILP). It is analogous to the classical Knapsack problem which is well known to be $\mathcal{NP}$-hard. The presence of these constraints in the newsvendor problem makes it even more difficult.

The problem is motivated by a real situation in a multi-national retail firm. It considers promotional ordering across multiple retailers of a supply chain network with single supplier as shown in Figure 7.1. The supplier does not have any capacity constraint. The sales forecasts are generated centrally and communicated to the retailers. The retailers then validate the forecast and plan their replenishment accordingly. The firm follows a decentralized inventory planning approach where planning for one or a set of retailers is done by one planner depending upon geography. Multiple orders are not allowed. Financial information regarding item costs, storage costs and salvage values is not available to the planner. The retailer sells multiple items and each item can be ordered through any one or more of available the packing choices. There are 3 distinct packing types for each item, and the packing choices available can be one or many of those. The 3 distinct packing (see Figure 4.7) types are:

T1. Individual units;

T2. Single-item prepack, which contains only one type of item;

T3. Multi-item prepack, which contains more than one type of items.

Business Rules. Although a global optimization can give the best result for a set of input conditions, the output may not be practically feasible. Therefore, to make the output more realistic and acceptable for practical use, we formulate certain business rules in consultation with the retail firm. These business rules are then included in the optimization model. We will

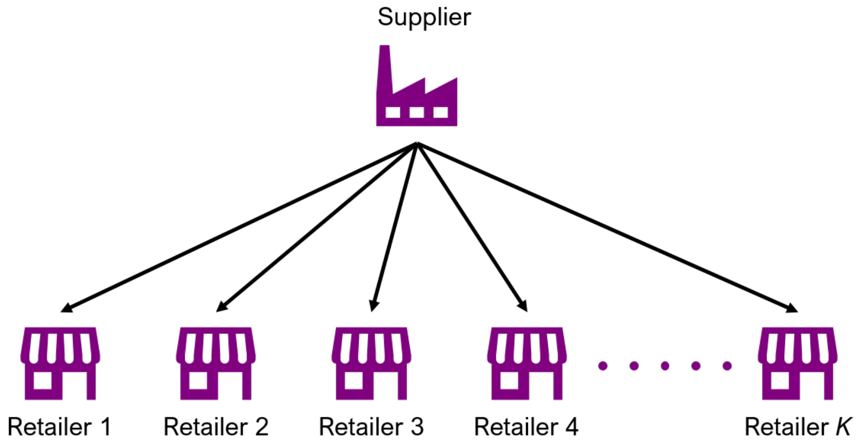


Figure 7.1: Supply chain network structure under study.

explain the incorporation of the rules in the next section. The business rules are mentioned below.

R1. Larger packagings are easier to handle, therefore should be given priority.
R2. All critical items must meet their respective minimum service level criterion.
R3. If an item has non-zero forecast it must be ordered in non-zero quantities.
R4. Maximum number of pack types that can be ordered per item per retailer is Γ.

Problem Class

The problem we consider has a single-echelon supply chain network, deterministic lead time and stochastic non-stationary demand. Excess demand is lost completely and the retailer must satisfy a minimum service level. We have multiple objectives, a dynamic inventory policy and multiple lot sizes. Also, multiple items are ordered from a single supplier in a single period ordering. This basic problem can be denoted using the nomenclature given in Chapter 2:
$\texttt{1;}\phi\texttt{||D;D||N;L;Y||M||D;M||J;S;S}$

This basic structure can have extensions, such as with back order, single batch size or a single objective. For example:
$\texttt{1;}\phi\texttt{||D;D||N;}\textit{B}\texttt{;Y||M||D;M||J;S;S}$,
$\texttt{1;}\phi\texttt{||D;D||N;L;Y||M||D;}\textit{S}\texttt{||J;S;S}$ and
$\texttt{1;}\phi\texttt{||D;D||N;L;Y||}\textit{C}\texttt{||D;M||J;S;S}$.

7.3. Literature Review

The promotional ordering problem involving multiple items is similar to the multi-product newsvendor (MPNV) problem for which a large amount of literature exists. The MPNV problem determines the optimal order quantities (or α-service level) for all items subject to certain constraints. It was introduced in Hadley and Whitin (1963) with a single budget constraint. More recently, Chernonog and Goldberg (2018) solve the MPNV problem with bounded demand. The problem we consider does not focus on finding the optimal α-service level for each product based on inventory holding cost and shortage cost, instead, it focuses on optimal combinations of prepacks for a deterministic demand that is a function of the service level. Presence of multiple items in one ordering unit has been addressed as product bundling (Sheikhzadeh and Elahi, 2013) and prepack (Gao *et al.*, 2014a) in the literature. While the prime objective of product bundling is to provide a discounted price of one item based on order quantity of other items, the use of prepacks primarily aims at minimizing the logistics costs. Product bundling also considers the possibility of the items also being sold as bundles in contrast to unit-wise selling of items in prepacks. Product bundling with price discount is considered in Rosenthal *et al.* (1995), who provide a mixed integer linear programming (MILP) approach for vendor selection. McCardle *et al.* (2007) provide analytical results on the optimal bundle prices, order quantities, and profits for deterministic demand. Ernst and Kouvelis (1999) study the optimal structure of stocking policy in the single period newsboy type problem in the presence of product bundles. However, their study is limited to selection of only one type of bundle for the ordering decision. They also assume that the bundles are sold to the customer in as-is condition. Gao *et al.* (2014a) study the optimal policy structure in the multi-period case, where items can be bought as units and/or prepacks and can be sold as units. They also consider only one type of prepack. Our research varies from the above in many aspects. We aim to address the situation where multiple prepacks for the same item can be ordered. One item may or may not be ordered by units. The actual choice depends on the suppliers' offerings.

We also develop an iterated local search (ILS) (Lourenço *et al.*, 2003) approach to solve the problem. Our solution approach encounters a general multidimensional knapsack problem (MDKP) (Puchinger *et al.*, 2010) in each iteration, and even the optimal solution of the MDKP included cannot ensure the optimal solution to the initial problem unless some conditions are satisfied (refer to Section 7.7.2). Even though the 0-1 MDKP has been extensively studied, studies on the general MDKP are scarce (Akçay *et al.*, 2007). Due to the nature of our problem and practical limitations, we only focus on literature which gives timely solutions. Akçay *et al.* (2007) provide a primal effective capacity heuristic (PECH) for the problem, which is greedy by principle. They also introduce a parameter to adjust the greediness. Although no clear relationship exists between the greediness and quality of the solution, a more greedy approach provides faster solutions. The PECH outperforms the two best known algorithms, the genetic algorithm in Chu and Beasley (1998) and the LP-based algorithm in Pirkul and Narasimhan (1986) in terms of computational effort. In our problem however, the final quality of the solution depends on the perturbations and the local search. Good perturbations are required to ensure the possibility that the optimal solution can be obtained from the incumbent solution, and an effective local search is required to obtain good solutions.

7.4. Mathematical Formulation

Probabilistic Forecast. In the classical single period newsvendor problem, demand is expressed as a random variable with a probability density function. However, in practice it is often difficult to express the demand probability density function accurately. In case of the multi-period newsvendor problem several works focus on demand estimation. Bensoussan *et al.* (2007) use dynamic programming and probability theory, Levi *et al.* (2007c) use monte-carlo simulations, and Levina *et al.* (2010) use online learning method. Kim *et al.* (2015) express non-stationary demand for multi-period newsvendor problem as discrete scenarios. We use the same approach, but for single-period model. The uncertainty is discretized into a set of scenarios, which denote the possible realizations of the random event and its

Table 7.1: Parameters and variables for the optimization model.

Sets	
$\mathcal{K}$	Set of retailers, $k \in \mathcal{K} = \{1, ..., K\}$
$\mathcal{N}$	Set of items, $i \in \mathcal{N} = \{1, ..., N\}$
$\mathcal{M}$	Set of prepacks, $m \in \mathcal{M} = \{1, ..., M\}$
$\mathcal{Z}$	Set of scenarios, $z \in \mathcal{Z} = \{1, ..., Z\}$
Parameters	
SL_{ki}	Minimum required service level at retailer k for item i
P_z	Occurrence probability of scenario z
F_{kiz}	Demand forecast at retailer k for item i in scenario z
R_{im}	Quantity of item i contained in prepack m
r_{im}	1 if item i contained in prepack m, 0 otherwise
Decision Variables	
Q_{km}	Order quantity at retailer k for prepack m
Internal Variables	
q_{ki}	Order quantity at retailer k for item i
I_{kiz}	Leftover inventory at retailer k for item i in scenario z
D_{kiz}	Demand fulfilled at retailer k for item i in scenario z
A_{km}	1 if order quantity at retailer k for prepack m is non-zero, 0 otherwise

corresponding probability. Scenario-based models make stochastic constraints into regular constraints (Kim *et al.*, 2015). We generate a forecast for each retailer-item combination. $\bar{F}_{ki}$ defines the point forecast for retailer k and item i. The probabilistic forecast is described by a set of discrete scenarios $\mathcal{Z}$ for each retailer-item combination. The number of scenarios for each retailer-item is Z and $z \in \mathcal{Z} = \{1, ..., Z\}$ describes the index of scenarios. P_z defines the probability of occurrence of scenario z and F_{kiz} the expectation of d_{ki} under scenario z, i.e. $Pr(d_{ki} = F_{kiz}) = P_z$, where d_{ki} is the demand at retailer k for item i.

Optimization Model. We develop the optimization model for a set of retailers. The model has three objectives. They are:

O1. To minimize lost sales;
O2. To minimize the leftover inventory;
O3. To minimize the total number of prepacks ordered.

Objective O3 reflects business rule R1. Indeed, minimizing the total number of prepacks will force the model to choose larger prepacks. Table 7.1 contains the parameters and variables for the optimization model.

Objectives

The first objective is to minimize total lost sales. Let $\mathcal{Q}$ represent the set of feasible solutions and $\mathbf{Q}$ be the vector representing the decision variables Q_{km}, then the first objective is

$$\min_{\mathbf{Q}\in\mathcal{Q}} O_1(\mathbf{Q}) = \sum_{k\in\mathcal{K}} \sum_{i\in\mathcal{N}} \sum_{z\in\mathcal{Z}} P_z(F_{kiz} - D_{kiz}) \tag{7.1}$$

The second objective is to minimize the leftover inventory

$$\min_{\mathbf{Q}\in\mathcal{Q}} O_2(\mathbf{Q}) = \sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{N}}\sum_{z\in\mathcal{Z}} P_z I_{kiz} \tag{7.2}$$

The third objective is to minimize the total quantity of prepacks ordered

$$\min_{\mathbf{Q}\in\mathcal{Q}} O_3(\mathbf{Q}) = \sum_{k\in\mathcal{K}}\sum_{m\in\mathcal{M}} Q_{km} \tag{7.3}$$

subject to

$$q_{ki} = \sum_{m\in\mathcal{M}} Q_{km}R_{im} \qquad \forall k, \forall i \tag{7.4}$$

$$I_{kiz} + D_{kiz} = q_{ki} \qquad \forall k, \forall i, \forall z \tag{7.5}$$

$$D_{kiz} \leq F_{kiz} \qquad \forall k, \forall i, \forall z \tag{7.6}$$

$$q_{ki} \geq SL_{ki}\sum_{z\in\mathcal{Z}} F_{kiz}P_z \qquad \forall k, \forall i \tag{7.7}$$

$$A_{km} = \begin{cases} 1 \text{ if } Q_{km} > 0 \\ 0 \text{ if } Q_{km} = 0 \end{cases} \qquad \forall k, \forall m \tag{7.8}$$

$$q_{ki} \begin{cases} \geq 1 \text{ if } \sum_{z\in\mathcal{Z}} P_z\bar{F}_{ki} > 0 \\ \geq 0 \text{ if } \sum_{z\in\mathcal{Z}} P_z\bar{F}_{ki} = 0 \end{cases} \qquad \forall k, \forall i \tag{7.9}$$

$$\sum_{m\in\mathcal{M}} r_{im}A_{km} \leq \Gamma \qquad \forall k, \forall i \tag{7.10}$$

$$Q_{km} \geq 0 \qquad \forall k, \forall m \tag{7.11}$$

$$q_{ki} \geq 0 \qquad \forall k, \forall i \tag{7.12}$$

$$I_{kiz} \geq 0 \qquad \forall k, \forall i, \forall z \tag{7.13}$$

Constraint 7.4 is a material balance equation indicating the equivalent purchase of items from order quantity of prepacks. Constraints 7.5 and 7.6 are used to calculate the leftover inventory and demand fulfilled at a given order quantity. Constraint 7.7 indicates that all items must meet the minimum service level criteria. This is required to follow business rule R2. In the following Section 7.5, we explain the details of inventory classification and minimum service level allocation. Constraint 7.8 is for calculating the binary variable A_{km} indicating whether the prepack m is ordered for retailer k. Constraint 7.9 ensures that the order quantities for items having non-zero forecast are non-zero (See Business rule R3). These two constraints 7.8 and 7.9 may be easily linearizable, but above presentation is easy to understand. Constraint 7.10 limits the number of prepacks ordered per item per retailer to Γ (See Business rule R4). Constraints 7.11, 7.12 and 7.13 are non-negativity constraints. The optimization model is denoted as M-0.

Also, it is noteworthy that, here we define the leftover inventory as the excess quantity above the forecast quantity. The phenomenon is depicted in Figure 7.2. Any order quantity q_i for item i will constitute leftover inventory if it is more that the forecast quantity F_{kzi}. This is justified because of the use of Type-2 service level (β service level). In this context, it is assumed that any order quantity up to the forecast quantity can be sold, and it will not create any leftover inventory after the promotion. Therefore, we will obtain excess inventory if and only if the order quantity is more than the forecast quantity.

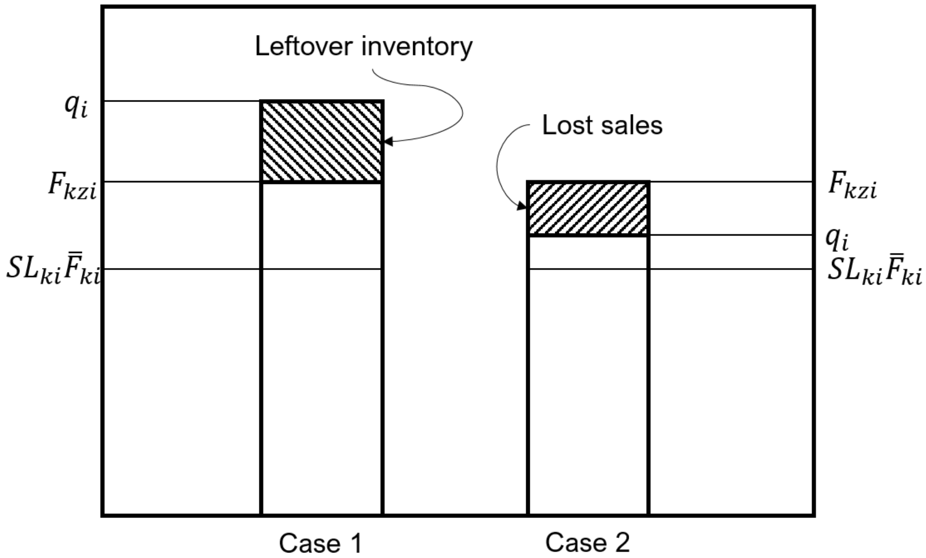


Figure 7.2: Illustration of the nature of inventory computation. In Case 1, the order quantity q_i for an item i is greater than its forecast F_{kiz}. It results in excess inventory $q_i - F_{kiz}$ as indicated in the shaded porting. This is because, in such cases 100% demand can be fulfilled. However, if the order quantity is less than or equal to the forecast quantity then, the excess inventory is zero as a maximum demand equal to the order quantity can be fulfilled. On the contrary, in Case 2, the shortfall from the forecast indicates the lost sales even if it fulfills the service level constraint.

7.5. Methodology

To deal with this promotional replenishment planning problem, we propose a methodology consisting of two stages:

1. Inventory classification and Minimum service level determination;
2. Multi-Objective optimization with probabilistic forecast.

7.5.1. Inventory Classification and Minimum Service Level Allocation

Inventory classification plays an important role in inventory management. It helps in defining inventory priorities clearly. The ABC and VED classifications are widely accepted. In our approach, we propose a mix of both approaches to classify the items depending upon individual forecast quantity and criticality. After classification, we define a threshold service levels based on inventory classes. It will be used as the minimum required service level in the optimization problem.

General ABC classification (Torabi *et al.*, 2012; Millstein *et al.*, 2014) is based on volume. We demonstrate an example in Figure 7.3. The high volume items contributing upto 50% of the cumulative volume are classified as Class-A items. The next high volume items contributing upto 80% of the cumulative volume constitute Class-B items. The remaining items fall in Class-C. While ABC classification is a quantitative approach, VED classification (Gupta *et al.*, 2007; Molenaers *et al.*, 2012) is qualitative. The criticality of the items is defined by the retail firm (Business rule R2). Items in the “Vital” class are most critical. The “Essential” class contains items which are relatively less critical, and the least critical items are grouped into the “Desired” class.

We use ABC and VED classifications simultaneously to exploit the benefits of both ap-

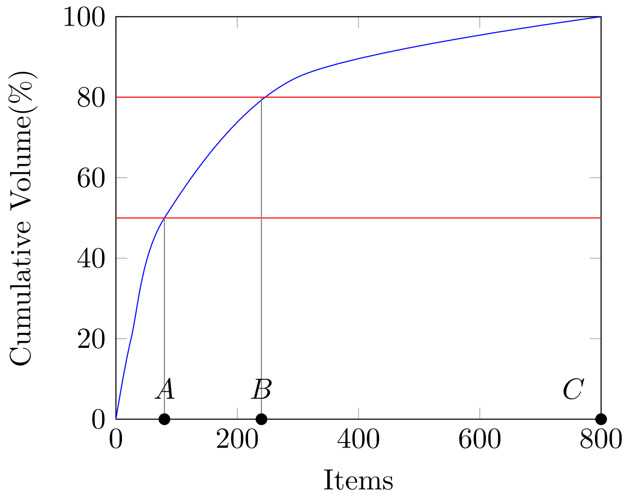


Figure 7.3: Illustration of ABC classification. X-axis represents items, and Y-axis represents cumulative volume after the items are sorted in descending order of sales volume. The interval from the origin to point A on the X-axis represents high volume items constituting 50% volume, i.e. class-A items. The interval from point A to point B on the X-axis represents the next high volume items constituting from top 50% to 80% volume, i.e. class-B items. All the remaining items belong to class-C.

proaches. It is logical for a retailer to achieve maximum service level for class-A and vital items. Therefore, for our case study we define certain minimum service level thresholds for these classes of items. Table 7.2 contains the minimum service level threshold for different classes of items used in our case study. For items which are not in class-A and are not Vital, there are no minimum service level thresholds. The service levels for them are decided by the optimization problem.

Table 7.2: Minimum service level thresholds for different classes of items. In this example, 95% of minimum service level is required for Class-A and Vital items. Differentiated service levels are used for other classes.

	ABC Class		
VED Class	A	B	C
Vital	95%	90%	85%
Essential	90%	85%	80%
Desired	85%	80%	75%

7.5.2. Multi-Objective Optimization with probabilistic forecast

In Section 7.4, we described the multi-objective optimization model. From the first stage we get the minimum service level, which is required by Constraint 7.7. As objective O_3 is not one of the major objectives, we add objectives O_3 to O_1 and formulate a bi-objective optimization model. To solve the bi-objective optimization problem, we use an ϵ-constrained Deb (2014) like method. This procedure overcomes some of the convexity problems of the weighted sum technique. This involves minimizing one objective and expressing the other objectives in the form of inequality constraints. We limit objective $O_2(\mathbf{Q})$ i.e. total leftover inventory, and minimize $O_1(\mathbf{Q}) + O_3(\mathbf{Q})$. Mathematically we express it as

$$\min_{\mathbf{Q}\in\mathcal{Q}}[O_1(\mathbf{Q}) + O_3(\mathbf{Q})] \tag{7.14}$$

$$\text{s.t. } O_2(\mathbf{Q}) \leq \epsilon \tag{7.15}$$

By solving 7.14 for different values of ϵ we obtain several Pareto optimal solutions. The suitable operating point can be selected by the user depending upon their lost sales and leftover inventory targets. Figure 7.4 describes the whole methodology through a flowchart.

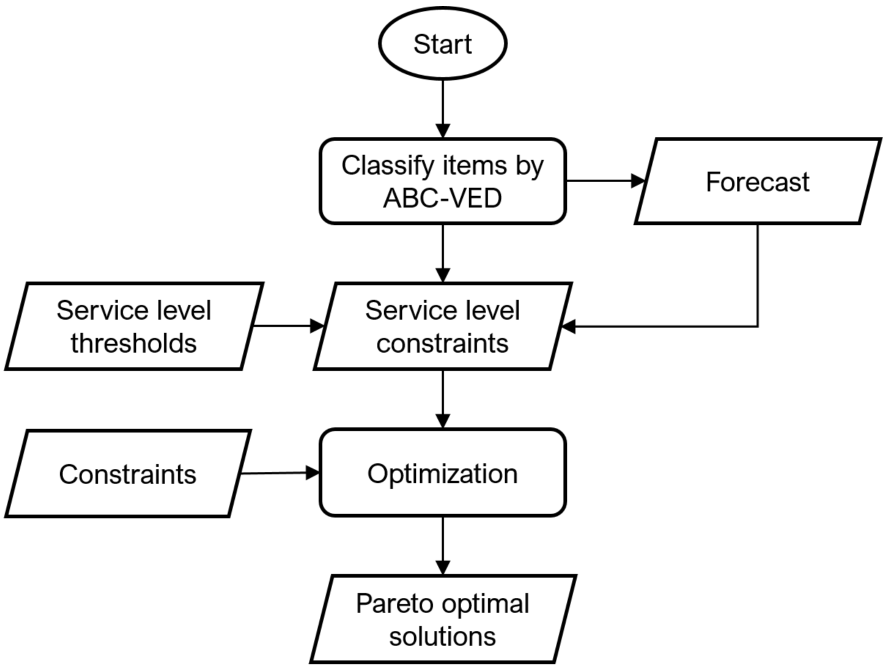


Figure 7.4: Flowchart illustrating the proposed methodology. Inputs are forecast and packaging constraints. ABC-VED classification is used to define minimum service level. The problem is formulated as a mixed integer linear programming (MILP) model, and an ϵ-constrained like method is used to generate Pareto optimal solutions.

