INVENTORY CONTROL

Second Edition

Recent titles in the **INTERNATIONAL SERIES IN OPERATIONS RESEARCH & MANAGEMENT SCIENCE**
Frederick S. Hillier, Series Editor, *Stanford University*

Maros/ *COMPUTATIONAL TECHNIQUES OF THE SIMPLEX METHOD*
Harrison, Lee & Neale/ *THE PRACTICE OF SUPPLY CHAIN MANAGEMENT: Where Theory and Application Converge*
Shanthikumar, Yao & Zijm/ *STOCHASTIC MODELING AND OPTIMIZATION OF MANUFACTURING SYSTEMS AND SUPPLY CHAINS*
Nabrzyski, Schopf & Węglarz/ *GRID RESOURCE MANAGEMENT: State of the Art and Future Trends*
Thissen & Herder/ *CRITICAL INFRASTRUCTURES: State of the Art in Research and Application*
Carlsson, Fedrizzi, & Fullér/ *FUZZY LOGIC IN MANAGEMENT*
Soyer, Mazzuchi & Singpurwalla/ *MATHEMATICAL RELIABILITY: An Expository Perspective*
Chakravarty & Eliashberg/ *MANAGING BUSINESS INTERFACES: Marketing, Engineering, and Manufacturing Perspectives*
Talluri & van Ryzin/ *THE THEORY AND PRACTICE OF REVENUE MANAGEMENT*
Kavadias & Loch/*PROJECT SELECTION UNDER UNCERTAINTY: Dynamically Allocating Resources to Maximize Value*
Brandeau, Sainfort & Pierskalla/ *OPERATIONS RESEARCH AND HEALTH CARE: A Handbook of Methods and Applications*
Cooper, Seiford & Zhu/ *HANDBOOK OF DATA ENVELOPMENT ANALYSIS: Models and Methods*
Luenberger/ *LINEAR AND NONLINEAR PROGRAMMING, 2nd Ed.*
Sherbrooke/ *OPTIMAL INVENTORY MODELING OF SYSTEMS: Multi-Echelon Techniques, Second Edition*
Chu, Leung, Hui & Cheung/ *4th PARTY CYBER LOGISTICS FOR AIR CARGO*
Simchi-Levi, Wu & Shen/ *HANDBOOK OF QUANTITATIVE SUPPLY CHAIN ANALYSIS: Modeling in the E-Business Era*
Gass & Assad/ *AN ANNOTATED TIMELINE OF OPERATIONS RESEARCH: An Informal History*
Greenberg/ *TUTORIALS ON EMERGING METHODOLOGIES AND APPLICATIONS IN OPERATIONS RESEARCH*
Weber/ *UNCERTAINTY IN THE ELECTRIC POWER INDUSTRY: Methods and Models for Decision Support*
Figueira, Greco & Ehrgott/ *MULTIPLE CRITERIA DECISION ANALYSIS: State of the Art Surveys*
Reveliotis/ *REAL-TIME MANAGEMENT OF RESOURCE ALLOCATIONS SYSTEMS: A Discrete Event Systems Approach*
Kall & Mayer/ *STOCHASTIC LINEAR PROGRAMMING: Models, Theory, and Computation*
Sethi, Yan & Zhang/ *INVENTORY AND SUPPLY CHAIN MANAGEMENT WITH FORECAST UPDATES*
Cox/ *QUANTITATIVE HEALTH RISK ANALYSIS METHODS: Modeling the Human Health Impacts of Antibiotics Used in Food Animals*
Ching & Ng/ *MARKOV CHAINS: Models, Algorithms and Applications*
Li & Sun/ *NONLINEAR INTEGER PROGRAMMING*
Kaliszewski/ *SOFT COMPUTING FOR COMPLEX MULTIPLE CRITERIA DECISION MAKING*
Bouyssou et al/ *EVALUATION AND DECISION MODELS WITH MULTIPLE CRITERIA: Stepping stones for the analyst*
Blecker & Friedrich/ *MASS CUSTOMIZATION: Challenges and Solutions*
Appa, Pitsoulis & Williams/ *HANDBOOK ON MODELLING FOR DISCRETE OPTIMIZATION*
Herrmann/ *HANDBOOK OF PRODUCTION SCHEDULING*