7.6. Case Study 1

7.6.1. Input Data and Experimental Setting

To test our methodology, we use the promotional data of a multi-national firm which needs to control its inventory at the retailers. The firm sells a variety of goods during a promotional event, which can either be completely new items or regular items. We summarize the components of the firm's supply chain in Table 7.3. There are no capacity constraints at the suppliers. There are multiple packaging choices for a single item. The firm can order only once for each promotion.

Table 7.3: Supply chain components.

Distribution structure	
Number of retailers	106
Promotion components	
Number of items	717
Number of prepacks	1202

Table 7.4: Problem size and computational time

Integer Variables	
Deterministic, 106 stores	406,828
20 scenarios, 106 stores	1,850,866
Computation time	
Deterministic, 106 stores	18 Minutes
20 scenarios, 106 stores	74 Minutes

We use the past sales data and promotional characteristics from all items to compute two types of forecast. First a point forecast, which is the single valued expectation of demand. Let us name the approach of using the point forecast in optimization as *deterministic approach*. Second is a probabilistic forecast described by a discrete distribution. Then a number of distinct scenarios are extracted from the probabilistic forecast. Let us name the approach of using probabilistic forecast for optimization as *probabilistic approach*. For the discussed case, we consider 20 scenarios. In each scenario the demand forecast is non-negative. The value of Γ in Equation 7.10 (see Business rule R4) is set equal to 2 in the studied case. As explained in the Section 7.5, we use an ϵ-constraint like method to solve the multi-objective optimization problem. Therefore, we constrain $O_2(\mathbf{Q})$ at different values and optimize $O_1(\mathbf{Q})+O_3(\mathbf{Q})$ using CPLEX 12.7.1 with 8GB memory. The details of the problem size and the computational time are summarized in Table 7.4. The service level thresholds are presented in Table 7.5.

Table 7.5: Service level thresholds for different class of items.

	ABC Class		
VED Class	A	B	C
Vital	95%	95%	95%
Essential	95%	None	None
Desired	95%	None	None

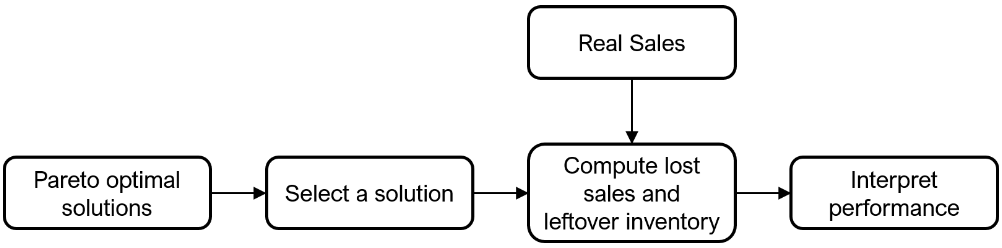


Figure 7.5: Flowchart illustrating the method used to evaluate the quality of solutions. We compare each solution with the corresponding real sales. If the order quantity obtained from the solution is higher than the real sales then it results in leftover inventory and if it is lower than the real sales then it results in lost sales.

Hence, from the optimization we obtain several Pareto optimal solutions in both the deterministic and probabilistic approaches. Each solution provides the suitable order quantity for the given leftover inventory constraint. In order to evaluate the quality of a solution, we use the actual sales data. Figure 7.5 illustrates the methodology followed to evaluate the solutions. Let us represent real sales for retailer k and item i as G_{ki}. If the order placed q_{ki} is greater than or equal to the real sales G_{ki}, then the lost sales is zero and the leftover inventory is the difference between order quantity and sales quantity, $q_{ki} - G_{ki}$. On the contrary, if the order quantity q_{ki} is less than the real sales G_{ki}, then the lost sales is calculated as the difference between real sales and order quantity, $G_{ki} - q_{ki}$. The leftover inventory in this case is zero. Total lost sales and total leftover inventory is obtained by summing each of these quantities for all retailers. For each Pareto optimal solution the values of total lost sales and total leftover inventory are calculated and plotted in Figure 7.6. From the plot leftover inventory vs lost sales, the performance is assessed. We choose to plot only $O_1(\mathbf{Q})$ and $O_2(\mathbf{Q})$ because of their higher importance than $O_3(\mathbf{Q})$ and to obtain a 2D representation (easier to read). The third objective which reflects business rule R1, will be discussed in Table 7.6.

7.6.2. Results and Interpretation

Figure 7.6 plots the total lost sales and leftover inventory across all retailers. Our objective is to minimize both. The circles indicate performance of the solutions obtained using the deterministic approach and the triangles indicate the performance of solutions obtained using the probabilistic approach. It is observed from Figure 7.6 that, each of the triangles outperforms at least one circle on both lost sales and leftover inventory. Moreover, none of the circles outperform any triangle on both lost sales and leftover inventory. Therefore, the probabilistic approach completely dominates the deterministic approach. Another observation that can be made is that, the solutions from the deterministic and probabilistic approaches perform equivalently (or converge) with decreasing leftover inventory. However, with increasing leftover inventory the solutions from probabilistic approach perform better and better.

Table 7.6: Comparison of solution performances.

	Lost Sales			Left-over Inventory			Larger Prepacks(%)	
Retailer	Current	Det. Appr.	Prob. Appr.	Current	Det. Appr.	Prob. Appr.	Det. Appr.	Prob. Appr.
S1	1946	**854**	911	10698	5223	**4928**	91%	91%
S2	1171	316	**235**	4056	2499	**2454**	95%	95%
S3	1456	1056	**899**	6928	4505	**4242**	90%	90%
S4	5602	1738	**1503**	**10185**	12075	10610	**97%**	92%
S5	2066	1526	**1293**	10942	5106	**5059**	**97%**	96%
S6	2973	1949	**1871**	12089	7116	**6919**	**99%**	97%
S7	1805	1215	**1033**	7390	5317	**5075**	**91%**	90%
S8	**710**	1015	882	7626	3724	**3595**	88%	88%
S9	**1574**	2440	2318	16003	**4367**	4395	**96%**	95%
S10	3406	2128	**1614**	16233	**8004**	8051	**96%**	93%
Average	2271	1424	**1256**	10215	5797	**5533**	94%	93%
Gain		37%	**45%**		43%	**46%**		

The performance of the solution currently used by the retailer is indicated by point A in Figure 7.6. Point B represents the ideal point that could be obtained if the exact demand was known in advance. Let us note that even in this case the leftover inventory cannot reach 0 due to packaging constraints. As seen from the figure, the solutions obtained from deterministic and probabilistic approaches perform better than the solution currently used by the retailer. When we compare the order quantity with real sales, deterministic approach gives 43% less lost sales than existing planning for the same inventory level. For the same lost sales level,

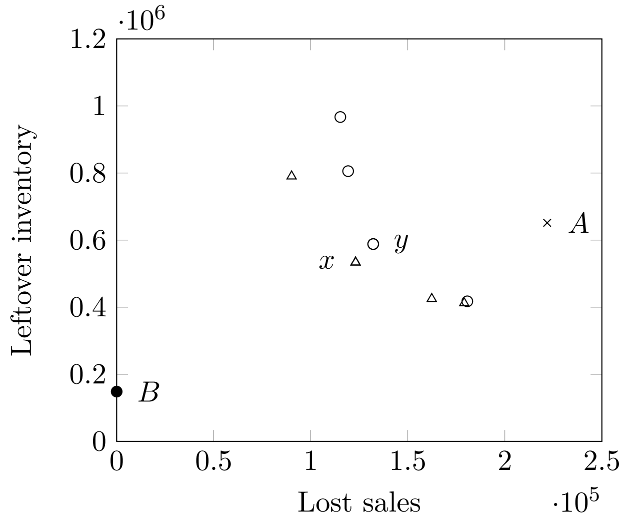


Figure 7.6: Performance of the models. The circles indicate the results from the deterministic approach. The triangles represent the results from the probabilistic approach. Point A indicates performance of the solution currently used by the firm. Point B indicates the ideal solution if the exact demand is known in advance. Results of top 10 retailers presented in Table 7.6 correspond to points x (Prob. Appr.) and y (Det. Appr.).

it gives 36% less leftover inventory. The probabilistic approach gives 55% less lost sales than the existing planning for the same inventory level. For the same lost sales level, it gives 37% less inventory.

For deeper analysis of results, we select the 10 retailers having the highest sales volume and present the values of different objectives in Table 7.6. These results correspond to points x and y in Figure 7.6. The third objective $O3$, which reflects Business rule R1 is presented as the percentage of larger prepacks in the order for better comparison. We focus on the first 2 objectives as they are more relevant financially. From Table 7.6 we can see that for several retailer (S2, S3, S5, S6, S7), the probabilistic approach performs better simultaneously for both, lost sales and leftover inventory. As far as S1, S4, S8 and S10 are concerned, the probabilistic approach performs better for only one of the two criteria, whereas for S9, the two other approaches manage to outperform the probabilistic one.

But the values obtained are close, and on average, the probabilistic approach manages to outperform the others for both criteria. Performance of both approaches are comparable for the third criterion. However, in the probabilistic approach the quantity is a little less. This can be attributed to additional procurement of smaller prepacks to reduce lost sales.

7.7. Problem Reformulation

Previously, we have presented a multi-objective model and a solution method for the promotional replenishment planning. The three objectives: lost sales quantity, $O1$, leftover inventory, $O2$ and number of prepacks, $O3$ were organized into a bi-objective optimization model. A case study of less-than-real problem size was solved using CPLEX 12.7.1. In the following section, we consider real-world problems that can be of larger size. Therefore, we propose another formulation of the problem in which, we drop $O3$ and reformulate $O1$ as a service level constraint. This allows to design a single-objective metaheuristic able to deal with larger size problems.

7.7.1. Optimization Model

We present an alternative MILP model for the problem described in Section 7.2 taking the demand to be equal to the expected demand, $\bar{F}_{ki} = \sum_{z\in\mathcal{Z}} P_z F_{kzi}$. This results in a deterministic optimization problem. The objective is to minimize the total leftover inventory while satisfying the service level. The parameter and variables for the optimization model are the ones already presented in Table 7.1. The optimization problem is formulated as follows.

$$\text{minimize} \sum_{k\in\mathcal{K}} \sum_{i\in\mathcal{N}} I_{ki} \tag{7.16}$$

subject to

$$q_{ki} = \sum_{m\in\mathcal{M}} R_{im} Q_{km} \qquad \forall i \in \mathcal{N}, k \in \mathcal{K} \tag{7.17}$$

$$I_{ki} \geq q_{ki} - \bar{F}_{ki} \qquad \forall i \in \mathcal{N}, k \in \mathcal{K} \tag{7.18}$$

$$q_{ki} \geq SL_{ki}\bar{F}_{ki} \qquad \forall i \in \mathcal{N}, k \in \mathcal{K} \tag{7.19}$$

$$q_{ki}, I_{ki} \geq 0 \qquad \forall i \in \mathcal{N}, k \in \mathcal{K} \tag{7.20}$$

$$Q_{km} \geq 0 \qquad \forall m \in \mathcal{M}, k \in \mathcal{K} \tag{7.21}$$

Constraints 7.17 and 7.19 are equivalent to constraints 7.4, 7.7 respectively. Constraint 7.18 is equivalent to constraints 7.5 and 7.6. Above optimization problem is separable for each retailer since we do not have any global budget constraint. Therefore, we can drop the subscript k for retailers. We refer to Equations 7.16-7.21 as model M-1.

7.7.2. Complexity Analysis

In this section, we determine the complexity of the above optimization problem under 100% service level, i.e. $SL_i\bar{F}_i = \bar{F}_i$. We analyze the process of obtaining the optimal solution of M-1 (with each retailer considered independently) starting from any feasible solution and assess the complexity. Let $\mathbf{Q}$ be a feasible solution so that each of its m-th element Q_m represents the order quantity of prepack m. Also, $\mathbf{Q}$ is chosen such that Q_m is greater than or equal to the optimal solution Q^*_m, $\forall m \in \mathcal{M}$. Then

$$q_i = \sum_{m\in\mathcal{M}} R_{im} Q_m \qquad \forall i \in \mathcal{N} \tag{7.22}$$

$$\hat{I}_i = [q_i - \bar{F}_i]^+ \qquad \forall i \in \mathcal{N} \tag{7.23}$$

where, $[x]^+ = 0$ for $x \leq 0$ and $[x]^+ = x$ for $x > 0$. The effective order quantities q_i can be lower than the forecast $\bar{F}_i$ depending upon the service level SL_i. For cases where $SL_i = 1$, $\forall i$, $q_i \geq \bar{F}_i$.

To reduce the total leftover inventory $\sum_i \hat{I}_i$, we can remove prepacks from the feasible solution $\mathbf{Q}$ (not optimal) in such a way that the modified order quantities for the items q_i do not go below their respective forecast quantities $\hat{F}_i$. Hence, a prepack m can only be removed if the leftover inventory of each of its items is greater than or equal to its pack quantity R_{im}. Let U_m be the number of items contained in prepack m, i.e. $U_m = \sum_{i\in\mathcal{N}} R_{im}$. Then, unit removal of prepack m will reduce the left-over inventory by U_m. Let l_m be the net removal quantity of prepack m. The following optimization problem M-2 maximizes the reduction quantity from the total leftover inventory generated by the current solution $\mathbf{Q}$.

$$\text{maximize} \sum_{m \in \mathcal{M}} U_m l_m \tag{7.24}$$

subject to

$$l_m \leq Q_m \qquad \forall m \in \mathcal{M} \tag{7.25}$$

$$\sum_m R_{im} l_m \leq \hat{I}_i \qquad \forall i \in \mathcal{N} \tag{7.26}$$

$$l_m \geq 0 \qquad \forall m \in \mathcal{M} \tag{7.27}$$

Proposition 1: The optimal solution of M-1 is $\mathbf{Q}^* = \mathbf{Q} - \mathbf{L}^*$, where $\mathbf{Q}$ is a feasible solution such that, each of its elements is greater than or equal to the optimal solution, and $\mathbf{L}^*$ is the optimal solution of M-2.

Proof. The proof is straightforward. The total leftover inventory for the feasible current solution $\mathbf{Q}$ is $\sum_i \hat{I}_i = \sum_i I_i + \sum_m U_m l_j$ by definition, and $\sum_i \hat{I}_i$ is a positive constant for any $\mathbf{Q}$. Maximization of the second term in the right-hand side will ensure minimization of the first term and vice-versa since they satisfy the same set of constraints. Therefore, the optimal solution of M-1 can be obtained by subtracting the optimal solution of M-2 from the current solution $\mathbf{Q}$. ■

To obtain the optimal solution, Q_m must be greater than or equal to Q^*_m. Under such a condition M-1 is as hard as M-2. However, if the condition is not satisfied for the initial feasible solution, we will require solving M-2 in each iteration. Hence, it can be concluded that M-1 is at least at hard as M-2. M-2 is a general multidimensional knapsack problem (MDKP) of M dimensions with U_m as unit rewards and $\hat{I}_i$ as available resources. The general MDKP is strongly $\mathcal{NP}$-hard (Puchinger *et al.*, 2010).

7.8. Metaheuristic Approach

Using the proposed mathematical formulation, solving medium to large size instances requires long computational time. In practice, we encounter problems of enormous size due to the large variety of items. For example, hard discount stores offer only 500 - 600 items (e.g. Aldi), whereas, hypermarkets offer over 100,000 items and department stores offer more than one million items (e.g. Macy's in New York) (Dujak *et al.*, 2017). Although, the retailer does not require to solve the problem very frequently, for large size problems the solution with a commercial solver becomes intractable. Therefore, in this section, we propose a metaheuristic for practical purposes. The proposed metaheuristic has three steps. In the first step, a feasible initial solution is generated using a greedy approach, then, it is improved using a local search. At last, perturbations are applied to the local optimum to move in the search space. The perturbations and local search are repeated alternatively until the stopping criterion is met. The solution to problem M-1 can be expressed as a vector $\mathbf{Q}$ of size M as defined in Section 7.7.2. Each element of the vector $\mathbf{Q}$ corresponds to the order quantity of the respective prepack. An example for the solution is depicted in Figure 7.7.

Q_1	Q_2	Q_3	$\cdots$	Q_m	$\cdots$	Q_M

Figure 7.7: Representation of a solution of problem M-1 with M prepacks as a vector of size M.

7.8.1. Initial Solution

Our methodology requires an initial feasible solution for further improvements. While a naive solution can be obtained by setting very high values for Q_m, this will result in high leftover inventory and further improvements may need a longer computational time. We propose two greedy heuristics to generate good quality initial solutions, GH1 and GH2. The main idea behind GH1 is to select an item randomly, then increase the order quantity of its largest prepack one by one until the service level constraint is fulfilled. It is elaborated in Appendix D. A better heuristic, GH2 (Algorithm 5), is also proposed which had lower complexity than GH1. It is also a greedy heuristic that chooses the prepacks which would create the minimum excess inventory in each iteration. The items are not selected randomly but, by their largest remaining quantities. Also, the prepacks are ordered as per a penalty involving excess inventory.

Algorithm 5: Greedy Heuristic 2 (GH2)

```
Input  : R, F, SL
Output: P
initialization P ← 0, L_i ← SL_i F̄_i, ∀i ∈ N
repeat
    select item î = arg max_i L_i, i ∈ N
    select M* = {m* : R_{îm*} > 0}
    calculate penalty score η_{îm}, ∀m ∈ M*
    find m̂ = arg min_m η_{îm}, m ∈ M*
    calculate increment quantity ζ
    Q_m̂ ← Q_m̂ + ζ
    update q_i, ∀i ∈ N
    L_i ← L_i − q_i, ∀i ∈ N
until L_i ≤ 0, ∀i ∈ N;
```

GH2 works in the following way:

2.0. GH2 generates a feasible solution by taking the prepack configurations, item forecasts and service levels as inputs.

2.1. Initialize by setting the values of order quantity Q_m to zero, and values of remaining quantity L_i their respective service level targets.

2.2. Select the item $\hat{i}$ with maximum remaining quantity and then define the set of all prepacks containing $\hat{i}$ as $\mathcal{M}^*$.

2.3. Compute the penalty score $\eta_{\hat{i}m} = \sum_{i:R_{im}>0}[R_{im} - L_i]^+/R_{\hat{i}m}$ for each prepack in the set $\mathcal{M}^*$.

2.4. Select the prepack having minimum penalty score and increase its order quantity by ζ. The increment quantity ζ is the maximum quantity of the selected prepack that can be ordered without exceeding the remaining forecast quantity of the selected item or one, i.e. $\zeta = \max(1, \lfloor \frac{L_{\hat{i}}}{R_{\hat{i}m}} \rfloor)$.

2.5. Update the order quantities q_i and the remaining quantities L_i.

2.6. Repeat 2.3-2.5 until the remaining quantity for each item reaches zero or becomes negative.

The number of iterations per item is at most two in case of GH2 due to the increment quantity ζ. Hence, the complexity of GH2 is $\mathcal{O}(2N)$.

7.8.2. Local Search

After the initial solution is constructed using GH2, we propose to use a local search methodology for its further improvement. For practical usability, the search space is restricted to feasible solutions space $\mathcal{Q}$ only. This enables us to select a solution at any point in time. Before going into the details of the local search procedure, some notations are introduced. The neighborhood operator is denoted by $\mathcal{V}$ and the neighborhood is by $\mathbf{V}(.)$. $w(.)$ denotes the fitness function that is formulated as follows.

$$w(\mathbf{Q}) = \sum_i \Big[\sum_m R_{im} Q_m - F_i \Big]^+ \tag{7.28}$$

LS1 (Algorithm 6) describes a simple local search procedure. It randomly selects a neighbor solution (detailed description of the neighborhood is given next). If the selected solution satisfies all the constraints and has a better fitness, then the solution is updated and the whole process is repeated until the whole neighborhood of any incumbent solution is explored. LS1 works using the first improvement strategy. The local optimum can be defined as any incumbent solution whose neighborhood does not contain any better solution than itself. Mathematically, the local optimum $\mathbf{Q}^v = \mathbf{Q}$ if $w(\hat{\mathbf{Q}}) \geq w(\mathbf{Q}), \forall \hat{\mathbf{Q}} \in \mathbf{V}(\mathbf{Q})$.

Algorithm 6: Local Search (LS1)

Input : $\mathbf{P}, \mathbf{R}, \mathbf{F}, \mathbf{SL}, \mathcal{V}$
Output: $\mathbf{Q}^v$
1 Initialization $\mathbf{Q}^v \leftarrow \mathbf{Q}$
2 $m \leftarrow \text{Random}[1, |\mathbf{V}(\mathbf{Q}^v)|]$, $\hat{\mathbf{Q}} \in \mathbf{V}(\mathbf{Q}^v)$
3 **for** $i \in \{m, m+1, ..., |\mathbf{V}(\mathbf{Q}^*)|, 1, 2, ..., m-1\}$ **do**
4   **if** $\hat{\mathbf{Q}}_i \in \mathcal{Q}$ & $w(\hat{\mathbf{Q}}_i) < w(\mathbf{Q}^v)$ **then**
        $\mathbf{Q}^v \leftarrow \hat{\mathbf{Q}}_i$
        Repeat 2
    end
end

Description of the Neighborhood

We consider several neighborhood operators. Their respective neighborhoods are as follows.

Case $\mathcal{V} = \{-1\}$: Neighborhood of any incumbent solution $\mathbf{Q}$ is represented in Figure 7.8 for $\mathcal{V} = \{-1\}$. In such cases, a neighbor can be defined with $Q_m \leftarrow Q_m - 1$ for any $m \in \mathcal{M}$. Hence, there is a maximum of M neighbors for any incumbent solution barring any m for which $Q_m = 0$. Zero values in a solution reduce the size of the neighborhood as negative order quantities are infeasible.

Case $\mathcal{V} = \{-1, +1\}$: Similarly, for $\mathcal{V} = \{-1, +1\}$ (Refer to Figure 7.9), a neighbor can be $Q_m \leftarrow Q_m - 1$ and $Q_m \leftarrow Q_m + 1$ for any m. Hence, there is a maximum of $2M$ neighbors.

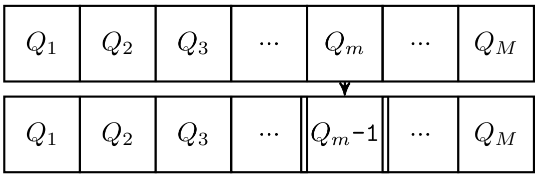


Figure 7.8: Neighborhood representation with $\mathcal{V} = \{-1\}$.

The $+1$ operator is particularly useful in case when $SL_i < 1$ for any i, as it helps exploring a larger search space.

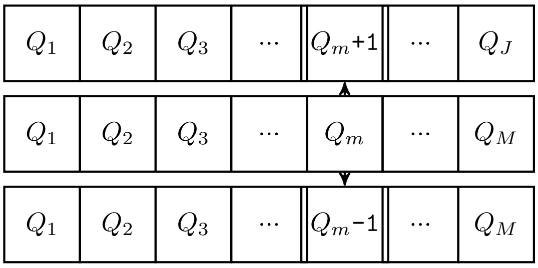


Figure 7.9: Neighborhood representation with $\mathcal{V} = \{-1, +1\}$.

Impact of Service Level

The service levels have a unique impact on the solution methodology due to our leftover inventory calculation procedure. $\mathcal{V} = \{-1\}$ is particularly useful when $SL = 1$ for all items. Such a case results in a situation, where removal of any prepack from an incumbent solution either reduces inventory (improves fitness) or violates the service level constraints. In that case, the neighborhood operator, $\mathcal{V} = \{-1\}$ is very interesting as it automatically ensures better fitness if a neighbor is feasible. In this particular case, as soon as a new feasible solution is found it has to be accepted by a first improve local search (without any need for evaluation), as we are sure that it improves the fitness. Only the local optima need to be evaluated. This makes the local search very efficient. This is, however, not the case when the service level is less than 100% for at least one item. In such cases, removal of any prepack without violating the service level constraints may or may not result in better fitness. This can be explained as follows. When the order quantity of at least one of the items present in the selected prepack is above its forecast quantity, a feasible removal of such prepack will improve the fitness. But, if all items of a prepack are ordered below the forecast quantity, a feasible removal of that prepack will not improve the solution.

7.8.3. Guided Search Heuristic

The local search procedure explained in Algorithm 6 adopts a first improvement strategy. It selects the first neighbor of any incumbent solution randomly and applies a first improvement strategy, i.e. selects the neighbor randomly that improves the fitness. In this section, we propose a "best-max" improvement strategy. It can be explained as follows.

We define "best-max" improvement strategy as selecting the neighbor that provides the maximum fitness improvement then applying the neighborhood operator maximum possible times at that point. For example, if reducing the prepack m in $\mathbf{Q}$ by one unit give the best improvement then, we reduce prepack m until the point it generates infeasible solutions. Afterwards we select the next best neighbor and repeat the process. We refer to the heuristic as the Guided Search Heuristic (GSH).

Algorithm 7: Guided Search Heuristic (GSH)

Input : $\mathbf{Q}, \mathbf{R}, \mathbf{F}, \mathbf{SL}$

Output: $\mathbf{Q}^*$

1 initialization $\mathbf{Q}^* \leftarrow \mathbf{Q}$, $L_m = 0, \forall m \in \mathcal{M}$, $q_i = \sum_m R_{im} Q_m, \forall i \in \mathcal{N}$

2 $\chi_m = Q_m - L_m$

3 $X_i^1 = [q_i - \bar{F}_i SL_i]^+, \forall i \in \mathcal{N}$

$X_i^2 = [q_i - \bar{F}_i]^+, \forall i \in \mathcal{N}$

4 $Y_m^1 = \min_i \left\lfloor \frac{X_i^1}{R_{im}} \right\rfloor, \forall m \in \mathcal{M}, R_{im} > 0$

$Y_m^2 = \min_i \left\lceil \frac{X_i^2}{R_{im}} \right\rceil, \forall m \in \mathcal{M}, R_{im} > 0$

5 $Z_m = \sum_i \min(X_i^2, R_{im}), \forall m \in \mathcal{M}$

repeat

6 $\quad \hat{m} = \arg\max_m \{Z_m | Y_m^2 \chi_m > 0\}$

$\quad a_{\hat{m}} \leftarrow \min(\chi_{\hat{m}}, \max(1, Y_{\hat{m}}^1))$

$\quad L_{\hat{m}} \leftarrow L_{\hat{m}} + a_{\hat{m}}$

7 $\quad X_i^1 \leftarrow [X_i^1 - L_{\hat{m}} R_{i\hat{m}}]^+, \forall i \in \mathcal{N}$

$\quad X_i^2 \leftarrow [X_i^2 - L_{\hat{m}} R_{i\hat{m}}]^+, \forall i \in \mathcal{N}$

$\quad Y_m^1 \leftarrow \min_i \left\lfloor \frac{X_i^1}{R_{im}} \right\rfloor, \forall m \in \mathcal{M}, R_{im} > 0$

$\quad Y_m^2 \leftarrow \min_i \left\lceil \frac{X_i^2}{R_{im}} \right\rceil, \forall m \in \mathcal{M}, R_{im} > 0$

$\quad Z_m \leftarrow \sum_i \min(X_i^2, R_{im}), \forall m \in \mathcal{M}$

$\quad \chi_m \leftarrow Q_m - L_m, \forall m \in \mathcal{M}$

until $\max_m Y_m^1 \chi_m Z_m = 0$;

8 $\mathbf{Q}^* \leftarrow \mathbf{Q} - \mathbf{L}$

The formal representation of GSH is set out in Algorithm 7. It operates as follows.

4.0. The GSH takes the incumbent solution, prepack configuration, forecasts and service levels as inputs.

4.1. The initialization phase sets the value of the initial feasible solution $\mathbf{L}$ (a vector of size M whose elements are L_m) of optimization problem M-2 (see Section 7.7.2) to zero, and determines the prepack volume U_m for each prepack.

4.2. Calculate the maximum slack quantities χ_m. For any incumbent solution, it is equal to the respective order quantities as the initiation quantity.

4.3. Determine two resource levels $\mathbf{X}^1$ and $\mathbf{X}^2$. $\mathbf{X}^1$ is defined as the excess quantity ordered for an item above its service level threshold. And, $\mathbf{X}^2$ is defined as the excess quantity ordered for an item above its forecast. This is due to the nature of our inventory calculation. As our expected sale quantity is the forecast $\bar{F}_i$ for each item i, any order quantity below it and above the service level threshold $SL_i\bar{F}_i$ will not amount to any increase in excess inventory. A service level below 100% is rational for items which are not "important" (see Section 7.5.1).

4.4. Determine the respective prepack reduction levels for the excess quantities $\mathbf{X}^1$ and $\mathbf{X}^2$. Those levels are $\mathbf{Y}^1$ and $\mathbf{Y}^2$. The notation $\lceil x \rceil$ denotes the smallest integer exceeding x, and the notation $\lfloor x \rfloor$ denotes the largest integer not exceeding x. Hence, $\mathbf{Y}^1$ can be defined as the maximum quantity of order that can be subtracted from the incumbent solution without violating the service level constraints. Similarly, $\mathbf{Y}^2$ can be defined as the minimum quantity of order that can be subtracted from the incumbent solution to reach exactly or just below the forecast quantity.

4.5. Determine the unit inventory reduction potential $\mathbf{Z}$. $\mathbf{Z}$ can be defined as the quantity of inventory reduction per unit reduction of any prepack. For example, reduction in order quantity of any prepack m by one unit will reduce the inventory by R_{im} if the item i is ordered in excess of R_{im}. Otherwise, the inventory reduction quantity will be only the excess, i.e. X_i^2. For any order quantity below forecast, any reduction in prepack order quantity will not reduce inventory. Therefore, the net reduction potential for any prepack m is the sum of minimum of R_{im} and X_i^2.

4.6. Select the prepack with maximum inventory reduction potential and positive slack quantity χ_m. Determine the maximum possible reduction quantity and update the value of the reduction quantity for the prepack.

4.7. Update the slack quantity, resource levels, reduction levels and inventory reduction potential as in 4.3 to 4.5. Then repeat 4.6. until no more reductions can be made or all of the resources are consumed. This gives the final value of $\mathbf{L}$.

4.8. Subtract $\mathbf{L}$ from the current solution $\mathbf{Q}$ to generate the output solution $\mathbf{Q}^*$.

7.8.4. Perturbation and Iterated Local Search

It is established in the complexity analysis in Section 7.7.2 that our approach requires each order quantity Q_m in the current feasible solution $\mathbf{Q}$ to be greater than or equal to the optimal order quantity Q_m^*. The proposed greedy heuristic GH2 used to construct an initial solutions does not necessarily satisfy this condition. The GSH produces better solutions by reducing the order quantities of the initial solutions. Therefore, they both cannot ensure that their outputs satisfy the previous condition. Therefore, we propose to apply random perturbations that increase the order quantities from the local optima to generate new feasible solutions. The perturbation methodology is formally presented in Algorithm 8.

Algorithm 8: Random Perturbation (RP)

Input : $\mathbf{Q}$, θ, γ,
Output: $\mathbf{Q}^*$

```
1 initialization Q* ← Q, i ← 1
2 repeat
      select m ∈ M pertubation position randomly
      Q_m ← Q_m + θ
      i ← i + 1
  until i = γ;
```

Random Perturbation (RP) uses two parameters, θ indicating the perturbation size and γ indicating number of perturbation. For any solution $\mathbf{Q}$, γ random positions are selected and increased by θ. The perturbations are comprised of only addition operations on a current feasible solution. Therefore, the new solutions obtained after the perturbations are also feasible. We expect the higher values of θ and γ to help us explore farther in the solution space. We propose to use RP and GSH alternatively to improve the solution. This procedure constitutes an iterated local search approach. We set time as stopping criterion. The iterated guided search heuristic (IGSH) is presented in Algorithm 9.

Algorithm 9: Iterated Guided Search Heuristic (IGSH)

Input : $\mathbf{Q}, \mathbf{R}, \mathbf{F}, \mathbf{SL}, \theta, \gamma$
Output: $\mathbf{Q}^*$

```
1 initialization Q* ← Q
2 repeat
      Q′ ← RP(Q, θ, γ)
      Q″ ← GSH(Q′, R, F, SL)
3     if w(Q″) ≤ w(Q*) then
          Q* ← Q″
      end
4     Q ← Q″
  until stopping criterion met;
```

As seen from the heuristics explained above, at no point during the execution of the heuristics we operate on infeasible solution space. This is a particularly appealing feature that enables its practical implementation. The stopping criterion can be set at any suitable time and the solution can be used.

Table 7.7: Instance sizes and parameters for numerical analysis. N is the number of items, M is the number of prepacks, $\tilde{R}$ is the number of item-prepack combinations, and **SL** is the vector of service levels. The forecast values **F** and pack contents **R**, and **SL** are chosen randomly from the given uniform distributions.

No.	N	M	$\tilde{R}$	**F**	**R**	SL
A	10	10	50	$\mathcal{U}(100, 200)$	$\mathcal{U}(1, 20)$	1
B				$\mathcal{U}(100, 200)$	$\mathcal{U}(1, 20)$	$\mathcal{U}(0.8, 1)$
C				$\mathcal{U}(100, 500)$	$\mathcal{U}(1, 20)$	1
D				$\mathcal{U}(100, 500)$	$\mathcal{U}(1, 20)$	$\mathcal{U}(0.8, 1)$
E				$\mathcal{U}(100, 1000)$	$\mathcal{U}(1, 20)$	1
F	20	20	50	$\mathcal{U}(100, 200)$	$\mathcal{U}(1, 20)$	1
G				$\mathcal{U}(100, 200)$	$\mathcal{U}(1, 20)$	$\mathcal{U}(0.8, 1)$
H				$\mathcal{U}(100, 500)$	$\mathcal{U}(1, 20)$	1
I				$\mathcal{U}(100, 500)$	$\mathcal{U}(1, 20)$	$\mathcal{U}(0.8, 1)$
J				$\mathcal{U}(100, 1000)$	$\mathcal{U}(1, 20)$	1
K	30	30	100	$\mathcal{U}(100, 200)$	$\mathcal{U}(1, 20)$	1
L				$\mathcal{U}(100, 200)$	$\mathcal{U}(1, 20)$	$\mathcal{U}(0.8, 1)$
M				$\mathcal{U}(100, 500)$	$\mathcal{U}(1, 20)$	1
N				$\mathcal{U}(100, 500)$	$\mathcal{U}(1, 20)$	$\mathcal{U}(0.8, 1)$
O				$\mathcal{U}(100, 1000)$	$\mathcal{U}(1, 20)$	1
P	50	50	450	$\mathcal{U}(100, 200)$	$\mathcal{U}(1, 20)$	1
Q				$\mathcal{U}(100, 200)$	$\mathcal{U}(1, 20)$	$\mathcal{U}(0.8, 1)$
R				$\mathcal{U}(100, 500)$	$\mathcal{U}(1, 20)$	1
S				$\mathcal{U}(100, 500)$	$\mathcal{U}(1, 20)$	$\mathcal{U}(0.8, 1)$
T				$\mathcal{U}(100, 1000)$	$\mathcal{U}(1, 20)$	1
U	50	50	451	$\mathcal{U}(200, 500)$	$\mathcal{U}(1, 10)$	1
V				$\mathcal{U}(200, 500)$	$\mathcal{U}(1, 10)$	$\mathcal{U}(0.8, 1)$
W				$\mathcal{U}(100, 1000)$	$\mathcal{U}(1, 10)$	1
X				$\mathcal{U}(100, 1000)$	$\mathcal{U}(1, 10)$	$\mathcal{U}(0.8, 1)$

7.9. Numerical Study for the Metaheuristic

In order to validate the methodology, an extensive numerical study has been conducted with synthetic as well as real datasets. In case of synthetic datasets, different heuristics and parameters were used to assess their performance. Then, those results were used to select the heuristics and parameters for the experiments with real data sets. The experiments conducted with synthetic data sets are presented in the upcoming subsections. There are two objectives for the experiments with synthetic data.

1. The first is to assess the performance of heuristics used to generate initial solutions, GH1 (see Appendix D) and GH2.
2. The second is to assess the performance of IGSH and the effects of the perturbation parameters.

The synthetic problem instances are presented in Table 7.7. The values of $\mathbf{F}, \mathbf{R}$ and $\mathbf{SL}$ are chosen randomly from uniform distributions with the specified ranges. Each problem instance is defined by the number of items N, number of prepacks M, composition of a per-pack $\mathbf{R}$, forecast $\mathbf{F}$ and service levels $\mathbf{SL}$. We introduce a binary parameter r_{im}, that is equal to 1 if R_{im} is greater than zero, and equal to 0 otherwise. We define instance size as $(N, M, \tilde{R})$, where $\tilde{R} = \sum_i \sum_m r_{im}$. $\tilde{R} \geq M$ since each prepack contains at least one item. $\tilde{R}$ is important in defining the difficulty in solving the problem. For example, if the problem contains prepacks of type $T1$ and $T2$ (see Section 7.2) only, then it can be solved for individual items with less computational effort. In such cases, $\tilde{R} = M$. However, in cases where $\tilde{R} > M$, the problem is more complex. Instances A-O are classified as small instances, and P-X are medium to large instances.

The performance is assessed by comparing the results with the optimal solution obtained using CPLEX 12.7.1. All of the runs were conducted using a Dell Latitude E5470 computer with 8GB memory. The performance criteria measured is defined as the percentage gap between the solution obtained and the optimal solution. If w^* and w^h are fitnesses (see Equation 7.28) of the optimal solution and the solution obtained using heuristic respectively then, the performance criterion Δ can be expressed as below. In case of stochastic output $\overline{\Delta}$ represents the mean value of Δ across multiple runs.

$$\Delta = \frac{w^h - w^*}{w^*} \times 100 \tag{7.29}$$

$$\overline{\Delta} = \frac{1}{totalruns} \sum_{b=1}^{totalruns} \Delta_b \tag{7.30}$$

7.9.1. Comparative Performances on GH1 and GH2

GH1 and GH2 are used for generating feasible initial solutions. Experiments were conducted to assess the quality of initial solutions generated by each of the above heuristics, and also to determine whether the quality of initial solution affects the quality of final solution.

Experimental Protocol

We first analyzed the performance of GH1 and GH2 on randomly generated instances presented in Table 7.7. The output of GH1 can vary depending on the generation of random numbers. Therefore, twenty independent runs were conducted for each instance and the optimality gap for each run was computed. Then the average performance was assessed. But, the output of GH2 is not stochastic and hence, multiple runs were not required. We also recorded the average runtime for each of the heuristics. Secondly, we generated the final solutions using the IGSH after generating the initial solutions using GH1 and GH2. This was done to test whether the initial solution quality affect the final solution. Similar to the earlier case, we took the average of twenty different runs.

Table 7.8: Performances of heuristics used for generating the initial solutions (GH1 and GH2). $\overline{\Delta}$ is the mean percentage gap between the initial solution and the optimal solution. t represents the runtime in seconds.

S.No.	N	M	$\tilde{R}$	t(sec.) GH1	$\overline{\Delta}$ GH1	t(sec.) GH2	$\overline{\Delta}$ GH2
A	10	10	50	3.67	426.00	0.80	152.79
B				3.33	368.63	0.87	39.66
C				4.49	113.86	0.85	39.00
D				4.16	105.90	0.78	35.48
E				9.31	119.39	0.88	79.27
F	20	20	50	6.36	105.98	1.26	48.53
G				5.36	65.55	1.23	108.42
H				12.93	54.41	1.24	53.16
I				11.45	69.54	1.36	52.34
J				19.44	170.16	1.30	72.44
K	30	30	100	11.22	27.72	1.72	22.75
L				9.79	19.18	1.73	31.68
M				21.49	18.63	1.67	6.93
N				18.78	17.65	1.76	13.22
O				45.18	12.41	1.74	9.94
P	50	50	450	8.21	175.52	2.39	155.92
Q				7.67	192.03	2.42	168.28
R				15.07	173.94	2.44	79.94
S				11.74	171.76	2.45	89.77
T				21.12	271.10	2.61	121.96
U	50	50	451	26.37	330.11	2.80	180.91
V				25.45	392.97	2.70	198.47
W				42.48	136.96	2.68	98.48
X				34.09	153.18	2.82	84.41
Average				15.79	153.86	1.77	80.99
Gain over GH1				-	-	89%	47%

Results and Discussion

In Table 7.8 we present the performances of heuristics GH1 and GH2 for initial solutions. In most of the instances the solution generated by GH1 is of inferior quality than that generated by GH2. Also, GH1 is significantly slower than GH2. The average time taken by GH2 for generating an initial solution is 89% lower than that taken by GH1 on average. It can also be observed that GH1 takes more and more time for the same instance size with increase in the forecast quantity. On the other hand, the runtime of GH2 is not affected by increase in the forecast quantity. The fitness of the solutions generated using GH2 is 47% better on average. The interest here is to examine the impact of selection GH1 or GH2 on the final runtime and final quality of the solution. Hence, only GH2 was chosen for the forthcoming experiments.

Table 7.9 represents the optimality gaps of the final solution obtained using the IGSH after GH1 and GH2. We fixed the perturbation parameters for the IGSH as $\theta = 5$, $\gamma = 10$. As the IGSH is stochastic, the average performances across twenty different iterations are compared.

Figure **??** presents the spreads of the fitnesses of the solutions in form of box-plots. One way ANOVA tests are conducted for each of the instances as well as on the average performance of all instances with a 95% significance level. The null hypothesis is defined as "the mean performances of both approaches are equal", and is rejected when *p*-value was less than or equal to 0.05. Although, the quality of final solutions obtained from both approaches are not very different, the difference is significant and we cannot conclude that both approaches perform equally. For individual instances a choice between GH1 and GH2 can be made based on solution quality and runtime. We expect the runtime to be a significant factor if the total runtime is very short. Individually in 21 out of 24 instances IGSH with GH2 gave better quality solutions. In 12 of the instances we reject the null hypothesis. However, GH2 is always faster than GH1 (see Table 7.8). Hence we selected GH2 for further used.

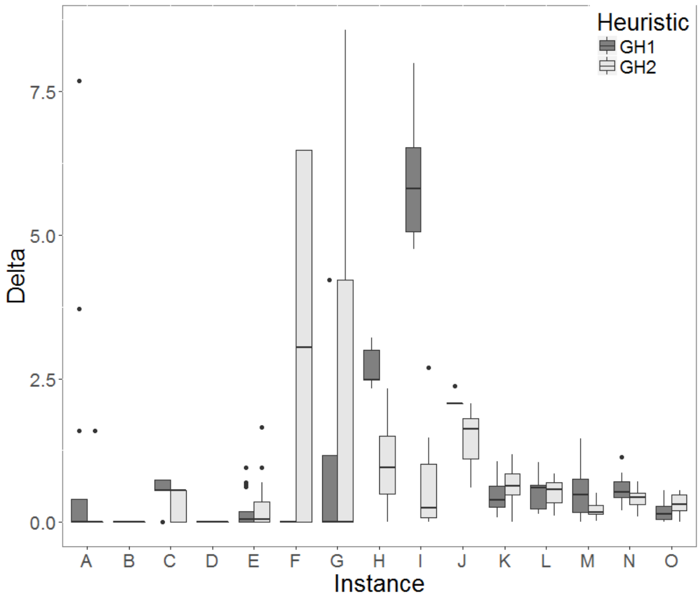


Figure 7.10: Effect of initial solution on final quality on instances A-O.