**** A list of the early publications in the series is at the end of the book ****

INVENTORY CONTROL

Second Edition

Sven Axsäter

Springer

Sven Axsäter
Lund University
Lund, Sweden

Library of Congress Control Number: 2006922871

ISBN-10: 0-387-33250-2 (HB) ISBN-10: 0-387-33331-2 (e-book)
ISBN-13: 978-0387-33250-5 (HB) ISBN-13: 978-0387-33331-1 (e-book)

Printed on acid-free paper.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

springer.com

About the Author

Sven Axsäter is Professor of Production Management at Lund University since 1993. He is also head of the Department of Industrial Management and Logistics. Before coming to Lund he held professorships at Linköping Institute of Technology and Luleå University of Technology. He served as Visiting Professor at North Carolina State University in 1980 and at Hong Kong University of Science and Technology in 2001.

The main focus of Sven Axsäter's research has been production and inventory control. Past and current interests include: hierarchical production planning, lot sizing, and most recently multi-echelon inventory systems. He has published numerous papers in the leading journals in his research area, and has taught various courses on production and inventory control at universities in different parts of the world. Sven Axsäter has also served in an editorial capacity in various journals, including many years of service as Associate Editor of both *Operations Research* and *Management Science.*

Sven Axsäter has been President of the International Society of Inventory Research, and Vice President of the Production and Operations Management Society. He is also a member of the Royal Swedish Academy of Engineering Sciences. In 2005 he was awarded the Harold Larnder Memorial Prize by the Canadian Operational Research Society for distinguished international achievement in Operational Research.

In addition, he has a vast consulting experience in the inventory management area. He has implemented inventory control in several companies, and has also developed software for commercial inventory control systems.

CONTENTS

Preface xv

1 INTRODUCTION 1

1.1 Importance and objectives of inventory control 1
1.2 Overview and purpose of the book 2
1.3 Framework 5
References 5

2 FORECASTING 7

2.1 Objectives and approaches 7
2.2 Demand models 8
2.2.1 Constant model 9
2.2.2 Trend model 9
2.2.3 Trend-seasonal model 10
2.2.4 Choosing demand model 10
2.3 Moving average 11
2.4 Exponential smoothing 12
2.4.1 Updating procedure 12
2.4.2 Comparing exponential smoothing to a moving average 13
2.4.3 Practical considerations and an example 14
2.5 Exponential smoothing with trend 16
2.5.1 Updating procedure 16
2.5.2 Practical considerations and an example 17
2.6 Winters' trend-seasonal method 18
2.6.1 Updating procedure 18
2.6.2 Practical considerations and an example 20
2.7 Using regression analysis 21
2.7.1 Forecasting demand for a trend model 21
2.7.2 Practical considerations and an example 23
2.7.3 Forecasts based on other factors 24
2.7.4 More general regression models 25
2.8 Sporadic demand 26

2.9 Box-Jenkins techniques 27
2.10 Forecast errors 29
2.10.1 Common error measures 29
2.10.2 Updating MAD or σ^2 30
2.10.3 Determining the standard deviation as a function of demand 32
2.10.4 Forecast errors for other time periods 32
2.10.5 Sales data instead of demand data 34
2.11 Monitoring forecasts 34
2.11.1 Checking demand 35
2.11.2 Checking that the forecast represents the mean 35
2.12 Manual forecasts 36
References 37
Problems 38

3 COSTS AND CONCEPTS 43
3.1 Considered costs and other assumptions 44
3.1.1 Holding costs 44
3.1.2 Ordering or setup costs 44
3.1.3 Shortage costs or service constraints 45
3.1.4 Other costs and assumptions 45
3.2 Different ordering systems 46
3.2.1 Inventory position 46
3.2.2 Continuous or periodic review 47
3.2.3 Different ordering policies 48
3.2.3.1 (R, Q) policy 48
3.2.3.2 (s, S) policy 49
References 50

4 SINGLE-ECHELON SYSTEMS: DETERMINISTIC LOT SIZING 51
4.1 The classical economic order quantity model 52
4.1.1 Optimal order quantity 52
4.1.2 Sensitivity analysis 54
4.1.3 Reorder point 54
4.2 Finite production rate 55
4.3 Quantity discounts 56
4.4 Backorders allowed 59
4.5 Time-varying demand 61
4.6 The Wagner-Whitin algorithm 63