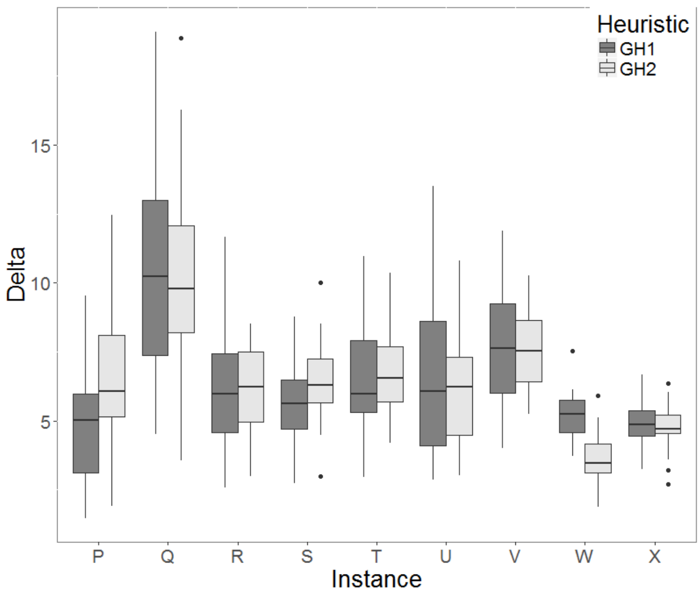


Figure 7.11: Effect of initial solution on final quality on instances P-X.

Table 7.9: Effect of initial solution on the quality of final solution. $\overline{\Delta}$ is the mean percentage gap between the initial solution and the optimal solution. t represents the runtime in seconds. σ is the standard deviation of Δ for 20 iterations. The $p-$value is calculated using one way ANOVA between the observations obtained using IGSH-GH1 and IGSH-GH2.

S.No.	t(sec.)	$\overline{\Delta}(\sigma)$ IGSH-GH1	$\overline{\Delta}(\sigma)$ IGSH-GH2	$p-$Value	Choice
A	60	0.81(0.02)	0.32(0.01)	0.28	GH2
B	60	0.00(0.00)	0.00(0.00)		GH2
C	60	0.55(0.00)	0.39(0.00)	0.03	GH2
D	60	0.00(0.00)	0.00(0.00)		GH2
E	60	0.20(0.00)	0.28(0.00)	0.49	GH1
F	60	0.00(0.00)	3.35(0.03)	<0.01	GH1
G	60	0.99(0.02)	1.59(0.03)	0.19	GH2
H	60	2.69(0.00)	0.95(0.01)	<0.01	GH2
I	60	5.82(0.01)	0.62(0.01)	<0.01	GH2
J	60	2.09(0.00)	1.34(0.00)	<0.01	GH2
K	60	0.45(0.00)	0.44(0.00)	0.04	GH2
L	60	0.51(0.00)	0.48(0.00)	0.87	GH2
M	60	0.50(0.00)	0.23(0.00)	<0.01	GH2
N	60	0.57(0.00)	0.46(0.00)	0.02	GH2
O	60	0.18(0.00)	0.27(0.00)	0.03	GH1
P	300	6.65(0.03)	7.26(0.03)	0.96	GH2
Q	300	12.03(0.04)	10.33(0.04)	0.13	GH2
R	300	6.73(0.01)	6.52(0.02)	0.19	GH2
S	300	5.82(0.02)	5.79(0.02)	0.23	GH2
T	300	5.71(0.01)	5.73(0.02)	0.03	GH2
U	300	5.72(0.02)	6.79(0.02)	0.40	GH2
V	300	6.77(0.01)	6.64(0.01)	0.09	GH2
W	300	5.03(0.01)	3.46(0.01)	<0.01	GH2
X	300	4.92(0.02)	4.08(0.01)	0.68	GH2
Avg.		3.11	2.80	0.02	GH2

7.9.2. Performance of IGSH

Experiments were conducted to assess the performance of IGSH. GH2 was selected for generating the initial solutions and IGSH was run with different perturbation parameters to assess the quality of final solutions.

Experimental Protocol

In previous experiments fixed values for the perturbation size θ and the number of perturbations γ were used. In order to assess their effects and also the robustness of the meta-heuristic, experiments were conducted with different values of perturbation parameters. The values of γ were set at 3, 5, 10, 15, and 20. For each value of γ, the values of θ were set at 2, 3, and 5. Therefore, 15 different configurations of IGSH for each problem instance were tested. Similar to the earlier case, 20 iterations were run for each configuration and the value of $\overline{\Delta}$ (see Equation 7.29, 7.30) was calculated.

Results and Discussion

Experimental results for smaller instances (A-O) and medium to large instances (P-X) are presented in Table 7.10 and 7.11 respectively. For smaller instances, the optimal solution can be obtained using using CPLEX in reasonable time. In such instances IGSH gives close to optimal solutions with the largest gap being 2% without any significant difference in runtime. However, in case of medium to larger instances, the runtime required to obtain the optimal solution using CPLEX extends up to 8 hours. In such cases, IGSH provides solutions which are within 8% of the optimal solution within 300 seconds.

Table 7.10: Effect of perturbation size θ and number of perturbations γ. Problem instances A-O. The statistically best results are indicated in bold font.

S.No.	t(sec.)	θ	$\overline{\Delta}(\%)$				
			$\gamma = 3$	$\gamma = 5$	$\gamma = 10$	$\gamma = 15$	$\gamma = 20$
A	60	2	2.41	0.08	0.08	**0**	**0**
	60	3	0.48	0.40	0.08	0.27	0.36
	60	5	0.89	0.58	0.32	0.24	0.73
B	60	2	0	0	0	0	0
	60	3	0	0	0	0	0
	60	5	0	0	0	0	0
C	60	2	0.66	**0.29**	0.44	0.40	0.52
	60	3	0.62	0.51	0.60	0.62	0.63
	60	5	0.54	0.30	0.39	0.50	0.51
D	60	2	**0**	**0**	**0**	0.47	0.19
	60	3	0.20	**0**	**0**	**0**	0.55
	60	5	0.16	**0**	**0**	**0**	**0**
E	60	2	7.74	1.64	0.58	0.34	0.47
	60	3	2.22	0.48	**0.15**	0.23	0.46
	60	5	1.55	0.23	0.28	0.27	0.85
F	60	2	5.76	6.06	6.08	5.77	6.32
	60	3	4.50	6.31	4.78	4.65	3.16
	60	5	5.53	4.47	3.35	3.01	**1.56**
G	60	2	8.15	7.73	7.99	8.15	8.15
	60	3	7.46	7.75	7.75	6.55	5.75
	60	5	5.95	5.22	1.59	1.48	**0.84**
H	60	2	0.51	0.54	0.56	0.64	0.82
	60	3	**0.46**	0.50	0.69	1.26	1.28
	60	5	0.61	0.58	0.95	1.02	1.33
I	60	2	0.28	0.49	0.42	0.53	0.76
	60	3	**0.15**	0.39	0.39	0.62	1.05
	60	5	0.57	0.26	0.62	0.51	1.13
J	60	2	2.02	1.55	2.20	2.03	2.00
	60	3	0.42	0.93	1.71	1.85	2.01
	60	5	**0.37**	1.10	1.34	1.93	2.01
K	60	2	1.57	1.35	0.80	0.79	0.72
	60	3	1.40	0.93	0.77	0.61	0.67
	60	5	1.30	0.96	**0.44**	0.47	0.56
L	60	2	1.78	1.29	0.67	0.70	0.75
	60	3	1.55	0.93	0.73	0.64	0.71
	60	5	1.32	0.69	**0.48**	0.56	0.57
M	60	2	0.68	0.47	0.30	0.38	0.59
	60	3	0.51	0.39	0.27	0.35	0.34
	60	5	0.39	0.33	**0.23**	0.24	0.33
N	60	2	0.99	0.75	0.70	0.57	0.64
	60	3	1.00	0.71	0.49	0.46	0.50
	60	5	0.88	0.58	0.46	0.42	**0.41**
O	60	2	1.67	0.94	0.68	0.63	0.56
	60	3	2.22	0.72	0.49	0.40	0.37
	60	5	0.95	0.62	0.27	0.25	**0.21**

Table 7.11: Effect of perturbation size θ and number of perturbations γ. Problem instance P-X. The statistically best results are indicated in bold font.

S.No.	t(sec.)	θ	$\overline{\Delta}(\%)$				
			$\gamma = 3$	$\gamma = 5$	$\gamma = 10$	$\gamma = 15$	$\gamma = 20$
P	300	2	14.69	5.55	4.37	5.37	5.29
	300	3	13.48	6.21	**3.76**	5.08	6.88
	300	5	14.15	7.00	7.26	8.99	10.44
Q	300	2	21.24	12.05	7.48	8.35	9.12
	300	3	23.68	11.46	**7.36**	9.14	11.68
	300	5	25.84	13.88	10.33	13.05	14.52
R	300	2	12.45	7.95	6.48	6.05	7.16
	300	3	12.35	7.21	**5.21**	5.73	6.37
	300	5	11.80	8.13	6.52	7.27	9.14
S	300	2	10.91	6.78	**5.45**	6.22	7.31
	300	3	10.07	7.03	5.53	5.59	6.02
	300	5	11.48	8.05	5.79	6.45	9.43
T	300	2	31.61	15.95	**4.65**	5.46	6.82
	300	3	25.00	9.19	5.63	5.86	7.24
	300	5	23.36	9.43	5.73	6.87	8.22
U	300	2	28.90	14.11	**5.12**	5.95	7.29
	300	3	23.21	11.49	5.47	5.39	6.74
	300	5	23.76	12.86	6.79	6.54	7.63
V	300	2	18.36	10.62	**5.89**	7.61	7.79
	300	3	17.86	9.21	6.19	7.05	7.65
	300	5	18.41	10.34	6.64	8.02	10.44
W	300	2	16.06	6.45	4.16	4.15	5.15
	300	3	13.14	5.75	4.05	4.38	4.88
	300	5	12.13	4.90	**3.46**	4.23	5.36
X	300	2	12.66	6.73	**4.06**	4.25	5.07
	300	3	9.81	5.46	4.09	4.19	5.99
	300	5	9.17	5.57	4.08	4.87	5.70

Table 7.12: Overall performance of different perturbation parameters across problem instances A-P. The best results are indicated in bold font.

θ	Average Rank, Instances A-O				
	$\gamma = 3$	$\gamma = 5$	$\gamma = 10$	$\gamma = 15$	$\gamma = 20$
2	10.80	7.86	7.26	8.06	9.20
3	9.33	7.60	5.66	6.20	8.20
5	9.46	5.13	**4.06**	**4.06**	5.93

As seen from Table 7.10 and Table 7.11, different perturbation parameters yield final solutions which are of varying fitness. This is because of the effect of those parameters on the capability to explore new search space. In our case, although we could not quantify the relationship between θ or γ and $\overline{\Delta}$, some clear observations can be made from the results presented in Table 7.10 and Table 7.11. For a fixed θ, an increase in γ improves the solution quality up to some extent and then deteriorates. Similarly, for a fixed γ, an increase in θ first improves and then deteriorates the solution quality. We plan to explore automatic or adaptive parameter selection in our future research.

To determine the relative effectiveness of different values of perturbation parameters, the rank of each combination based on its $\overline{\Delta}$ has been determined. The details are presented in Table 7.12 for smaller instances and in Table 7.13 for medium to larger instances. A lower rank means it has better $\overline{\Delta}$ (i.e. lower excess inventory). We observe that, in each of the problem instance one of the combinations of the perturbation parameters gives the best fitness. $\gamma = 10, 15, \theta = 5$ perform best for smaller instances, and $\gamma = 10, \theta = 2, 3$ perform best for medium to larger instances.

Table 7.13: Overall performance of different perturbation parameters across problem instances P-X. The best results are indicated in bold font.

θ	Average Rank, Instances P-X				
	$\gamma = 3$	$\gamma = 5$	$\gamma = 10$	$\gamma = 15$	$\gamma = 20$
2	14.55	10.33	**2.00**	4.00	6.89
3	13.56	8.77	**2.00**	3.77	7.00
5	13.89	10.11	4.78	7.56	10.78

Below in Figure 7.12, convergence curves for problem instances A and X are plotted as examples of small and large instances. The plots for remaining instances are provided in Appendix E. It can be seen that the initial rate of convergence is fast and gradually slows down.

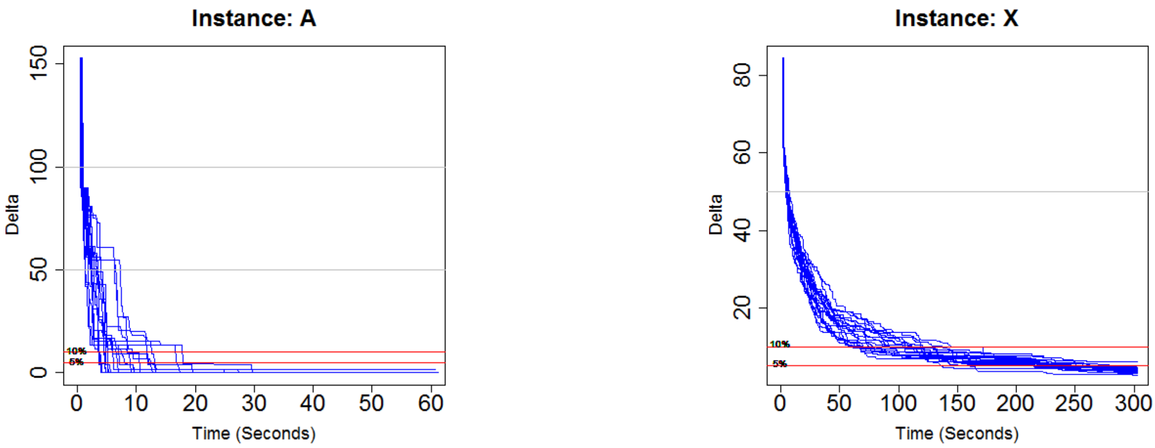


Figure 7.12: Convergence curves for problem instances A and X.

7.10. Case Study 2

In this section, we present a case study about a multinational retailer firm, that encounters the problem discussed above during promotions. The firm conducts multiple promotional activities throughout the year. In each of those promotions, the firm sells more than 10,000 items in about 500 retailers. For validation purposes the retailer provided us with a subset of those items. The suppliers of the concerned firm provide multiple packaging options for the items and determine the configuration of the prepacks. The suppliers also require about 12 weeks of lead time to fulfill the orders. Hence, the firm effectively places only one order for every promotion. It tries to minimize its leftover inventory after each promotion without loosing sales.

7.10.1. Data and Experimental Protocol

The dataset is detailed in Table 7.14. There are 408 retailers and each of them sells 843 items during a promotion. For the 843 items, there are 1238 prepacks and 1538 item-prepack combinations. Each retailer has the same packaging choices.

The parameter values for IGSH were set at $\gamma = 10$ and $\theta = 5$. Similar to earlier analysis, 20 different runs were used to assess performance. As the problems were independent for each retailer, each of them was solved separately using IGSH and CPLEX, and the overall performance was compared. Δ and total runtime reduction were calculated.

Table 7.14: Parameters for the case study.

Network structure	
Number of retailers	408
Promotional components	
Number of items	843
Number of prepacks	1238
Number of item-prepack combinations	1538
Prepack size	[1, 504]
Total forecast (Units)	2,770,518
Metaheuristic parameters	
γ	10
θ	5

7.10.2. Results and Discussion

The results are analyzed in two parts. First, the results are analyzed for each retailer. The retailer specific results are presented in Table 7.15. The following values are calculated. The mean percentage gap over different runs is calculated for each retailer and denoted by $\overline{\Delta}_k$ (refer to Equation 7.30. Here, it is used with the subscript for retailer). The average mean gap $\overline{\overline{\Delta}}$ is defined as the average value of $\overline{\Delta}_k$. The average minimum $\overline{\Delta}_{min}$ is defined as the average value minimum Δ for each retailer across 20 runs. At last, the average maximum $\overline{\Delta}_{max}$ is defined as the average values of maximum Δ for each retailer across 20 runs. The notation b is used for individual runs.

$$\overline{\overline{\Delta}} = \frac{1}{K} \sum_{k \in \mathcal{K}} \overline{\Delta}_k \tag{7.31}$$

$$\overline{\Delta}_{min} = \frac{1}{K} \sum_{k \in \mathcal{K}} \left[\min_{b \in \{1..20\}} \Delta_{kb} \right] \tag{7.32}$$

$$\overline{\Delta}_{max} = \frac{1}{K} \sum_{k \in \mathcal{K}} \left[\max_{b \in \{1..20\}} \Delta_{kb} \right] \tag{7.33}$$

On average IGSH gives solutions within 4.79% of the optimum. The minimum performance gap is 1.79% on average and the maximum performance gap is 7.58% on average. The minimum value of mean percentage gap is 0.30% and the maximum value of mean percentage gap is 21.63%. In 85 out of 408 retailers, the mean gaps are within 2% of the optimal solution and in 287 retailers our mean results are within 5% of the optimal solution. Figure 7.13 depicts the histogram for mean performances for all of the 408 retailers.

We also analyze the overall performance across all retailers and the runtime. The results are summarized in Table 7.16. The overall results are obtained by summing the individual results of each retailer. The optimal solution gives 1,357,025 units of excess inventory, which is 48.98% of the total demand. Meanwhile, the solution obtained using IGSH gives 1,419,012 units of excess inventory, which is 51.21% of the total demand. In absolute terms, the IGSH gives 4.56% higher excess inventory than the optimal solution. With CPLEX the runtime is about 51 hours. Although, this seems to be a reasonable computation time considering less frequent promotions, for the complete dataset it becomes intractable. We also obtain 94.93% reduction in runtime for the 408-retailer use case.

Table 7.15: retailer specific results of the case study.

retailer specific results	
$\overline{\overline{\Delta}}$	4.79%
$\overline{\Delta}_{min}$	1.79%
$\overline{\Delta}_{max}$	7.58%
Minimum $\overline{\Delta}_k$	0.30%
Maximum $\overline{\Delta}_k$	21.63%
Number of retailers with $\overline{\Delta}_k \leq 2\%$	85 (20.83%)
Number of retailers with $\overline{\Delta}_k \leq 5\%$	278 (68.13%)
Number of retailers with $\overline{\Delta}_k \leq 10\%$	360 (88.23%)

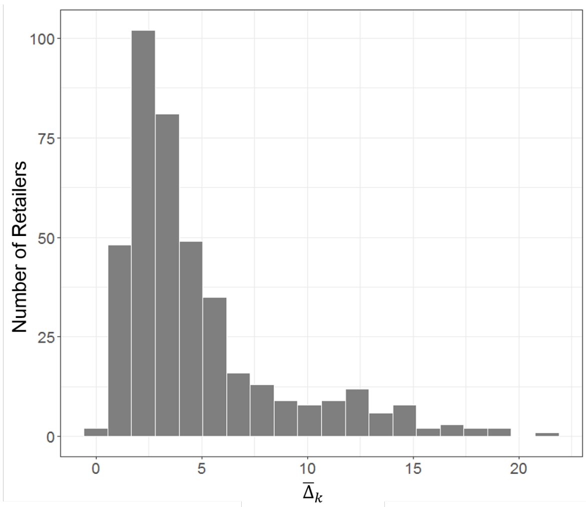


Figure 7.13: Histogram of $\overline{\Delta}_k$ for different retailers of the case study

7.11. Conclusions

In this chapter, we addressed the promotional ordering problem for retail supply chains as an extension of the general newsvendor problem. We considered a real world problem, where during a promotional event the retailer has the option of ordering the items in different prepacks and there is lack of financial information. We considered the demand to be stochastic and introduce the use of a discrete probabilistic forecast similar to scenarios in the single period newsvendor problem to address uncertainty. We also described some business rules for real world applications. These business rules guarantee that the proposed solution will be manageable by retailers. They address basic distribution priority for larger pack sizes, consider criticality of items while ordering and limit the type of prepacks per item.

Table 7.16: Overall results of the case study.

Overall results	
w(CPLEX)	1,357,025
	(48.98%)
w(IGSH)	1,419,012
	(51.21%)
Overall Δ(%)	4.56
Reduction in runtime (%)	94.93

First, we formulated the problem as a three-objective MILP model and incorporated the business rules as objectives and constraints. We then proposed a 2-stage methodology to solve the optimization problem. In the first stage, we used an ABC-VED inventory classification method to define minimum service levels thresholds for different classes of items at different retailers. In the second stage, we transformed the tri-objective optimization model to a bi-objective model. Then we solved it using an ϵ-constrained like method.

We applied our methodology to a real world use case. Results show that the probabilistic approach always outperforms the deterministic approach and both approaches outperform the solution currently used by the firm. At lower leftover inventory levels both approaches converge. However, the probabilistic approach performs much better when the limit on leftover inventory level increases.

Secondly, to be more prudent in practical application we reformulated the problem to minimize the leftover inventory at a given service level. Complexity analysis of the problem suggests that, the solution method encounters a multi-dimensional knapsack problem due to the presence of multi-item prepacks. Therefore, the problem is difficult to solve optimally for practical purposes.

We have provided an iterated guided search approach to obtain near optimal solutions. The methodology uses a greedy heuristic to generate initial solutions and a guided search heuristic for obtaining the local optimum. Perturbations are applied to the local optimum to explore the search space. Our methodology ensures the whole search to be conducted in the feasible space only to enable practical usage.

We first conducted extensive numerical analysis on random realistic problem instances. The proposed metaheuristic provides solutions within a 2% optimality gap within 60 seconds for small instances, and within a 8% gap for medium to large instance. We also tested our methodology on a real test case of 408 retailers. In that case we obtain an overall gap of 4.56% with less than a 10% gap in 88.23% of the retailers.