4.7 The Silver-Meal heuristic 66
4.8 A heuristic that balances holding and ordering costs 68
4.9 Exact or approximate solution 70
References 70
Problems 72

5 SINGLE-ECHELON SYSTEMS: REORDER POINTS 77
5.1 Discrete stochastic demand 77
5.1.1 Compound Poisson demand 77
5.1.2 Logarithmic compounding distribution 80
5.1.3 Geometric compounding distribution 82
5.1.4 Smooth demand 83
5.1.5 Fitting discrete demand distributions in practice 85
5.2 Continuous stochastic demand 85
5.2.1 Normally distributed demand 85
5.2.2 Gamma distributed demand 86
5.3 Continuous review (*R*, *Q*) policy - inventory level distribution 88
5.3.1 Distribution of the inventory position 88
5.3.2 An important relationship 90
5.3.3 Compound Poisson demand 90
5.3.4 Normally distributed demand 91
5.4 Service levels 94
5.5 Shortage costs 95
5.6 Determining the safety stock for given S_1 96
5.7 Fill rate and ready rate constraints 97
5.7.1 Compound Poisson demand 97
5.7.2 Normally distributed demand 98
5.8 Fill rate - a different approach 99
5.9 Shortage cost per unit and time unit 101
5.9.1 Compound Poisson demand 101
5.9.2 Normally distributed demand 103
5.10 Shortage cost per unit 106
5.11 Continuous review (*s*, *S*) policy 107
5.12 Periodic review - fill rate 109
5.12.1 Basic assumptions 110
5.12.2 Compound Poisson demand - (R, Q) policy 111
5.12.3 Compound Poisson demand - (s, S) policy 112
5.12.4 Normally distributed demand - (R, Q) policy 113
5.13 The newsboy model 114
5.14 A model with lost sales 117

5.15 Stochastic lead-times 119

5.15.1 Two types of stochastic lead-times 119

5.15.2 Handling sequential deliveries independent of the lead-time demand 120

5.15.3 Handling independent lead-times 122

5.15.4 Comparison of the two types of stochastic lead-times 123

References 124

Problems 126

6 SINGLE-ECHELON SYSTEMS: INTEGRATION - OPTIMALITY 129

6.1 Joint optimization of order quantity and reorder point 129

6.1.1 Discrete demand 129

6.1.1.1 (R, Q) policy 130

6.1.1.2 (s, S) policy 132

6.1.2 An iterative technique 133

6.1.3 Fill rate constraint - a simple approach 135

6.2 Optimality of ordering policies 137

6.2.1. Optimality of (R, Q) policies when ordering in batches 138

6.2.2 Optimality of (s, S) policies 140

6.3 Updating order quantities and reorder points in practice 140

References 145

Problems 146

7 COORDINATED ORDERING 149

7.1 Powers-of-two policies 150

7.2 Production smoothing 154

7.2.1 The Economic Lot Scheduling Problem (ELSP) 155

7.2.1.1 Problem formulation 155

7.2.1.2 The independent solution 156

7.2.1.3 Common cycle time 157

7.2.1.4 Bomberger's approach 159

7.2.1.5 A simple heuristic 160

7.2.1.6 Other problem formulations 163

7.2.2 Time-varying demand 163

7.2.2.1 A generalization of the classical dynamic lot size problem 163

7.2.2.2 Application of mathematical programming approaches 170

7.2.3 Production smoothing and batch quantities 170

7.3 Joint replenishments 172
7.3.1 A deterministic model 173
7.3.1.1 Approach 1. An iterative technique 174
7.3.1.2 Approach 2. Roundy's 98 percent approximation 176
7.3.2 A stochastic model 180
References 181
Problems 183

8 MULTI-ECHELON SYSTEMS: STRUCTURES AND ORDERING POLICIES 187

8.1 Inventory systems in distribution and production 188
8.1.1 Distribution inventory systems 188
8.1.2 Production inventory systems 189
8.1.3 Repairable items 192
8.1.4 Lateral transshipments in inventory systems 192
8.1.5 Inventory models with remanufacturing 194
8.2 Different ordering systems 195
8.2.1 Installation stock reorder point policies and KANBAN policies 196
8.2.2 Echelon stock reorder point policies 197
8.2.3 Comparison of installation stock and echelon stock policies 198
8.2.4 Material Requirements Planning 204
8.2.5 Ordering system dynamics 213
References 215
Problems 217

9 MULTI-ECHELON SYSTEMS: LOT SIZING 221

9.1 Identical order quantities 222
9.1.1 Infinite production rates 222
9.1.2 Finite production rates 223
9.2 Constant demand 225
9.2.1 A simple serial system with constant demand 225
9.2.2 Roundy's 98 percent approximation 230

9.3 Time-varying demand 236
9.3.1 Sequential lot sizing 236
9.3.2 Sequential lot sizing with modified parameters 238
9.3.3 Other approaches 240
9.3.4 Concluding remarks 241
References 242
Problems 243

10 MULTI-ECHELON SYSTEMS: REORDER POINTS 247

10.1 The Clark-Scarf model 248
10.1.1 Serial system 249
10.1.2 The Clark-Scarf approach for a distribution system 256
10.2 The METRIC approach for distribution systems 261
10.3 Two exact techniques 266
10.3.1 Disaggregation of warehouse backorders 266
10.3.2 A recursive procedure 267
10.4 Optimization of ordering policies 271
10.5 Batch-ordering policies 273
10.5.1 Serial system 273
10.5.2 Distribution system 276
10.5.2.1 Some basic results 277
10.5.2.2 METRIC type approximations 278
10.5.2.3 Disaggregation of warehouse backorders 279
10.5.2.4 Following supply units through the system 280
10.5.2.5 Practical considerations 280
10.6 Other assumptions 281
10.6.1 Guaranteed service model approach 281
10.6.2 Coordination and contracts 283
10.6.2.1 The newsboy problem with two firms 284
10.6.2.2 Wholesale-price contract 285
10.6.2.3 Buyback contract 286
References 287
Problems 291

11 IMPLEMENTATION 295

11.1 Preconditions for inventory control 295
11.1.1 Inventory records 296
11.1.1.1 Updating inventory records 296
11.1.1.2 Auditing and correcting inventory records 297
11.1.2 Performance evaluation 298

11.2 Development and adjustments 299
11.2.1 Determine the needs 300
11.2.2 Selective inventory control 301
11.2.3 Model and reality 302
11.2.4 Step-by-step implementation 303
11.2.5 Simulation 304
11.2.6 Short-run consequences of adjustments 305
11.2.7 Education 306
References 307

APPENDIX 1
ANSWERS AND HINTS TO PROBLEMS 309

APPENDIX 2
NORMAL DISTRIBUTION TABLES 321

INDEX 325

Preface

Modern information technology has created new possibilities for more sophisticated and efficient control of supply chains. Most organizations can reduce their costs associated with the flow of materials substantially. Inventory control techniques are very important components in this development process. A thorough understanding of relevant inventory models is a prerequisite for successful implementation. I hope that this book will be a useful tool in acquiring such an understanding.

The book is primarily intended as a course textbook. It assumes that the reader has a good basic knowledge of mathematics and probability theory, and is therefore most suitable for industrial engineering and management science/operations research students. The book can be used both in undergraduate and more advanced graduate courses.

About fifteen years ago I wrote a Swedish book on inventory control. This book is still used in courses in production and inventory control at several Swedish engineering schools and has also been appreciated by many practitioners in the field. Positive reactions from many readers made me contemplate writing a new book in English on the same subject. Encouraging support of this idea from the Springer Editors Fred Hillier and Gary Folven finally convinced me to go ahead with that project six years ago.

The resulting first edition of this book was published in 2000 and contained quite a lot of new material that was not included in its Swedish predecessor. It has since then been used in quite a few university courses in different parts of the world, and I have received many positive reactions. Still some readers have felt that the book was too compact and some have asked for additional topics. Some of those who have used the book as a textbook have also requested more problems to be solved by the students.