In this dissertation we have addressed the single-item stochastic replenishment problem, its extension to supplier selection and multi-item promotional ordering. The promotional ordering is a single-period problem. We plan to conduct additional research in the area of multi-item replenishment with multiple time periods. Its deterministic version is proven to be strongly $\mathcal{NP}$-hard. We have already conducted preliminary studies on our proposed use of a learning based stochastic local search. The prosed approach used the properties of local optima to construct new solutions so to navigate the search space in less time. Our initial results are promising and the approach already outperforms existing best known methods. In the next chapter, we discuss some of the industrial aspects of our research and conclude this dissertation.

Chapter 8

Extensions, Industrialization, Conclusions and Perspectives

Stochastic replenishment planning is a challenge when we consider real-world conditions. The optimal polices are known only for limited cases. Moreover, only 5% of the literature in this area has been useful in practice. Apart from replenishment planning, two other areas related to inventory optimization which are important in practice are: classification of those problems and performance assessment. And, they are not widely investigated as well. Our work in this dissertation addresses the above three research topics.

This chapter is arranged as follows. In Section 8.1, we summarize our contributions and provide the conclusions. As the proposed solution methods are shown to be effective for the "base" problems, in Section 8.2, we various possible extensions to those methods and detail the industrialization status of our proposed works. In the end, we provide the perspectives about future related research areas in Section 8.3.

8.1. Summary of the Main Contributions

In this dissertation, we have addressed three broad research areas related to inventory optimization.

1. Problem classification.
   (a) Proposition of a classification scheme for inventory optimization problems .
2. Performance evaluation.
   (a) Selection of suitable key performance indicators (KPIs) for industrial inventory management systems (IMS) and proposition of a simulation framework for performance assessment.
3. Replenishment planning methods.
   (a) Proposition of sampling-based optimization method for single-item replenishment planning problems.
   (b) Proposition of generalized sampling-based method for single-item replenishment planning with supplier selection.
   (c) Proposition of of scenario based optimization method for promotional joint replenishment problems with prepacks and development of metaheuristics.

Indeed in Chapter 2, we proposed a classification scheme for inventory optimization problems. We took inspirations from Graham's notations (Graham *et al.*, 1979) for scheduling problems. The proposed scheme has six fields and each fields has sub-fields. Depending on the specific type of sub-field the problem are named. Then in Chapter 3, we listed some of the important KPIs for inventory management systems. We categorized those KPIs under three groups: financial, operational and service. Evaluation of those KPIs for a single IMS is simple in practice provided all the required data is recorded. However, comparative evaluation of two or more IMSs is difficult. Comparison with an existing IMS lacks identical problem setting or requires a longer time horizon. In Chapter 3, we also proposed a simulation method (Δ-method) (Sahu *et al.*, 2020a) to have an effective comparative assessment of multiple IMSs. In Chapter 4, we defined the replenishment planning problems in detail. Various types of problems were identified and organized to form a modular IMS framework. Form Chapter 5 onwards, we focus on actual solution methods.

In Chapter 5, we proposed a sampling-based method for the stochastic single-item replenishment planning problem. Our proposed method optimizes the approximate expected cost for the first period in the planning horizon. We call this method as immediate expected cost approach (IECA). We first proposed an enumerative heuristic for the optimal immediate expected cost. Then we proved the convexity of cost function and proposed an improved heuristic (OH-IECA) based on this property. Numerical results suggest that OH-IECA gives 0.463% excess cost over the optimal for the benchmark instances in Veinott Jr and Wagner (1965). The proposed method is also useful in case of non-stationary demand but, gives comparatively higher cost. We also proposed its extensions in the presence of minimum order quantity and higher variance. The proposed method requires little modification to include batch-size. For stationary demand, the optimality gap in the presence of batch size is less than 1%. Moreover, the proposed method requires very less computational effort, about 7 milliseconds as compared to 3 minutes with dynamic programming.

In Chapter 6, we addressed the single-item replenishment planning problem with non-stationary demand and the single-item replenishment planning and supplier selection problem. We extended the works from Chapter 5 to formulate a dynamic program with fewer state space. The proposed method was tested for full backorder scenario with average optimality gap of 3-4%. The proposed method can also be adapted for the extensions of the basic problem to address lost sales and batch sizes. The second problem addressed was also a multi-stage stochastic program. We proposed two approaches for supplier selection: common supplier selection and dynamic supplier selection (Sahu *et al.*, 2020b). We first conducted the financial benefit analysis of dynamic supplier selection versus selecting a common supplier for the planning horizon. Then we proposed an approximation framework for both approaches. A common supplier for the whole planning horizon is a practically more appealing feature. However, the dynamic supplier selection results in higher economic benefits. Both of the aforementioned problems are multi-stage stochastic optimization problems. Numerical analysis suggest that the dynamic supplier supplier selection approach always outperforms the approach with one common supplier, especially when the inventory holding costs and the backorder costs are very different, and when the suppliers impose a minimum order quantity constraints. Finding the optimal solutions of any of the above approaches is time consuming. Hence, we developed an approximation framework based on dynamic programming. The proposed solution framework works in two stages. The first stage gives the optimal order quantity and cost for discrete coverage period. Afterwards, a dynamic programming approach optimizes the total cost for the rolling horizon. Numerical analysis for synthetic instances suggests an average the optimality gap is 3.5%. The average computation time is 6 milliseconds.

In Chapter 7, we addressed the promotional ordering problem for retail supply chains (Sahu *et al.*, 2018) as an extension of the general newsvendor problem. We considered the demand to be stochastic and introduce the use of a discrete probabilistic forecast similar to scenarios in the single-period newsvendor problem to address the uncertainty. We first formulated the problem as a multi-objective MILP model and then proposed a 2-stage methodology to solve the optimization problem. In the first stage, we use ABC-VED inventory classification method to define minimum service levels thresholds. Then in the second stage, we transformed the multi-objective optimization model to a bi-objective model. Then we solved it using an ϵ-constrained like method. Results show that the probabilistic approach always outperforms the deterministic approach. At lower leftover inventory levels both approaches converge. However, the probabilistic approach performs much better when the limit on leftover inventory level increases. Analysis on a case study suggests 43-55% reduction in lost sales for same inventory levels and 36% reduction in leftover inventory for the same lost sales level. On the second part, we then reformulated the problem to minimize the leftover inventory at a given service level. Complexity analysis suggested that, the solution method encounters a multi-dimensional knapsack problem due to the presence of multi-item prepacks. We then developed an iterated guided search approach to obtain near optimal solutions. The methodology uses a greedy heuristic to generate initial solutions and a guided search heuristic for obtaining the local optimum. During numerical analysis on random realistic problem instances, the proposed metaheuristic provided solutions within a 2% optimality gap within 60 seconds for small instances, and within a 8% gap for medium to large instance.

8.2. Extensions and Industrialization

The parameters associated with real-world replenishment problems vary from those considered for the mathematical modeling. This is to simplify the solution procedure. In this section, we present some of the possible extensions to the solution methods presented earlier in this dissertation. They are summarized in Table 8.1. The status of their industrialization at Vekia is also indicated.

Most extensions are proposed for the single-item replenishment planning (SRP) and the single-item replenishment planning with supplier selection (SRPSS) problems. The proposed base models for the SRP problem address zero lead time, stationary and non-stationary demand, backorder, unit batch size and equal uncertainty. Extensions can address deterministic and stochastic lead times, lost sale, a mix of lost sales and backorder, service level constraints, increased uncertainty, minimum and maximum stock constraints and minimum order quantity (MOQ) constraint. Similarly for the SRPSS, major extensions are to include stochastic lead times and lost sales.

8.3. Perspectives

Based on numerical analysis in each chapter and how we imagine future supply chain optimization processes at Vekia, we identify the opportunities for future research as the following.

The proposed sampling-based method for the SRP problem is highly flexible, however, like in the literature, it is build on a major assumption. The method assumes that the demand forecast is accurate, i.e., we draw random samples from the forecast. It can be adapted only if the actual uncertainty is higher. Unfortunately, almost all forecasts come up with some errors. These errors have a major impact on the replenishment quality. Our current research is focused on including the effect of forecast error in replenishment planning or to tune the replenishment planning process based on forecast error. This would be a particularly interesting topic as this would lead to a segregation of the effect of forecast accuracy and efficient optimization on the final cost.

Table 8.1: Possible extensions and industrialization status. The three problems: single-item replenishment planning (SRP), single-item replenishment planning with supplier selection (SRPSS) and the promotional joint replenishment problem (PJRP) are under industrialization. The "Base model" refers to the problem defined in this dissertation without any extension.

Concerned Problem	Problem Parameter	Option	Base Model/ Extension	Industrialization Status
SRP	Lead time	Zero	Base model	Yes
		Deterministic	Extension	Yes
		Stochstic	Extension	Ongoing
	Demand	Stationary	Base model	Yes
		Non-Stationary	Base model	Yes
	Shortage	Backorder	Base model	Yes
		Zero	Extension	Ongoing
		Lost sale	Extension	Yes
		Mixed	Extension	Ongoing
	Service level	Alpha	Extension	Ongoing
		Beta	Extension	Yes
	Batch size	Unit	Base model	Yes
		Single	Extension	Yes
		Multiple	Extension	Ongoing
	Uncertainty	Equal	Base model	Yes
		Increased	Extension	Ongoing
	Min. stock	No	Base model	Yes
		Yes	Extension	Yes
	Max. stock	No	Base model	Yes
		Yes	Extension	Yes
	MOQ	No	Base model	Yes
		Yes	Extension	Yes
SRPSS	Lead time	Zero	Base model	Yes
		Deterministic	Extension	Yes
		Stochstic	Extension	Ongoing
	Demand	Stationary	Base model	Yes
		Non-Stationary	Base model	Yes
	Shortage	Backorder	Base model	Yes
		Lost sale	Extension	Yes
	Service level	Alpha	Extension	Ongoing
		Beta	Extension	Yes
	Batch size	Unit	Base model	Yes
		Single	Extension	Yes
	Min. stock	No	Base model	Yes
		Yes	Extension	Yes
	Max. stock	No	Base model	Yes
		Yes	Extension	Yes
	MOQ	No	Base model	Yes
		Yes	Extension	Yes
PJRP	-	-	Base model	Ongoing

In most real-world problems, we encounter non-stationary demand. With our proposed one-step method IECA we obtain a relatively higher optimality gap. Even though using the RDP in Chapter 6 we can have a lower gap, further research is needed to have a one-step method.

Supplier selection is another interesting topic of research. We have showed that a dynamic supplier selection approach is cost effective. However, practical ease must also be taken care of. We also have not addressed the effect of discount on supplier selection. That might hold important outcomes since it is widely used by the suppliers. Next, for the promotional ordering problem, we have addressed a novel problem even if the use of prepacks is very much evident nowadays. Future research can investigate the effect of presence of prepacks on multi-period replenishment planning.

Apart from the PJRP prblem, the JRP problems are very much evident in multi-period planning scenario. Examples of some use case are: presence of joint fixed cost, multi-item quantity discount, multi-item discount based on purchase amount, multi-item truckload optimization, etc. We have conducted some preliminary research on the multi-period JRP problem. Initial results with a synchronized neighborhood operator (SNO) (Sahu and Veerapen, 2019) has shown some promising results by outperforming the existing best know methods. This requires further research.

Finally, very few of the supply chain networks are single-echelon. Our next focus is to implement the proposed work in a multi-echelon setting. This has two reasons. First one is that, capacitated multi-echelon inventory optimization is notoriously hard to solve. With real-world constraints, it is even more so. The second reason is user preference. Although, a multi-echelon approach can be cost effective, it requires fully centralized planning. However, recent trend suggests, organizations still want to retain some decentralized control. Therefore, the optimal degree of centralization is also another interesting area of research. Different modes of fulfillment are emerging nowadays. One such mode is transshipment. It has tremendous potential to improve service level. Proactive transshipment may incur higher transportation cost but, organization can explore this option to reduce their overall cost. Transshipment can be seen as an alternative to ordering from a supplier with higher lead time. This will not only improve service quality (with lower lead time) but also, reduce the inventory held in the supply chain. Future research could explore the possibility of incorporating the same into regular replenishment planning.

With the current crisis due to COVID-19, organizations are seeing unprecedented disruptions in their supply chains. One way to mitigate the risk of such disruptions is to carefully integrate them into their supply chain planning. However, it remains extremely difficult to predict such events. Another way organizations can mitigate the effect is to increase visibility and gain the ability to simulate the effects of such events. Supply chain digital twin is one approach that seems extremely appealing. Not only can it simulate the effect of various actions, it can also help analyze the current processes. This, in turn can also improve visibility.

Lastly, coming to the use of artificial intelligence (AI) in supply chain planning, we foresee a substantial impact. While AI has proven to be very efficient in providing reliable future information so that the planning can be done in advance, and accurately, the users often require an explanation of the proposed solution. The demand for explainable artificial intelligence (XAI) is growing day-by-day. Particularly in case of inventory management, AI and optimization methods have been proven to be helpful in lowering the inventory cost and increasing the service level. Other areas where these could be helpful include, predictive tracking, prediction and optimization of perishable inventory, etc. Automating the planning process is another benefit which is seeing increasing importance today.

List of Publications

International Journals

IJ1. Rabin Sahu, Clarisse Dhaenens, Nadarajen Veerapen, and Manuel Davy. A Sampling-Based Optimization Approach for Stochastic Replenishment Planning. *In Preparation*

IJ2. Rabin Sahu, Clarisse Dhaenens, Nadarajen Veerapen, and Manuel Davy. Retail Promotional Ordering: Mathematical Formulation and Metaheuristic Approach. *In Preparation*

International Conferences with Proceedings

CP1. (Sahu et al., 2020b) Rabin Sahu, Clarisse Dhaenens, Nadarajen Veerapen, and Manuel Davy. An approximate method for integrated stochastic replenishment planning with supplier selection. In International Conference on Operations Research and Enterprise Systems- ICORES, 2020.

CP2. (Sahu et al., 2018) Rabin Sahu, Alexandre Gerussi, Clarisse Dhaenens, Jianhua Zhang, and Manuel Davy. Multi-objective optimization of retail promotional ordering with probabilistic forecast. In ILS2018-International Conference on Information Systems, Logistics and Supply Chain, 2018.

CP3. Tarakanta Barik, and Rabin Sahu. Hybrid Manufacturing: Production Plan, Cost Optimization and Location Decisions (Part I & II). International Conference on Industrial Engineering (ICIE-2015), SVNIT Surat ISBN: 978-93-84935-56-6 Page No- 778-791.

International Conferences and Workshops

CW1. (Sahu et al., 2020a) Rabin Sahu, Clarisse Dhaenens, and Nadarajen Veerapen. A performance evaluation framework for inventory management systems. In Annual conference of the Belgian Operational Research Society-ORBEL, 2020.

CW2. (Sahu and Veerapen, 2019) Rabin Sahu and Nadarajen Veerapen. Stochastic joint replenishment planning optimization with an iterated local search. In International Workshop on Stochastic Local Search Algorithms-SLS, 2019.

Bibliography

(Abdel-Malek *et al.*, 2008) Layek Abdel-Malek, Roberto Montanari, and Diego Meneghetti. The capacitated newsboy problem with random yield: The gardener problem. *International Journal of Production Economics*, 115(1):113–127, 2008.

(Adivar *et al.*, 2019) Burcu Adivar, Işık Özge Yumurtacı Hüseyinoğlu, and Martin Christopher. A quantitative performance management framework for assessing omnichannel retail supply chains. *Journal of Retailing and Consumer Services*, 48:257–269, 2019.

(Akçay *et al.*, 2007) Yalçın Akçay, Haijun Li, and Susan H Xu. Greedy algorithm for the general multidimensional knapsack problem. *Annals of Operations Research*, 150(1):17, 2007.

(APICS) APICS. Supply chain operations reference (scor) model. http://www.apics.org/apics-for-business/frameworks/scor. Accessed: 2019-01-17.

(Arrow *et al.*, 1951) Kenneth J Arrow, Theodore Harris, and Jacob Marschak. Optimal inventory policy. *Econometrica: Journal of the Econometric Society*, pages 250–272, 1951.

(Askin, 1981) Ronald G Askin. A procedure for production lot sizing with probabilistic dynamic demand. *AIIE Transactions*, 13(2):132–137, 1981.

(Balfaqih *et al.*, 2016) Hasan Balfaqih, Zulkifli Mohd Nopiah, Nizaroyani Saibani, and Malak T Al-Nory. Review of supply chain performance measurement systems: 1998–2015. *Computers in Industry*, 82:135–150, 2016.

(Balintfy, 1964) Joseph L Balintfy. On a basic class of multi-item inventory problems. *Management science*, 10(2):287–297, 1964.

(Beamon, 1999) Benita M Beamon. Measuring supply chain performance. *International journal of operations & production management*, 1999.

(Benjaafar *et al.*, 2011) Saif Benjaafar, Mohsen ElHafsi, Chung-Yee Lee, and Weihua Zhou. Optimal control of an assembly system with multiple stages and multiple demand classes. *Operations Research*, 59(2):522–529, 2011.

(Bensoussan *et al.*, 2007) Alain Bensoussan, Metin Çakanyıldırım, and Suresh P. Sethi. A Multiperiod Newsvendor Problem with Partially Observed Demand. *Mathematics of OR*, 32(2):322–344, May 2007.

(Berling and Martínez-de Albéniz, 2015) Peter Berling and Victor Martínez-de Albéniz. Dynamic speed optimization in supply chains with stochastic demand. *Transportation Science*, 50(3):1114–1127, 2015.

(Berling and Martínez-de Albéniz, 2016) Peter Berling and Victor Martínez-de Albéniz. A characterization of optimal base-stock levels for a multistage serial supply chain. *Naval Research Logistics (NRL)*, 63(1):32–46, 2016.

(Bertsimas and Thiele, 2006) Dimitris Bertsimas and Aurélie Thiele. A robust optimization approach to inventory theory. *Operations research*, 54(1):150–168, 2006.

(Beyer *et al.*, 2016) Betsy Beyer, Chris Jones, Jennifer Petoff, and Niall Richard Murphy. *Site Reliability Engineering: How Google Runs Production Systems.* " O'Reilly Media, Inc.", 2016.

(Bollapragada and Morton, 1999) Srinivas Bollapragada and Thomas E Morton. A simple heuristic for computing nonstationary (s, S) policies. *Operations Research*, 47(4):576–584, 1999.

(Bowersox *et al.*, 2002) Donald J Bowersox, David J Closs, and M Bixby Cooper. *Supply chain logistics management*, volume 2. McGraw-Hill New York, NY, 2002.

(Boysen *et al.*, 2007) Nils Boysen, Malte Fliedner, and Armin Scholl. A classification of assembly line balancing problems. *European journal of operational research*, 183(2):674–693, 2007.

(Brewer, 2000) Peter C Brewer. Using the balanced scorecard to measure supply chain performance peter c brewer; thomas wspeh. *Journal of Business logistics*, 21(1):75, 2000.

(Brown, 2011) Marisa Brown. Inventory optimization: Show me the money. *Supply Chain Management Review*, 15(4):47–49, 2011.

(Brucker *et al.*, 1999) Peter Brucker, Andreas Drexl, Rolf Möhring, Klaus Neumann, and Erwin Pesch. Resource-constrained project scheduling: Notation, classification, models, and methods. *European journal of operational research*, 112(1):3–41, 1999.

(Buzacott and Shanthikumar, 1993) John A Buzacott and J George Shanthikumar. *Stochastic models of manufacturing systems*, volume 4. Prentice Hall Englewood Cliffs, NJ, 1993.

(Byrd *et al.*, 2011) Richard H Byrd, Gillian M Chin, Will Neveitt, and Jorge Nocedal. On the use of stochastic hessian information in unconstrained optimization. *SIAM Journal on Optimization*, 21(3):977–995, 2011.

(Cannella and Ciancimino, 2010) Salvatore Cannella and Elena Ciancimino. On the bullwhip avoidance phase: supply chain collaboration and order smoothing. *International Journal of Production Research*, 48(22):6739–6776, 2010.

(Cheaitou and Van Delft, 2013) Ali Cheaitou and Christian Van Delft. Finite horizon stochastic inventory problem with dual sourcing: Near myopic and heuristics bounds. *International Journal of Production Economics*, 143(2):371–378, 2013.

(Chernonog and Goldberg, 2018) Tatyana Chernonog and Noam Goldberg. On the multi-product newsvendor with bounded demand distributions. *International Journal of Production Economics*, 203:38–47, 2018.

(Chopra *et al.*, 2013) Sunil Chopra, Peter Meindl, and Dharam Vir Kalra. *Supply chain management: strategy, planning, and operation*, volume 232. Pearson Boston, MA, 2013.

(Chu and Beasley, 1998) Paul C Chu and John E Beasley. A genetic algorithm for the multidimensional knapsack problem. *Journal of heuristics*, 4(1):63–86, 1998.

(Clark and Scarf, 1960) Andrew J Clark and Herbert Scarf. Optimal policies for a multi-echelon inventory problem. *Management science*, 6(4):475–490, 1960.

(Cohen-Hillel and Yedidsion, 2018) Tamar Cohen-Hillel and Liron Yedidsion. The periodic joint replenishment problem is strongly NP-hard. *Mathematics of Operations Research*, 2018.

(Copil *et al.*, 2017) Karina Copil, Martin Wörbelauer, Herbert Meyr, and Horst Tempelmeier. Simultaneous lotsizing and scheduling problems: a classification and review of models. *OR spectrum*, 39(1):1–64, 2017.

(CoresightResearch, 2019) CoresightResearch. Making the case for inventory optimization. https://coresight.com/research/us-retailer-survey-revealing-the-hidden-costs-of-poor-inventory-management-2/, 2019. Accessed: 2020-05-05.

(Cuthbertson and Piotrowicz, 2008) Richard Cuthbertson and Wojciech Piotrowicz. Supply chain best practices–identification and categorisation of measures and benefits. *International Journal of Productivity and Performance Management*, 57(5):389–404, 2008.

(de Kok *et al.*, 2018) Ton de Kok, Christopher Grob, Marco Laumanns, Stefan Minner, Jörg Rambau, and Konrad Schade. A typology and literature review on stochastic multi-echelon inventory models. *European Journal of Operational Research*, 269(3):955–983, 2018.

(Deb, 2014) Kalyanmoy Deb. Multi-objective Optimization. In *Search Methodologies*, pages 403–449. Springer, Boston, MA, 2014. DOI: 10.1007/978-1-4614-6940-7_15.

(Defourny *et al.*, 2012) Boris Defourny, Damien Ernst, and Louis Wehenkel. Multistage stochastic programming: A scenario tree based approach to planning under uncertainty. In *Decision theory models for applications in artificial intelligence: concepts and solutions*, pages 97–143. IGI Global, 2012.

(Deloitte, 2019) Deloitte. Supply chain planning 2025. `https://www2.deloitte.com/content/dam/Deloitte/de/Documents/energy-resources/Supply-Chain-Planning-2025-Whitepaper-Deloitte.pdf`, 2019. Accessed: 2019-12-10.

(Diks and De Kok, 1998) Erik Bas Diks and AG De Kok. Optimal control of a divergent multi-echelon inventory system. *European journal of operational research*, 111(1):75–97, 1998.

(Dong, 2018) James Dong. Robust multi-product newsvendor problem under a global budget of uncertainty. 2018.

(Dujak *et al.*, 2017) Davor Dujak, Marina Kresoja, and Jelena Franjković. Space management in category management: A comparative analysis of retailers in the subcategory of pickled and preserved vegetables. *Strategic Management*, 22(1):60–72, 2017.

(Dyckhoff, 1990) Harald Dyckhoff. A typology of cutting and packing problems. *European Journal of Operational Research*, 44(2):145–159, 1990.

(Eppen, 1981) Gary Eppen. Centralized ordering policies in a multi-warehouse system with lead times and random demand. *Multi-Level Production/Inventory Control Systems*, pages 51–67, 1981.

(Erlebacher, 2000) Steven J Erlebacher. Optimal and heuristic solutions for the multi-item newsvendor problem with a single capacity constraint. *Production and Operations Management*, 9(3):303–318, 2000.

(Ernst and Kouvelis, 1999) Ricardo Ernst and Panagiotis Kouvelis. The effects of selling packaged goods on inventory decisions. *Management Science*, 45(8):1142–1155, 1999.

(Estampe *et al.*, 2013) Dominique Estampe, Samir Lamouri, Jean-Luc Paris, and Sakina Brahim-Djelloul. A framework for analysing supply chain performance evaluation models. *International Journal of Production Economics*, 142(2):247–258, 2013.

(Feng and Xiao, 2000) Youyi Feng and B Xiao. A new algorithm for computing optimal (s, S) policies in a stochastic single item/location inventory system. *IIE Transactions*, 32(11):1081–1090, 2000.

(Feng *et al.*, 2015) Haolin Feng, Qi Wu, Kumar Muthuraman, and Vinayak Deshpande. Replenishment policies for multi-product stochastic inventory systems with correlated demand and joint-replenishment costs. *Production and Operations Management*, 24(4):647–664, 2015.

(Firoozi, 2018) Mehdi Firoozi. *Multi-echelon Inventory optimization under supply and demand uncertainty*. PhD thesis, 2018.

(Firouz *et al.*, 2017) Mohammad Firouz, Burcu B Keskin, and Sharif H Melouk. An integrated supplier selection and inventory problem with multi-sourcing and lateral transshipments. *Omega*, 70:77–93, 2017.

(Fox *et al.*, 2006) Edward J Fox, Richard Metters, and John Semple. Optimal inventory policy with two suppliers. *Operations Research*, 54(2):389–393, 2006.

(Galasso *et al.*, 2016) François Galasso, Yves Ducq, Matthieu Lauras, Didier Gourc, and Mamadou Camara. A method to select a successful interoperability solution through a simulation approach. *Journal of Intelligent Manufacturing*, 27(1):217–229, 2016.

(Gao *et al.*, 2014a) Long Gao, Douglas J Thomas, and Michael B Freimer. Optimal inventory control with retail pre-packs. *Production and Operations Management*, 23(10):1761–1778, 2014.

(Gao *et al.*, 2014b) Long Gao, Douglas J. Thomas, and Michael B. Freimer. Optimal Inventory Control with Retail Pre-Packs. *Prod Oper Manag*, 23(10):1761–1778, October 2014.

(Gawankar *et al.*, 2013) Shradha Gawankar, Sachin S Kamble, and Rakesh Verma. Effect of supply chain management practices on supply chain profitability: an empirical investigation using structural equation modelling in indian retail sector. *International Journal of Services and Operations Management*, 16(2):145–173, 2013.

(Gayraud *et al.*, 2015) Fabrice Gayraud, Nathalie Grangeon, Laurent Deroussi, and Sylvie Norre. Notation and classification for logistic network design models. *RAIRO-Operations Research*, 49(1):195–214, 2015.

(Gilotte *et al.*, 2018) Alexandre Gilotte, Clément Calauzènes, Thomas Nedelec, Alexandre Abraham, and Simon Dollé. Offline a/b testing for recommender systems. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, pages 198–206. ACM, 2018.

(Glasserman and Tayur, 1994) Paul Glasserman and Sridhar Tayur. The stability of a capacitated, multi-echelon production-inventory system under a base-stock policy. *Operations Research*, 42(5):913–925, 1994.

(Glasserman and Tayur, 1995) Paul Glasserman and Sridhar Tayur. Sensitivity analysis for base-stock levels in multiechelon production-inventory systems. *Management Science*, 41(2):263–281, 1995.

(Gopal and Thakkar, 2012) PRC Gopal and Jitesh Thakkar. A review on supply chain performance measures and metrics: 2000-2011. *International Journal of Productivity and Performance Management*, 61(5):518–547, 2012.

(Graham *et al.*, 1979) Ronald L Graham, Eugene L Lawler, Jan Karel Lenstra, and AHG Rinnooy Kan. Optimization and approximation in deterministic sequencing and scheduling: a survey. In *Annals of discrete mathematics*, volume 5, pages 287–326. Elsevier, 1979.

(Grahl *et al.*, 2016) Jörn Grahl, Stefan Minner, and Daniel Dittmar. Meta-heuristics for placing strategic safety stock in multi-echelon inventory with differentiated service times. *Annals of Operations Research*, 242(2):489–504, 2016.

(Graves and Schoenmeyr, 2016) Stephen C Graves and Tor Schoenmeyr. Strategic safety-stock placement in supply chains with capacity constraints. *Manufacturing & Service Operations Management*, 18(3):445–460, 2016.

(Graves and Willems, 2000) Stephen C Graves and Sean P Willems. Optimizing strategic safety stock placement in supply chains. *Manufacturing & Service Operations Management*, 2(1):68–83, 2000.

(Graves and Willems, 2005) Stephen C Graves and Sean P Willems. Optimizing the supply chain configuration for new products. *Management science*, 51(8):1165–1180, 2005.

(Graves and Willems, 2008) Stephen C Graves and Sean P Willems. Strategic inventory placement in supply chains: Nonstationary demand. *Manufacturing & Service Operations Management*, 10(2):278–287, 2008.

(Gunasekaran *et al.*, 2001) Angappa Gunasekaran, Chaitali Patel, and Ercan Tirtiroglu. Performance measures and metrics in a supply chain environment. *International journal of operations & production Management*, 2001.

(Gunasekaran *et al.*, 2004) Angappa Gunasekaran, Christopher Patel, and Ronald E McGaughey. A framework for supply chain performance measurement. *International journal of production economics*, 87(3):333–347, 2004.

(Gupta *et al.*, 2007) R Gupta, KK Gupta, BR Jain, and RK Garg. ABC and VED analysis in medical stores inventory control. *Medical Journal Armed Forces India*, 63(4):325–327, 2007.

(Hadley and Whitin, 1963) George Hadley and Thomson M Whitin. Analysis of inventory systems. Technical report, 1963.

(Halman, 2015) Nir Halman. Provably near-optimal approximation schemes for sample-based dynamic programs with emphasis on stochastic inventory control models, 2015.

(Harris, 1913) Ford W Harris. How many parts to make at once. 1913.

(Hilli *et al.*, 2007) Petri Hilli, Matti Koivu, Teemu Pennanen, and Antero Ranne. A stochastic programming model for asset liability management of a finnish pension company. *Annals of Operations Research*, 152(1):115, 2007.

(Hindsberger and Philpott, 2014) Magnus Hindsberger and Andy B Philpott. ReSa: A method for solving multistage stochastic linear programs. *Journal of Applied Operational Research*, 6(1):2–15, 2014.

(Homem-de Mello and Bayraksan, 2014) Tito Homem-de Mello and Güzin Bayraksan. Monte carlo sampling-based methods for stochastic optimization. *Surveys in Operations Research and Management Science*, 19(1):56–85, 2014.

(Humair and Willems, 2006) Salal Humair and Sean P Willems. Optimizing strategic safety stock placement in supply chains with clusters of commonality. *Operations Research*, 54(4):725–742, 2006.

(Humair and Willems, 2011) Salal Humair and Sean P Willems. Optimizing strategic safety stock placement in general acyclic networks. *Operations Research*, 59(3):781–787, 2011.

(Ignall, 1969) Edward Ignall. Optimal continuous review policies for two product inventory systems with joint setup costs. *Management Science*, 15(5):278–283, 1969.

(Janakiraman and Muckstadt, 2009) Ganesh Janakiraman and John A Muckstadt. A decomposition approach for a class of capacitated serial systems. *Operations Research*, 57(6):1384–1393, 2009.

(Jirutitijaroen and Singh, 2008) Panida Jirutitijaroen and Chanan Singh. Reliability constrained multi-area adequacy planning using stochastic programming with sample-average approximations. *IEEE Transactions on Power Systems*, 23(2):504–513, 2008.

(Johansen and Melchiors, 2003) Søren Glud Johansen and Philip Melchiors. Can-order policy for the periodic-review joint replenishment problem. *Journal of the Operational Research Society*, 54(3):283–290, 2003.

(Jucker and Rosenblatt, 1985) James V Jucker and Meir J Rosenblatt. Single-period inventory models with demand uncertainty and quantity discounts: Behavioral implications and a new solution procedure. *Naval Research Logistics Quarterly*, 32(4):537–550, 1985.

(Kaplan and Norton, 1996) Robert S Kaplan and David P Norton. Using the balanced scorecard as a strategic management system, 1996.

(Kaplan, 2009) Robert S Kaplan. Conceptual foundations of the balanced scorecard. *Handbooks of management accounting research*, 3:1253–1269, 2009.

(Karlin, 1960) Samuel Karlin. Dynamic inventory policy with varying stochastic demands. *Management Science*, 6(3):231–258, 1960.

(Kendall, 1953) David G Kendall. Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded markov chain. *The Annals of Mathematical Statistics*, pages 338–354, 1953.

(Kenyon and Morton, 2003) Astrid S Kenyon and David P Morton. Stochastic vehicle routing with random travel times. *Transportation Science*, 37(1):69–82, 2003.

(Khouja and Mehrez, 1996) Moutaz Khouja and Abraham Mehrez. A multi-product constrained newsboy problem with progressive multiple discounts. *Computers & industrial engineering*, 30(1):95–101, 1996.

(Khouja, 1995) Moutaz Khouja. The newsboy problem under progressive multiple discounts. *European Journal of Operational Research*, 84(2):458–466, 1995.

(Khouja, 2000) Moutaz J Khouja. Optimal ordering, discounting, and pricing in the single-period problem. *International Journal of Production Economics*, 65(2):201–216, 2000.

(Kim *et al.*, 2015) Gitae Kim, Kan Wu, and Edward Huang. Optimal inventory control in a multi-period newsvendor problem with non-stationary demand. *Advanced Engineering Informatics*, 29(1):139–145, January 2015.

(Klosterhalfen and Minner, 2007) Steffen Klosterhalfen and Stefan Minner. Comparison of stochastic-and guaranteed-service approaches to safety stock optimization in supply chains. In *Operations Research Proceedings 2006*, pages 485–490. Springer, 2007.

(Klosterhalfen and Minner, 2010) Steffen Klosterhalfen and Stefan Minner. Safety stock optimisation in distribution systems: a comparison of two competing approaches. *International Journal of Logistics: Research and Applications*, 13(2):99–120, 2010.

(Klosterhalfen *et al.*, 2013) Steffen T Klosterhalfen, Daniel Dittmar, and Stefan Minner. An integrated guaranteed-and stochastic-service approach to inventory optimization in supply chains. *European Journal of Operational Research*, 231(1):109–119, 2013.

(Kurien and Qureshi, 2011) Georgy P Kurien and Muhammad N Qureshi. Study of performance measurement practices in supply chain management. *International Journal of Business, Management and Social Sciences*, 2(4):19–34, 2011.

(Lau and Lau, 1995) Hon-Shiang Lau and Amy Lau. The multi-product multi-constraint newsboy problem: Applications, formulation and solution. *Journal of operations management*, 13(2):153–162, 1995.

(Lau and Lau, 1996) Hon-Shiang Lau and Amy Hing-Ling Lau. The newsstand problem: A capacitated multiple-product single-period inventory problem. *European Journal of Operational Research*, 94(1):29–42, 1996.

(Leblanc, 2017) R Leblanc. How amazon is changing supply chain management. *The Balance Small Business*, 2017.

(Lesnaia, 2004) Ekaterina Lesnaia. *Optimizing safety stock placement in general network supply chains*. PhD thesis, Massachusetts Institute of Technology, 2004.

(Levi *et al.*, 2006) Retsef Levi, Robin O Roundy, and David B Shmoys. Provably near-optimal sampling-based algorithms for stochastic inventory control models. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 739–748. ACM, 2006.

(Levi *et al.*, 2007a) Retsef Levi, Martin Pál, Robin O Roundy, and David B Shmoys. Approximation algorithms for stochastic inventory control models. *Mathematics of Operations Research*, 32(2):284–302, 2007.

(Levi *et al.*, 2007b) Retsef Levi, Robin O Roundy, and David B Shmoys. Provably near-optimal sampling-based policies for stochastic inventory control models. *Mathematics of Operations Research*, 32(4):821–839, 2007.

(Levi *et al.*, 2007c) Retsef Levi, Robin O. Roundy, and David B. Shmoys. Provably Near-Optimal Sampling-Based Policies for Stochastic Inventory Control Models. *Mathematics of OR*, 32(4):821–839, November 2007.

(Levina *et al.*, 2010) Tatsiana Levina, Yuri Levin, Jeff McGill, Mikhail Nediak, and Vladimir Vovk. Weak aggregating algorithm for the distribution-free perishable inventory problem. *Operations Research Letters*, 38(6):516–521, November 2010.

(Li and Jiang, 2012) Haitao Li and Dali Jiang. New model and heuristics for safety stock placement in general acyclic supply chain networks. *Computers & Operations Research*, 39(7):1333–1344, 2012.

(Linoff and Berry, 2011) Gordon S Linoff and Michael JA Berry. *Data mining techniques: for marketing, sales, and customer relationship management.* John Wiley & Sons, 2011.

(Lockamy III and McCormack, 2004) Archie Lockamy III and Kevin McCormack. Linking scor planning practices to supply chain performance: An exploratory study. *International journal of operations & production management*, 24(12):1192–1218, 2004.

(Lourenço *et al.*, 2003) Helena R Lourenço, Olivier C Martin, and Thomas Stützle. Iterated local search. In *Handbook of metaheuristics*, pages 320–353. Springer, 2003.

(Maestrini *et al.*, 2017) Vieri Maestrini, Davide Luzzini, Paolo Maccarrone, and Federico Caniato. Supply chain performance measurement systems: A systematic review and research agenda. *International Journal of Production Economics*, 183:299–315, 2017.

(Magnanti *et al.*, 2006) Thomas L Magnanti, Zuo-Jun Max Shen, Jia Shu, David Simchi-Levi, and Chung-Piaw Teo. Inventory placement in acyclic supply chain networks. *Operations Research Letters*, 34(2):228–238, 2006.

(Martin and others, 2010) Benjamin Martin et al. Inventory management, metrics, and simulation. 2010.

(McCardle *et al.*, 2007) Kevin F McCardle, Kumar Rajaram, and Christopher S Tang. Bundling retail products: Models and analysis. *European Journal of Operational Research*, 177(2):1197–1217, 2007.

(Millstein *et al.*, 2014) Mitchell A Millstein, Liu Yang, and Haitao Li. Optimizing abc inventory grouping decisions. *International Journal of Production Economics*, 148:71–80, 2014.

(Minner, 1997) Stefan Minner. Dynamic programming algorithms for multi-stage safety stock optimization. *Operations-Research-Spektrum*, 19(4):261–271, 1997.

(Minner, 2001) Stefan Minner. Strategic safety stocks in reverse logistics supply chains. *International journal of production economics*, 71(1-3):417–428, 2001.

(Minner, 2012) Stefan Minner. *Strategic safety stocks in supply chains*, volume 490. Springer Science & Business Media, 2012.

(Mohammadivojdan and Geunes, 2018) Roshanak Mohammadivojdan and Joseph Geunes. The newsvendor problem with capacitated suppliers and quantity discounts. *European Journal of Operational Research*, 271(1):109–119, 2018.

(Molenaers *et al.*, 2012) An Molenaers, Herman Baets, Liliane Pintelon, and Geert Waeyenbergh. Criticality classification of spare parts: A case study. *International Journal of Production Economics*, 140(2):570–578, 2012.

(Moon and Silver, 2000) Ilkyeong Moon and Edward Allen Silver. The multi-item newsvendor problem with a budget constraint and fixed ordering costs. *Journal of the Operational Research Society*, 51(5):602–608, 2000.

(MTESFrance, 2019) MTESFrance. France logistique 2025. https://www.ecologique-solidaire.gouv.fr/france-logistique-2025, 2019. Accessed: 2020-14-20.

(Muharremoglu and Tsitsiklis, 2008) Alp Muharremoglu and John N Tsitsiklis. A single-unit decomposition approach to multiechelon inventory systems. *Operations Research*, 56(5):1089–1103, 2008.

(Nadar *et al.*, 2014) Emre Nadar, Mustafa Akan, and Alan Scheller-Wolf. Optimal structural results for assemble-to-order generalized m-systems. *Operations Research*, 62(3):571–579, 2014.

(Neely and Huang, 2010) Michael J Neely and Longbo Huang. Dynamic product assembly and inventory control for maximum profit. In *49th IEEE Conference on Decision and Control (CDC)*, pages 2805–2812. IEEE, 2010.

(Özen *et al.*, 2012) Ulaş Özen, Mustafa K Doğru, and S Armagan Tarim. Static-dynamic uncertainty strategy for a single-item stochastic inventory control problem. *Omega*, 40(3):348–357, 2012.

(Özkaya *et al.*, 2006) Banu Yüksel Özkaya, Ülkü Gürler, and Emre Berk. The stochastic joint replenishment problem: a new policy, analysis, and insights. *Naval Research Logistics (NRL)*, 53(6):525–546, 2006.

(Pan and Nagi, 2010) Feng Pan and Rakesh Nagi. Robust supply chain design under uncertain demand in agile manufacturing. *Computers & operations research*, 37(4):668–683, 2010.

(Paterson *et al.*, 2011) Colin Paterson, Gudrun Kiesmüller, Ruud Teunter, and Kevin Glazebrook. Inventory models with lateral transshipments: A review. *European Journal of Operational Research*, 210(2):125–136, 2011.

(Petropoulos *et al.*, 2019) Fotios Petropoulos, Xun Wang, and Stephen M Disney. The inventory performance of forecasting methods: Evidence from the m3 competition data. *International Journal of Forecasting*, 35(1):251–265, 2019.

(Pettey, 2019) Christy Pettey. Gartner top 8 supply chain technology trends for 2018. https://www.gartner.com/smarterwithgartner/gartner-top-8-supply-chain-technology-trends-for-2019/, 2019. Accessed: 2019-12-24.

(Pirkul and Narasimhan, 1986) Hasan Pirkul and Sridhar Narasimhan. Efficient algorithms for the multiconstraint general knapsack problem. *IIE transactions*, 18(2):195–203, 1986.

(Porter, 1985) Michael E Porter. Value chain. *The Value Chain and Competitive advantage: creating and sustaining superior performance*, 1985.

(Protopappa-Sieke and Seifert, 2010) Margarita Protopappa-Sieke and Ralf W Seifert. Interrelating operational and financial performance measurements in inventory control. *European Journal of Operational Research*, 204(3):439–448, 2010.

(Puchinger *et al.*, 2010) Jakob Puchinger, Günther R Raidl, and Ulrich Pferschy. The multidimensional knapsack problem: Structure and algorithms. *INFORMS Journal on Computing*, 22(2):250–265, 2010.

(Puterman, 2014) Martin L Puterman. *Markov Decision Processes.: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 2014.

(Qin *et al.*, 2011) Yan Qin, Ruoxuan Wang, Asoo J Vakharia, Yuwen Chen, and Michelle MH Seref. The newsvendor problem: Review and directions for future research. *European Journal of Operational Research*, 213(2):361–374, 2011.

(Rahdar *et al.*, 2018) Mohammad Rahdar, Lizhi Wang, and Guiping Hu. A tri-level optimization model for inventory control with uncertain demand and lead time. *International Journal of Production Economics*, 195:96–105, 2018.

(Rao and Rao, 2009) Madhusudhana C Rao and K Prahlada Rao. Inventory turnover ratio as a supply chain performance measure. *Serbian Journal of Management*, 4(1):41–50, 2009.

(Ren *et al.*, 2017) Shuyun Ren, Hau-Ling Chan, and Pratibha Ram. A Comparative Study on Fashion Demand Forecasting Models with Multiple Sources of Uncertainty. *Ann Oper Res*, 257(1-2):335–355, October 2017.

(Rosenthal *et al.*, 1995) Edward C Rosenthal, James L Zydiak, and Sohail S Chaudhry. Vendor selection with bundling. *Decision Sciences*, 26(1):35–48, 1995.