The Springer Editors Fred Hillier and Gary Folven finally convinced me to publish a Second Edition of my book. This new edition is quite different from the previous one. The text has been expanded by more than 50 percent. My main goal has been to make the new book more suitable as a textbook. There are eleven chapters compared to six in the previous version. The explanations of different results are more detailed, and a considerable number of exercises have been added. I have also included several new topics. The additions include: alternative forecasting techniques, more material on different stochastic demand processes and how they can be fitted to empirical data, generalized treatment of single-echelon periodic review systems, capacity constrained lot sizing, short sections on lateral transshipments and on remanufacturing, coordination and contracts.

When working with the book I have been much influenced by other textbooks and various scientific articles. I would like to thank the authors of these books and papers for indirectly contributing to my book.

There are also a number of individuals that I would like to thank. Before I started to work on the revision, Springer helped me to arrange a review process, where a number of international scholars were asked to suggest suitable changes in the book. These scholars were: Shoshana Anily, Tel Aviv University, Saif Benjaafar, University of Minnesota, Eric Johnson, Dartmouth College, George Liberopoulos, University of Thessaly, Suresh Sethi, University of Texas-Dallas, Jay Swaminathan, University of North Carolina, Ruud Teunter, Lancaster University, Geert-Jan Van Houtum, Eindhoven University of Technology, Luk Van Wassenhove, INSEAD, and Yunzeng Wang, Case Western Reserve University. Some of them had used the first edition of the book in their classes. The review process resulted in most valuable suggestions for improvements, and I want to thank all of you very much.

Several colleagues of mine at Lund University have helped me a lot. I would especially like to mention Johan Marklund for much and extremely valuable help with both editions, and Kaj Rosling (now at Växjö University) for his important suggestions concerning the first edition. Furthermore, Jonas Andersson, Peter Berling, Fredrik Olsson, Patrik Tydesjö, and Stefan Vidgren have reviewed the manuscript at different stages and offered valuable suggestions which have improved this book considerably. Thank you so much.

Finally, I would also like to thank the Springer people: Fred Hillier, Gary Folven, and Carolyn Ford for their support, and Sharon Bowker for polishing my English.

Sven Axsäter

1 INTRODUCTION

1.1 Importance and objectives of inventory control

For more or less all organizations in any sector of the economy, *Supply Chain Management*, i.e., the control of the material flow from suppliers of raw material to final customers, is a crucial problem. The strategic importance of this area is today fully recognized by top management. The total investment in inventories is enormous, and the control of capital tied up in raw material, work-in-progress, and finished goods offers a very important potential for improvement. Scientific methods for inventory control can give a significant competitive advantage. This book deals with a wide range of different inventory models that can be used when developing inventory control systems.

Advances in information technology have drastically changed the possibilities to apply efficient inventory control techniques. Furthermore, the recent progress in research has resulted in new and more general methods that can reduce the supply chain costs substantially. The field of inventory control has indeed changed during the last decades. It used to mean application of simple decision rules, which essentially could be carried out manually. Modern inventory control is based on quite advanced and complex decision models, which may require considerable computational efforts.

Inventories cannot be decoupled from other functions, for example purchasing, production, and marketing. As a matter of fact, the objective of inventory control is often to balance conflicting goals. One goal is, of course, to keep stock levels down to make cash available for other purposes. The purchasing manager may wish to order large batches to get volume discounts. The production manager similarly wants long production runs to avoid time-consuming setups. He also prefers to have a large raw material inventory to avoid stops in production due to missing materials. The marketing manager would like to have a high stock of finished goods to be able to provide customers a high service level.

It is seldom trivial to find the best balance between such goals, and that is why we need inventory models. In most situations some stocks are required. The two main reasons are *economies of scale* and *uncertainties*. Economies of scale mean that we need to order in *batches*. Uncertainties in supply and demand together with lead-times in production and transportation inevitably create a need for *safety stocks*. Still, most organizations can reduce their inventories without increasing other costs by using more efficient inventory control tools.

There are important inventory control problems in all supply chains. For those who are working with logistics and supply chains, it is difficult to think of any qualification that is more essential than a thorough understanding of basic inventory models.

1.2 Overview and purpose of the book

The main purpose of this book is that it should be useful as a course textbook. The structure of the book is illustrated in Figure 1.1.

After this introduction we consider different *forecasting techniques* in Chapter 2. We focus on methods like exponential smoothing and moving average procedures for estimating the future demand from historical demand data. We also provide techniques for evaluating the size of forecast errors.

Chapters 3 - 6 deal with basic inventory problems for a *single installation* and *items that can be handled independently*. More precisely, Chapter 3 presents various basic concepts. Chapter 4 deals with *deterministic lot sizing* and Chapter 5 with *safety stocks* and *reorder points*. In Chapter 6 we discuss *integration* and *optimality*.

The contents in Chapters 2 - 6 provide the foundation for an efficient standard inventory control system, which can include:

- A forecasting module, which periodically updates demand forecasts and evaluates forecast errors.
- A module for determination of reorder points and order quantities.
- Continuous or periodic monitoring of inventory levels and outstanding orders. Triggering of suggested orders when reaching the reorder points.

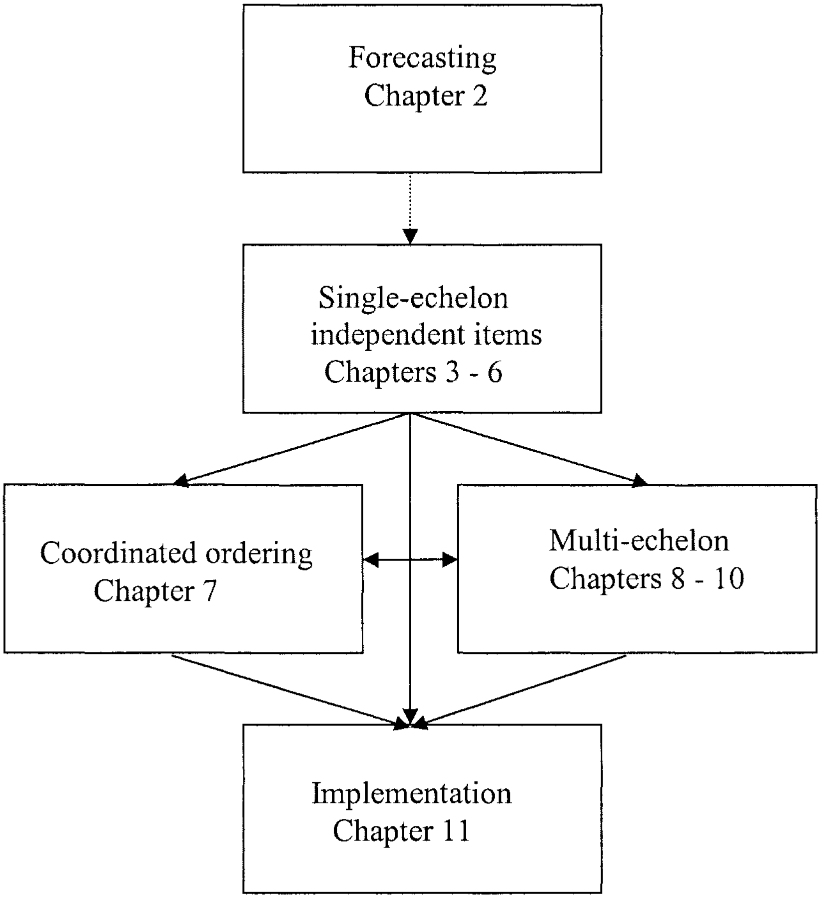

Figure 1.1 Structure of the book.

In Chapter 7 we leave the assumption of independent items and consider *coordinated replenishments*. Both production smoothing models and so-called joint replenishment problems are analyzed.

Chapters 8 - 10 focus on multi-echelon inventory systems, i.e., on several installations which are coupled to each other. The installations can represent, for example, stocks of raw materials, components, work-in-process, and final products in a production system, or a central warehouse and a number of retailers in a distribution system. In Chapter 8 we consider *structures* and *ordering policies*. Chapter 9 deals with *lot sizing* and Chapter 10 with *safety stocks* and *reorder points*.

Finally, in Chapter 11 we discuss various practical problems in connection with *implementation* of inventory control systems.