(Rosling, 1989) Kaj Rosling. Optimal inventory policies for assembly systems under random demands. *Operations Research*, 37(4):565–579, 1989.

(Rossi *et al.*, 2015) Roberto Rossi, Onur A Kilic, and S Armagan Tarim. Piecewise linear approximations for the static–dynamic uncertainty strategy in stochastic lot-sizing. *Omega*, 50:126–140, 2015.

(Royset and Polak, 2004) Johannes O Royset and E Polak. Reliability-based optimal design using sample average approximations. *Probabilistic Engineering Mechanics*, 19(4):331–343, 2004.

(Sahu and Veerapen, 2019) Rabin Sahu and Nadarajen Veerapen. Stochastic joint replenishment planning optimization with an iterated local search. In *International Workshop on Stochastic Local Search Algorithms-SLS*, 2019.

(Sahu *et al.*, 2018) Rabin Sahu, Alexandre Gerussi, Clarisse Dhaenens, Jianhua Zhang, and Manuel Davy. Multi-objective optimization of retail promotional ordering with probabilistic forecast. In *ILS2018-International Conference on Information Systems, Logistics and Supply Chain*, 2018.

(Sahu *et al.*, 2020a) Rabin Sahu, Clarisse Dhaenens, and Nadarajen Veerapen. A performance evaluation framework for inventory management systems. In *Annual conference of the Belgian Operational Research Society-ORBEL*, 2020.

(Sahu *et al.*, 2020b) Rabin Sahu, Clarisse Dhaenens, Nadarajen Veerapen, and Manuel Davy. An approximate method for integrated stochastic replenishment planning with supplier selection. In *International Conference on Operations Research and Enterprise Systems-ICORES*, 2020.

(Santoso *et al.*, 2005) Tjendera Santoso, Shabbir Ahmed, Marc Goetschalckx, and Alexander Shapiro. A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, 167(1):96–115, 2005.

(Scarf, 1959) Herbert Scarf. The optimality of (5, 5) policies in the dynamic inventory problem. 1959.

(Shapiro, 2003) Alexander Shapiro. Inference of statistical bounds for multistage stochastic programming problems. *Mathematical Methods of Operations Research*, 58(1):57–68, 2003.

(Shapiro, 2011) Alexander Shapiro. Analysis of stochastic dual dynamic programming method. *European Journal of Operational Research*, 209(1):63–72, 2011.

(Sheikhzadeh and Elahi, 2013) Mehdi Sheikhzadeh and Ehsan Elahi. Product bundling: Impacts of product heterogeneity and risk considerations. *International Journal of Production Economics*, 144(1):209–222, 2013.

(Shepherd and Günter, 2010) Craig Shepherd and Hannes Günter. Measuring supply chain performance: current research and future directions. In *Behavioral Operations in Planning and Scheduling*, pages 105–121. Springer, 2010.

(Shu and Karimi, 2009) Jia Shu and IA Karimi. Efficient heuristics for inventory placement in acyclic networks. *Computers & Operations Research*, 36(11):2899–2904, 2009.

(Silver and Moon, 2001) Edward A Silver and Ilkyeong Moon. The multi-item single period problem with an initial stock of convertible units. *European Journal of Operational Research*, 132(2):466–477, 2001.

(Silver *et al.*, 1998) Edward Allen Silver, David F Pyke, and Rein Peterson. *Inventory management and production planning and scheduling*, volume 3. Wiley New York, 1998.

(Simchi-Levi *et al.*, 2000) D Simchi-Levi, P Kaminsky, and E Simchi-Levi. Designing and managing the supply chain: Concepts. *Strategies, and Case Studies*, 2000.

(Simpson Jr, 1958) Kenneth F Simpson Jr. In-process inventories. *Operations Research*, 6(6):863–873, 1958.

(Stangl and Thonemann, 2017) Tobias Stangl and Ulrich W Thonemann. Equivalent inventory metrics: A behavioral perspective. *Manufacturing & Service Operations Management*, 19(3):472–488, 2017.

(Taleizadeh *et al.*, 2008) Ata Allah Taleizadeh, Seyed Taghi Akhavan Niaki, and V Hosseini. The multi-product multi-constraint newsboy problem with incremental discount and batch order. *Asian Journal of Applied Sciences*, 1:110–122, 2008.

(Torabi *et al.*, 2012) Seyed Ali Torabi, Seyed Morteza Hatefi, and B Saleck Pay. ABC inventory classification in the presence of both quantitative and qualitative criteria. *Computers & Industrial Engineering*, 63(2):530–537, 2012.

(Vairaktarakis, 2000) George L Vairaktarakis. Robust multi-item newsboy models with a budget constraint. *International Journal of Production Economics*, 66(3):213–226, 2000.

(Van Houtum and Kranenburg, 2015) Geert-Jan Van Houtum and Bram Kranenburg. *Spare parts inventory control under system availability constraints*, volume 227. Springer, 2015.

(Veinott Jr and Wagner, 1965) Arthur F Veinott Jr and Harvey M Wagner. Computing optimal (s, S) inventory policies. *Management Science*, 11(5):525–552, 1965.

(Veinott, 1966) Arthur F Veinott, Jr. On the opimality of (s,s) inventory policies: New conditions and a new proof. *SIAM Journal on Applied Mathematics*, 14(5):1067–1083, 1966.

(Verweij *et al.*, 2003) Bram Verweij, Shabbir Ahmed, Anton J Kleywegt, George Nemhauser, and Alexander Shapiro. The sample average approximation method applied to stochastic routing problems: a computational study. *Computational Optimization and Applications*, 24(2-3):289–333, 2003.

(Viswanathan, 1997) S Viswanathan. Note. periodic review (s, S) policies for joint replenishment inventory systems. *Management Science*, 43(10):1447–1454, 1997.

(Wäscher *et al.*, 2007) Gerhard Wäscher, Heike Haußner, and Holger Schumann. An improved typology of cutting and packing problems. *European journal of operational research*, 183(3):1109–1130, 2007.

(Xiang *et al.*, 2018) Mengyuan Xiang, Roberto Rossi, Belen Martin-Barragan, and S Armagan Tarim. Computing non-stationary (s, S) policies using mixed integer linear programming. *European Journal of Operational Research*, 271(2):490–500, 2018.

(Yang and Kim, 2018) Young Hyeon Yang and Jong Soo Kim. An adaptive joint replenishment policy for items with non-stationary demands. *Operational Research*, pages 1–20, 2018.

(Yao and Minner, 2017) Man Yao and Stefan Minner. Review of multi-supplier inventory models in supply chain management: An update. *Available at SSRN 2995134*, 2017.

(Zhang *et al.*, 2012) Wei Zhang, Zhongsheng Hua, and Saif Benjaafar. Optimal inventory control with dual-sourcing, heterogeneous ordering costs and order size constraints. *Production and Operations Management*, 21(3):564–575, 2012.

(Zhang, 2010) Guoqing Zhang. The multi-product newsboy problem with supplier quantity discounts and a budget constraint. *European Journal of Operational Research*, 206(2):350–360, 2010.

(Zheng and Federgruen, 1991) Yu-Sheng Zheng and Awi Federgruen. Finding optimal (s, S) policies is about as simple as evaluating a single policy. *Operations Research*, 39(4):654–665, 1991.

(Zheng, 1992) Yu-Sheng Zheng. On properties of stochastic inventory systems. *Management science*, 38(1):87–103, 1992.

Appendices

A. Proof for Property 1

Property 1: *X, Y and Ξ_z are piecewise linear convex functions in q for any s_0, Ω and $q > 0$.*

Proof. Here we prove that the functions have increasing slopes. Let

$$X_0 = \sum_{t=1}^{\zeta_0} H\Big(s_0 + q - \sum_{\tau=1}^{t} F_\tau\Big) \tag{1}$$

$$X_1 = \sum_{t=1}^{\zeta_1} H\Big(s_0 + q + 1 - \sum_{\tau=1}^{t} F_\tau\Big) \tag{2}$$

$$X_2 = \sum_{t=1}^{\zeta_2} H\Big(s_0 + q + 2 - \sum_{\tau=1}^{t} F_\tau\Big) \tag{3}$$

$$\Delta_1 = X_1 - X_0 \tag{4}$$

$$\Delta_2 = X_2 - X_1 \tag{5}$$

The following inequalities hold true regarding the above equations.

$$X_2 \geq X_1 \geq X_0 \tag{6}$$

$$\zeta_2 \geq \zeta_1 \geq \zeta_0 \tag{7}$$

$$\Delta_1 \geq 0 \tag{8}$$

$$\Delta_2 \geq 0 \tag{9}$$

The relationships between Δ_1 and Δ_2 are as follows

$$\zeta_1 \geq \Delta_1 \geq \zeta_0 \tag{10}$$

$$\zeta_2 \geq \Delta_2 \geq \zeta_1 \tag{11}$$

Therefore,

$$\Delta_2 \geq \Delta_1 \tag{12}$$

This implies that X has an increasing slope and therefore, it is convex.
Similarly, let

$$Y_0 = \sum_{t=\zeta_0+1}^{T} W\Big(\sum_{\tau=1}^{t} F_\tau - s_0 - q\Big) \tag{13}$$

$$Y_1 = \sum_{t=\zeta_1+1}^{T} W\Big(\sum_{\tau=1}^{t} F_\tau - s_0 - q - 1\Big) \tag{14}$$

$$Y_2 = \sum_{t=\zeta_2+1}^{T} W\Big(\sum_{\tau=1}^{t} F_\tau - s_0 - q - 2\Big) \tag{15}$$

$$\Delta_1 = Y_1 - Y_0 \tag{16}$$

$$\Delta_2 = Y_2 - Y_1 \tag{17}$$

A. Proof for Property 1

The following statements can be made about the above equations.

$$Y_2 \leq Y_1 \leq Y_0 \tag{18}$$

$$\zeta_2 \geq \zeta_1 \geq \zeta_0 \tag{19}$$

$$\Delta_1 \leq 0 \tag{20}$$

$$\Delta_2 \leq 0 \tag{21}$$

The relationships between Δ_1 and Δ_2 are as follows

$$T - \zeta_1 \leq -\Delta_1 \leq T - \zeta_0 \tag{22}$$

$$T - \zeta_2 \leq -\Delta_2 \leq T - \zeta_1 \tag{23}$$

Therefore,

$$\Delta_2 \geq \Delta_1 \tag{24}$$

This implies that Y has an increasing slope and therefore, it is convex.
$K\alpha$ is constant and equal to K for $q > 0$.
As $\Xi_z = X_z + Y_z + K\alpha$, it is also convex. ■

B. Explanation of Equation 5.25

Proof. Recalling the equation

$$\zeta_z^* = \arg\min_{\zeta} \Xi_z(s_0, \Omega, q) = \left\lceil \frac{W\Omega}{H+W} \right\rceil \tag{25}$$

$$\Xi_z(s_0, q, \Omega) = \sum_{t=1}^{\zeta_z} H\Big(s_0 + q - \sum_{\tau=1}^{t} F_{z\tau}\Big) + \sum_{t=\zeta_z+1}^{\Omega} W\Big(\sum_{\tau=1}^{t} F_{z\tau} - s_0 - q\Big) + K\alpha \tag{26}$$

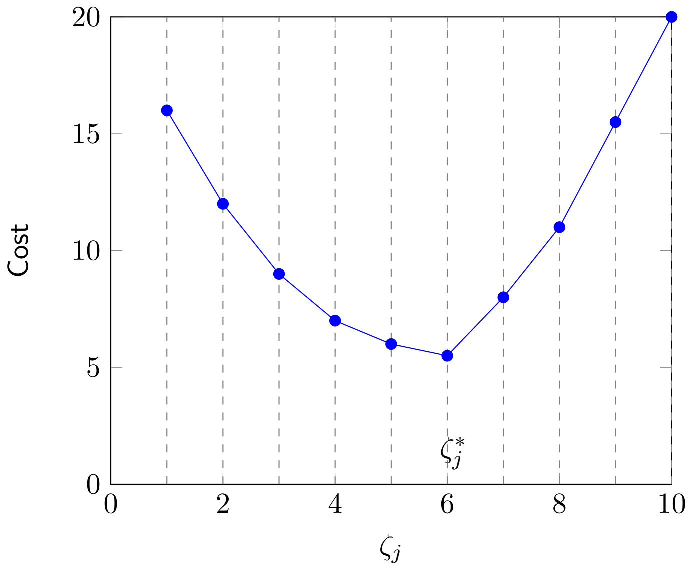


Figure 1: Illustration of demand samples. We present three distinct demand samples, which are random realizations of the demand during time $t = 1$ to $T = 10$.

As the above cost function is not differentiable everywhere, we find the minimum by finding the inflection point, i.e. the maximum q where the slope is less than or equal to zero. ζ_z can only take discrete values. The slope of the cost function is constant between two consecutive ζ_z. We can write

$$q^* = \arg\min \Xi_z(s_0, q, \Omega) \tag{27}$$

$$= \min q | (\Xi_z(s_0, q+1, \Omega) - \Xi_z(s_0, q, \Omega)) \geq 0 \tag{28}$$

Let us consider the case where

$$\Xi_z(s_0, q+1, \Omega) - \Xi_z(s_0, q, \Omega) = 0 \tag{29}$$

Solving (29) we get

$$W(\Omega - \zeta_z) - H\zeta_z = 0 \tag{30}$$

$$\zeta_z = \frac{W\Omega}{H+W} \tag{31}$$

The above equation does not necessarily yield an integer value for ζ_z. Since ζ_z takes only integer values, one of the following choices of ζ_z^* corresponds to the optimal order quantity q_z^*: $\zeta_z^* = \lfloor \zeta_z \rfloor$, $\zeta_z^* = \lceil \zeta_z \rceil$. Here, $\lceil x \rceil$ is the smallest integer greater than or equal to x and $\lfloor x \rfloor$ is the largest integer less than or equal to x. Let

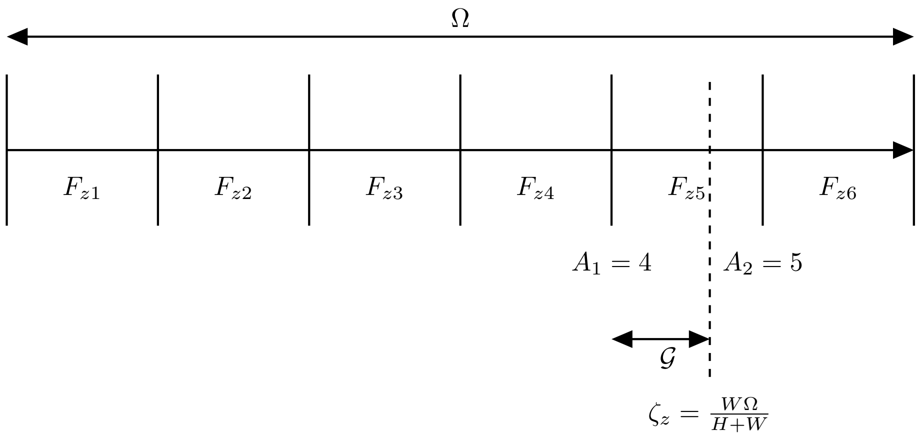
$$A_1 = \lfloor \zeta_z \rfloor \tag{32}$$

$$A_2 = \lceil \zeta_z \rceil \tag{33}$$

$$B_1 = \sum_{t=1}^{A_1} F_{zt} \tag{34}$$

$$B_2 = \sum_{t=1}^{A_2} F_{zt} \tag{35}$$

Figure 2 illustrates the cost components for different choices of ζ_z^* for $\Omega = 6$ and arbitrarily chosen $\frac{W}{W+H}$. Let $\mathcal{G} = \zeta_z - A_1$ be the fractional part. Then choosing the optimal period



$\zeta_j^* = \lfloor \zeta_z \rfloor, B = B_1$

Inventory	$F_{z2} + .. + F_{z4}$	$F_{z3} + F_{z4}$	F_{z4}	0	0	0
Backorder	0	0	0	0	F_{z5}	$F_{z5} + F_{z6}$

$\zeta_z^* = \lceil \zeta_z \rceil, B = B_2$

Inventory	$F_{z2} + .. + F_{z5}$	$F_{z3} + .. + F_{z5}$	$F_{z4} + F_{z5}$	F_{z5}	0	0
Backorder	0	0	0	0	0	F_{z6}

$\zeta_z^* = \zeta_z, B = B_1 + \mathcal{G}F_{z5}$

Inventory	$F_{z2} + .. + \mathcal{G}F_{z5}$	$F_{z3} + .. + \mathcal{G}F_{z5}$	$F_{z4} + \mathcal{G}F_{z5}$	$\mathcal{G}F_{z5}$	0	0
Backorder	0	0	0	0	$(1 - \mathcal{G})F_{z5}$	$(1 - \mathcal{G})F_{z5} + F_{z6}$

Figure 2: Illustration of cost components when we choose $\zeta_z^* = \lfloor \zeta_z \rfloor$, $\zeta_z^* = \lceil \zeta_z \rceil$ or $\zeta_z^* = \zeta_z$.

equal to A_1 will have a gain by reduction in inventory holding cost and a loss by increase in backorder penalty. The net increase in cost will be

$$W(\Omega - A_1)\mathcal{G}F_{zA_2} - HA_1\mathcal{G}F_{zA_2} \tag{36}$$

Similarly, if we choose the time period to be equal to A_2, we will have an increase in inventory holding cost and a decrease in backorder penalty. The net increase in cost will be

$$HA_1(1-\mathcal{G})F_{zA_2} - W(\Omega - A_1)(1-\mathcal{G})F_{zA_2} \tag{37}$$

If we replace A_1 in Equation 36 by $\zeta_z = \frac{W\Omega}{H+W}$, we get

$$= W(\Omega - \frac{W\Omega}{H+W})\mathcal{G}F_{zA_2} - H\frac{W\Omega}{H+W}\mathcal{G}F_{zA_2} \tag{38}$$

$$= W(\frac{H\Omega + W\Omega - W\Omega}{H+W})\mathcal{G}F_{zA_2} - H\frac{W\Omega}{H+W}\mathcal{G}F_{zA_2} \tag{39}$$

$$= \frac{WH\Omega}{H+W}\mathcal{G}F_{zA_2} - \frac{HW\Omega}{H+W}\mathcal{G}F_{zA_2} \tag{40}$$

$$= 0 \tag{41}$$

Since $A_1 \leq \zeta_z$, in Equation 36

$$W(\Omega - A_1)\mathcal{G}F_{zA_2} \geq W(\Omega - \zeta_z)\mathcal{G}F_{zA_2} \tag{42}$$

$$HA_1\mathcal{G}F_{zA_2} \leq H\zeta_z\mathcal{G}F_{zA_2} \tag{43}$$

From Equation 41, 42 and 43

$$W(\Omega - A_1)\mathcal{G}F_{zA_2} - HA_1\mathcal{G}F_{zA_2} \geq 0 \tag{44}$$

Similarly, if we replace A_1 in Equation 37 by $\zeta_z = \frac{W\Omega}{H+W}$, we get

$$= H\frac{W\Omega}{H+W}(1-\mathcal{G})F_{zA_2} - W(\Omega - \frac{W\Omega}{H+W})(1-\mathcal{G})F_{zA_2} \tag{45}$$

$$= \frac{HW\Omega}{H+W}(1-\mathcal{G})F_{zA_2} - W(\frac{H\Omega + W\Omega - W\Omega}{H+W})(1-\mathcal{G})F_{zA_2} \tag{46}$$

$$= \frac{HW\Omega}{H+W}(1-\mathcal{G})F_{zA_2} - \frac{HW\Omega}{H+W}(1-\mathcal{G})F_{zA_2} \tag{47}$$

$$= 0 \tag{48}$$

Since $A_1 \leq \zeta_z$, in Equation 37

$$HA_1(1-\mathcal{G})F_{zA_2} \leq H\zeta_z(1-\mathcal{G})F_{zA_2} \tag{49}$$

$$W(\Omega - A_1)(1-\mathcal{G})F_{zA_2} \geq W(\Omega - \zeta_z)(1-\mathcal{G})F_{zA_2} \tag{50}$$

From Equation 48, 49 and 50

$$HA_1(1-\mathcal{G})F_{zA_2} - W(\Omega - A_1)(1-\mathcal{G})F_{zA_2} \leq 0 \tag{51}$$

For any order quantity B such that $B_1 < B < B_2$, using Equation 44 and 51 we can conclude that

$$\tilde{C}_\Omega(s_0, B) \leq \tilde{C}_\Omega(s_0, B_1) \tag{52}$$

$$\tilde{C}_\Omega(s_0, B_2) \leq \tilde{C}_\Omega(s_0, B) \tag{53}$$

Any increase in q beyond B_2 will also have increase in inventory cost and a reduction in

B. Explanation of Equation 5.25

backorder cost. The net change will be

$$HA_2 - W(\Omega - A_2) \geq 0 \qquad \forall A_2 \geq \frac{W\Omega}{H+W} \tag{54}$$

$$\tag{55}$$

Therefore,

$$q^*_\Omega = B_2 \tag{56}$$

$$\zeta^*_z = A_2 = \left\lceil \frac{W\Omega}{H+W} \right\rceil \tag{57}$$

■

C. Additional Results from Chapter 5

Table 2: Comparative performance of different approaches for the problem instances when the realized standard deviation is higher than predicted. σ = Forecast standard deviation. $\hat{\sigma}$ = Realized standard deviation. It presents the ratio of C^{∞}_{π} to the minimum C^{∞} among the approaches.