Over the years a substantial number of excellent books and overview papers dealing with various inventory control topics have been published. A selection of these publications is listed at the end of this chapter. A natural question then is why this book is needed. To explain this, note first that this book is different from most other books because it also covers very recent advances in inventory theory, for example new techniques for multi-echelon inventory systems and Roundy's 98 percent approximation. Furthermore, this book is also different from most other books because it assumes a reader with a good basic knowledge of mathematics and probability theory. This makes it possible to present different inventory models in a compact and hopefully more efficient way. The book attempts to explain fundamental ideas in inventory modeling in a simple but still rigorous way. However, to simplify, several models are less general than they could have been.

Because the book assumes a good basic knowledge of mathematics and probability theory, it is most suitable for industrial engineering and management science/operations research students. It can be used in a basic undergraduate course, and/or in a more advanced graduate course.

Chapter 2 may be omitted in a course which is strictly focused on inventory control. If it is included, it should probably be the first part of the course. Chapters 3 - 6 should precede Chapters 7 - 10. Chapter 7 can either precede or succeed Chapters 8 - 10. Chapter 11 should come at the end of the course.

An *undergraduate course* can, for example, be based on the following parts of the book: Sections 2.1 - 2.6, Sections 2.10 - 2.12, Chapters 3 - 4, Section 5.1.1, Section 5.2.1, Sections 5.3 - 5.8, Section 5.13, Section 6.3, Section 7.2.1, Section 8.1, Sections 8.2.1 - 8.2.2, Sections 8.2.4 - 8.2.5, Section 9.1, Section 9.2.1, Chapter 11.

For students that have taken the suggested undergraduate course, or a corresponding course, a *graduate course* can build on a selection of the remaining parts of the book, e.g., Sections 5.1.2 - 5.1.5, Section 5.2.2. Sections 5.9 - 5.12, Sections 5.14 - 5.15, Sections 6.1 - 6.2, Section 7.1, Section 7.3, Section 8.2.3, Sections 9.2 - 9.3, Chapter 10.

A graduate course for students that have no prior knowledge of inventory control but a good mathematical background should include most of the material suggested for the undergraduate course, but can exclude some of the sections suggested for the graduate course.

Another purpose of this book is to describe and explain efficient inventory control techniques for practitioners, and in that way simplify and promote implementation in practice. The book can, e.g., be used as a *handbook* when implementing and adjusting inventory control systems.

1.3 Framework

Models and methods in this book are based on the cost structure that is most common in industrial applications. We consider holding costs including opportunity costs of alternative investments, ordering or setup costs, and shortage costs or service level constraints. We will not deal with, for example, inventory problems related to financial speculation, i.e., when the value of an item can be expected to increase, or with aggregate planning models for smoothing production in case of seasonal demand variations. The interaction with production is recognized through setup costs but also in some models by explicit capacity constraints. The book does not cover production planning settings that are not directly related to inventory control.

The models considered in the book assume that the basic conditions for inventory control are given, for example in the form of demand distributions, lead-times, service requirements, and holding and ordering costs. In practice, most of these conditions can be changed at least in the long run. There are, consequently, many important questions concerning inventories that are related to the structure and organization of the inventory control system. Such questions may concern evaluation of investments to reduce setup costs, or whether the customers should be served through a single-stage or a multi-stage inventory system. Although we do not treat such questions directly, it is important to note that a correct evaluation must always be based on inventory models of the type considered in this book. The question is always whether the savings in inventory-related costs are larger than the costs for changing the structure of the system.

References

Brown, R. G. 1967. *Decision Rules for Inventory Management*, Holt, Rinehart and Winston, New York.

Chikán, A. Ed. 1990. *Inventory Models*, Kluwer Academic Publishers, Boston.

De Kok, A. G., and S. C. Graves. Eds. 2003. *Supply Chain Management: Design, Coordination and Operation*, Handbooks in OR & MS, Vol.11, North Holland, Amsterdam.

Graves, S. C., A. Rinnooy Kan, and P. H. Zipkin. Eds. 1993. *Logistics of Production and Inventory*, Handbooks in OR & MS, Vol.4, North Holland, Amsterdam.

Hadley, G., and T. M. Whitin. 1963. *Analysis of Inventory Systems*, Prentice-Hall, Englewood Cliffs, NJ.

Hax, A., and D. Candea. 1984. *Production and Inventory Management*, Prentice-Hall, Englewood Cliffs, NJ.