H	W	K	λ	$\hat{\sigma}/\sigma$	DP vs. Minimum	ECA vs. Minimum	RCA vs. Minimum	OH-IECA vs. Minimum
1	10	0	5	1.05	1.000	1.273	1.269	1.007
				1.25	1.000	1.304	1.302	1.008
				1.50	1.004	1.271	1.284	1.000
				2.00	1.000	1.247	1.241	1.017
			10	1.05	1.044	1.208	2.830	1.000
				1.25	1.098	1.303	2.728	1.000
				1.50	1.125	1.329	2.491	1.000
				2.00	1.134	1.307	2.117	1.000
			15	1.05	1.000	1.302	1.110	1.004
				1.25	1.006	1.334	1.131	1.000
				1.50	1.009	1.322	1.148	1.000
				2.00	1.000	1.276	1.131	1.008
			20	1.05	1.000	1.210	2.150	1.006
				1.25	1.000	1.246	2.068	1.003
				1.50	1.004	1.260	1.955	1.000
				2.00	1.000	1.222	1.701	1.001
			50	1.05	1.000	1.133	2.296	1.025
				1.25	1.000	1.182	2.245	1.044
				1.50	1.000	1.213	2.079	1.065
				2.00	1.000	1.185	1.814	1.051
1	10	10	5	1.05	1.004	1.105	1.000	1.022
				1.25	1.018	1.118	1.000	1.036
				1.50	1.023	1.118	1.000	1.036
				2.00	1.044	1.162	1.000	1.061
			10	1.05	1.039	1.071	1.050	1.000
				1.25	1.012	1.046	1.011	1.000
				1.50	1.006	1.036	1.000	1.050
				2.00	1.000	1.038	1.002	1.051
			15	1.05	1.002	1.136	1.046	1.000
				1.25	1.003	1.158	1.061	1.000
				1.50	1.002	1.183	1.078	1.000
				2.00	1.003	1.177	1.074	1.000
			20	1.05	1.001	1.095	1.535	1.000
				1.25	1.000	1.130	1.558	1.003
				1.50	1.000	1.149	1.567	1.003
				2.00	1.002	1.149	1.478	1.000
			50	1.05	1.000	1.070	1.719	1.012
				1.25	1.000	1.112	1.754	1.023
				1.50	1.000	1.133	1.725	1.033
				2.00	1.000	1.147	1.630	1.049
0.1	1	10	5	1.05	1.029	1.035	1.036	1.000
				1.25	1.029	1.033	1.034	1.000
				1.50	1.032	1.033	1.032	1.000
				2.00	1.036	1.023	1.034	1.000
			10	1.05	1.001	1.001	1.025	1.000
				1.25	1.003	1.002	1.023	1.000
				1.50	1.001	1.001	1.020	1.000
				2.00	1.000	1.001	1.017	1.004
			15	1.05	1.010	1.003	1.017	1.000
				1.25	1.009	1.000	1.012	1.001

	1.50	1.011	1.002	1.010	1.000
	2.00	1.015	1.002	1.000	1.002
20	1.05	1.000	1.013	1.036	1.008
	1.25	1.000	1.006	1.031	1.008
	1.50	1.000	1.001	1.024	1.008
	2.00	1.008	1.000	1.025	1.013
50	1.05	1.001	1.007	1.481	1.000
	1.25	1.002	1.012	1.392	1.000
	1.50	1.000	1.011	1.302	1.001
	2.00	1.001	1.008	1.183	1.000

D. Additional Heuristic for Chapter 7

Algorithm 10: Greedy Heuristic 1 (GH1)

Input : $\mathbf{R}, \mathbf{F}, \mathbf{SL}$
Output: $\mathbf{Q}$
1 initialization $\mathbf{Q} \leftarrow 0, L_i \leftarrow SL_i F_i, \forall i \in \mathcal{N}$
repeat
2   select item $\hat{i}$ randomly : $\hat{i} \in \mathcal{N}$
  define $\mathcal{M}^* = \{m^* : R_{\hat{i}m^*} > 0\}$
  repeat
3     $\hat{m} = \arg\max_m R_{\hat{i}m}, m \in \mathcal{M}^*$
4   **if** $R_{\hat{i}\hat{m}} \leq L_{\hat{i}}$ *OR* $|\mathcal{M}^*| = 1$ **then**
      $Q_{\hat{m}} \leftarrow Q_{\hat{m}} + 1$
      $q_i \leftarrow \sum_{m \in \mathcal{M}} Q_m R_{im}, \forall i \in \mathcal{N}$
      $L_i \leftarrow L_i - q_i, \forall i \in \mathcal{N}$
5   **else**
        remove $\hat{m}$ from $\mathcal{M}^*$
      end
    end
  until $L_{\hat{i}} \leq 0$;
  remove $\hat{i}$ from $\mathcal{N}$
until $|\mathcal{N}| = 0$;

The detailed description of GH1 is as follows.

1.0. GH1 generates a feasible solution by taking the prepack configurations, item forecasts and service levels as inputs.

1.1. Initialize by setting the values of order quantity Q_m to zero, and values of remaining quantity L_i their respective service level targets.

1.2. Select an item $\hat{i}$ at random from the set $\mathcal{N}$ and then define the set of all prepacks containing item $\hat{i}$ as $\mathcal{M}^*$.

1.3. Select the prepack $\hat{m}$ from the set $\mathcal{M}^*$ that contains the maximum quantity of item $\hat{i}$.

1.4. If the quantity of item $\hat{i}$ contained in selected prepack $\hat{m}$ does not exceed its remaining quantity, or if the selected prepack is the last remaining prepack in $\mathcal{M}^*$, then increase the order quantity of prepack $\hat{m}$ by one. Then, update order quantities q_i and remaining quantities L_i.

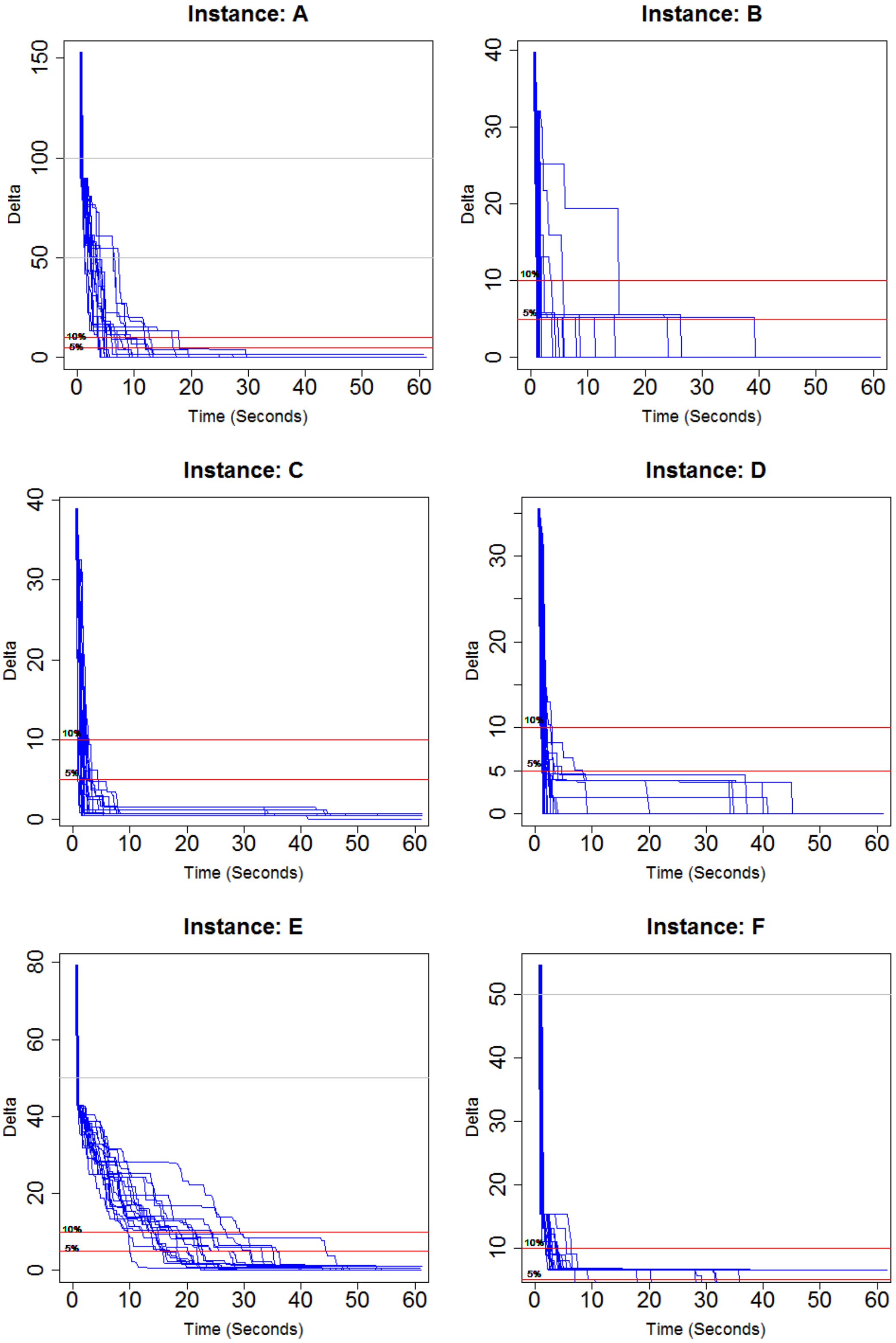
1.5. Else, remove the selected prepack $\hat{m}$ from $\mathcal{M}^*$.

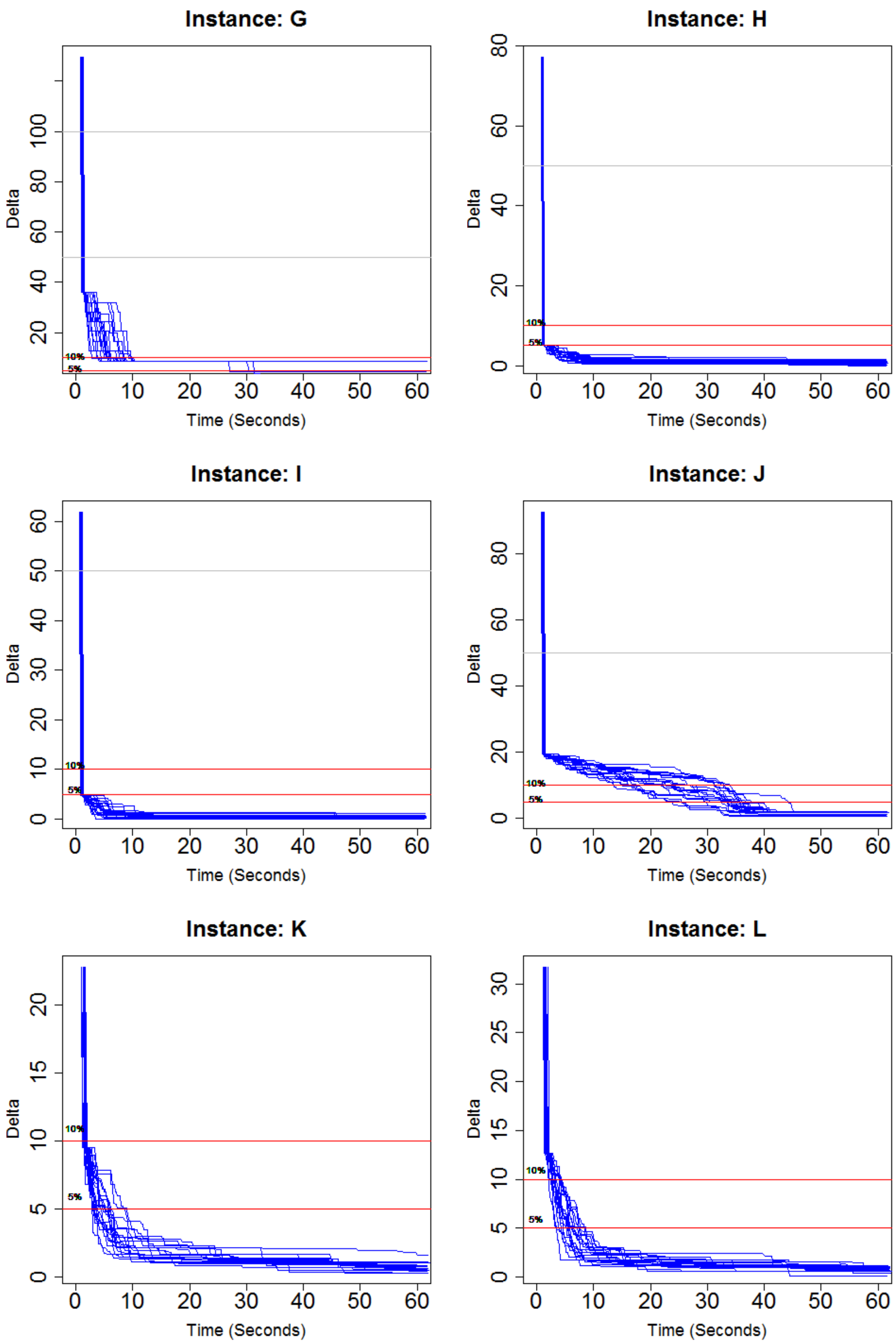
1.6. Repeat 1.3-1.5 until the remaining quantity for item $\hat{i}$ reaches zero or becomes negative. Then, remove $\hat{i}$ from $\mathcal{N}$.

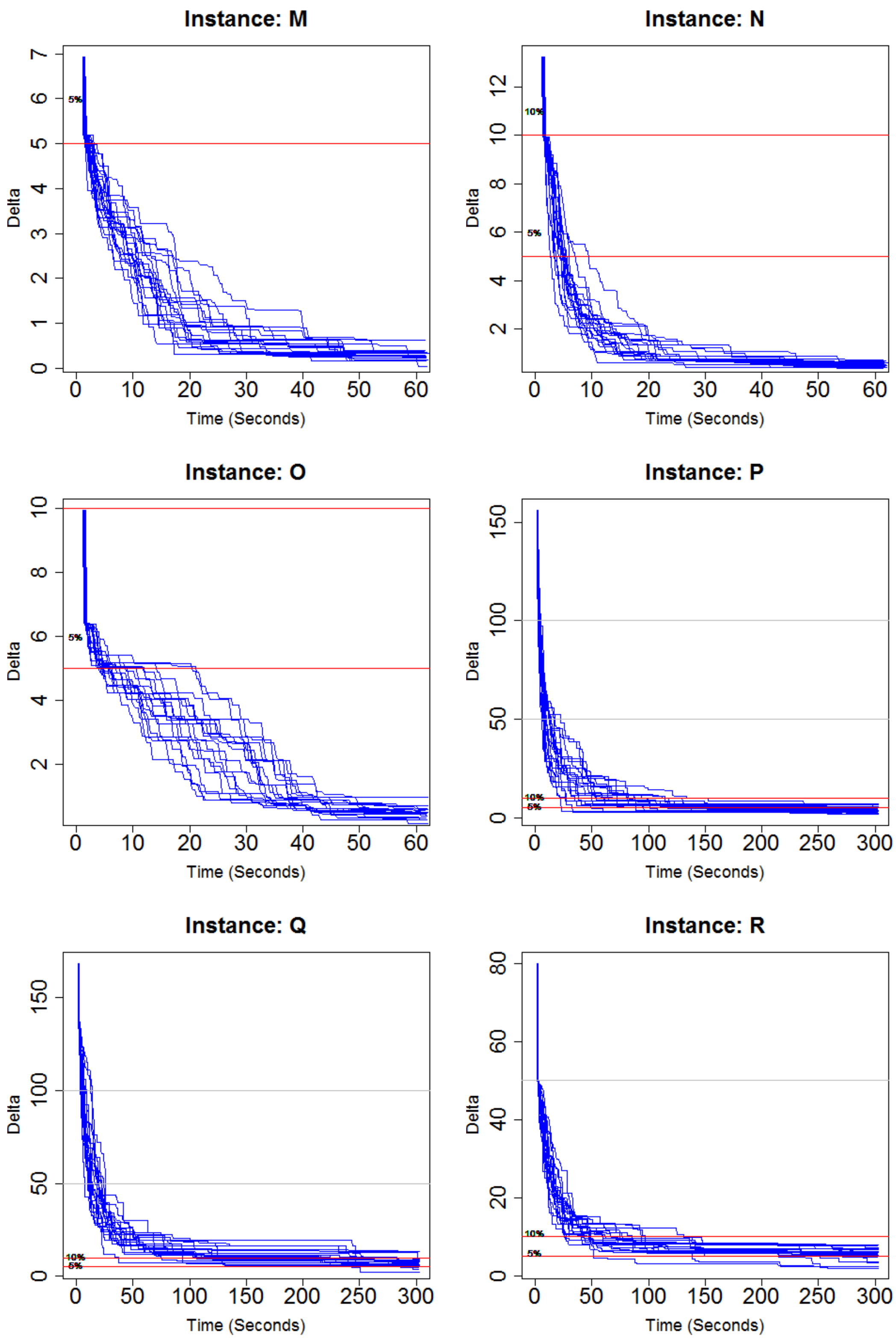
1.7. Repeat 1.2-1.6 until $\mathcal{N}$ is empty.

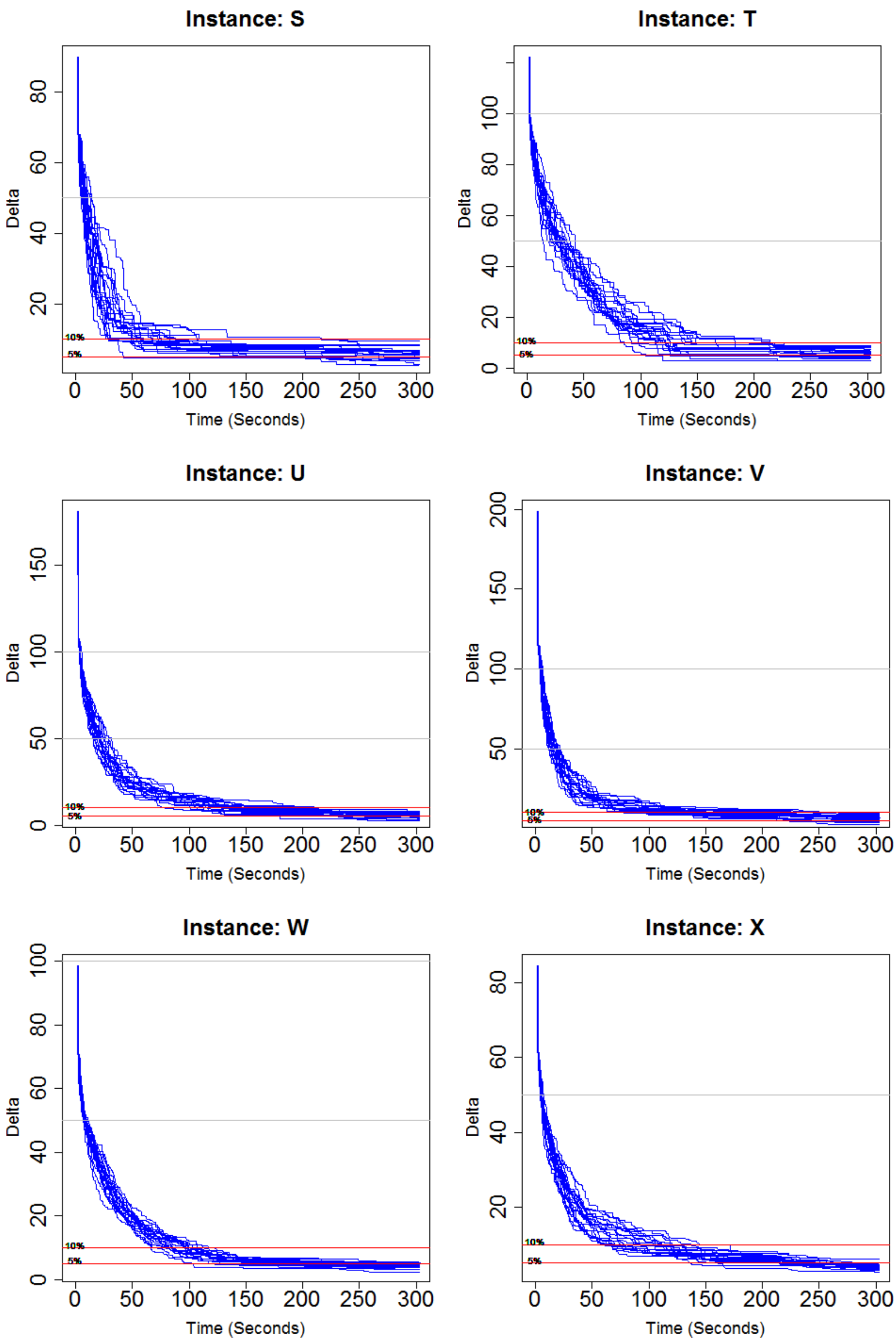
The complexity of GH1 depends on the forecast quantities and the prepack sizes. Let $n_i = \frac{\bar{F}_i}{\min(R_{im})}, \forall i$ and $\hat{n} = \max n_i$. Then, the maximum number of iterations required by GH1 is bounded by $\mathcal{O}(\hat{n}N)$. One drawback of the above methodology is that, once an item is ordered to fulfill its service level, it still has the chance to be ordered more while other retailers are being ordered since they may share common prepacks.

E. Additional Plots for Chapter 7




Instance: G
Instance: H
Instance: I
Instance: J
Instance: K
Instance: L
Delta
Time (Seconds)
10%
5%


Instance: M
Delta
0 1 2 3 4 5 6 7
5%
0 10 20 30 40 50 60
Time (Seconds)
Instance: N
Delta
2 4 6 8 10 12
10%
5%
0 10 20 30 40 50 60
Time (Seconds)
Instance: O
Delta
2 4 6 8 10
5%
0 10 20 30 40 50 60
Time (Seconds)
Instance: P
Delta
0 50 100 150
10%
5%
0 50 100 150 200 250 300
Time (Seconds)
Instance: Q
Delta
0 50 100 150
10%
5%
0 50 100 150 200 250 300
Time (Seconds)
Instance: R
Delta
0 20 40 60 80
10%
5%
0 50 100 150 200 250 300
Time (Seconds)


Instance: S
Instance: T
Instance: U
Instance: V
Instance: W
Instance: X
Delta
Time (Seconds)
10%
5%