Johnson, L. A., and D. C. Montgomery. 1974. *Operations Research in Production Planning, Scheduling, and Inventory Control*, Wiley, New York.

Love, S. F. 1979. *Inventory Control*, McGraw-Hill, New York.
McClain, J. O., and L. J. Thomas. 1980. *Operations Management: Production of Goods and Services*, Prentice-Hall, Englewood Cliffs, NJ.
Muller, M. 2003. *Essentials of Inventory Management*, AMACOM, New York.
Naddor, E. 1966. *Inventory Systems*, Wiley, New York.
Nahmias, S. 1997. *Production and Operations Analysis*, 3rd edition, Irwin, Boston.
Orlicky, J. 1975. *Material Requirements Planning*, McGraw-Hill, New York.
Plossl, G. W., and O. W. Wight. 1985. *Production and Inventory Control*, Prentice-Hall, Englewood Cliffs, NJ.
Porteus, E. L. 2002. *Foundations of Stochastic Inventory Theory*, Stanford University Press.
Sherbrooke, C. C. 2004. *Optimal Inventory Modeling of Systems*, 2nd edition, Kluwer Academic Publishers, Boston.
Silver, E. A., D. F. Pyke, and R. Peterson. 1998. *Inventory Management and Production Planning and Scheduling*, 3rd edition, Wiley, New York.
Tersine, R. J. 1988. *Principles of Inventory and Materials Management*, 3rd edition, North-Holland, New York.
Veinott, A. 1966. The Status of Mathematical Inventory Theory, *Management Science*, 12, 745-777.
Vollman, T. E., W. L. Berry, and D. C. Whybark. 1997. *Manufacturing Planning and Control Systems*, 4th edition, Irwin, Boston.
Wagner, H. M. 1962. *Statistical Management of Inventory Systems*, Wiley, New York.
Zipkin, P. H. 2000. *Foundations of Inventory Management*, McGraw-Hill, Singapore.

2 FORECASTING

There are two main reasons why an inventory control system needs to order items some time before customers demand them. First, there is nearly always a *lead-time* between the ordering time and the delivery time. Second, due to certain ordering costs, it is often necessary to order in *batches* instead of unit for unit. This means that we need to look ahead and forecast the future demand. A demand forecast is an estimated average of the demand size over some future period. But it is not enough to estimate the average demand. We also need to determine how uncertain the forecast is. If the forecast is more uncertain, a larger safety stock is required. Consequently, it is also necessary to estimate the forecast error, which may be represented by the standard deviation or the Mean Absolute Deviation (*MAD*).

2.1 Objectives and approaches

In this chapter we shall consider forecasting methods that are suitable in connection with inventory control. Typical for such forecasts is that they concern a relatively short time horizon. Very seldom is it necessary to look more than one year ahead. In general, there are then two types of approaches that may be of interest:

- Extrapolation of historical data

When extrapolating historical data, the forecast is based on previous demand data. The available techniques are grounded in statistical methods for analysis of time series. Such techniques are easy to apply and use in computerized inventory control systems. It is no problem to regularly update forecasts for thousands of items, which is a common requirement in connection with practical inventory control. Extrapolation of historical data is the most common and important approach to obtain forecasts over a short horizon, and we shall devote the main part of this chapter to such techniques.

- Forecasts based on other factors

It is very common that the demand for an item depends on the demand for some other items. Consider, for example, an item that is used exclusively as a component when assembling some final products. It is then often natural to first forecast the demand for these final products, for example by extrapolation of historical data. Next we determine a production plan for the products. The demand for the considered component is then obtained directly from the production plan. This technique to "forecast" demand for dependent items is used in *Material Requirements Planning* (*MRP*) that is dealt with in Section 8.2.4.

But there are also other factors that might be reasonable to consider when forecasting demand. Assume, for example, that a sales campaign is just about to start or that a competing product is introduced on the market. Clearly this can mean that historical data are no longer representative when looking ahead. It is normally difficult to take such factors into account in the forecasting module of a computerized inventory control system. It is therefore usually most practical to adjust the forecast manually in case of such special events.

It is also possible, at least in principle, to use other types of dependencies. A forecast for the demand of ice cream can be based on the weather forecast. Consider, as another example, forecasting of the demand for a spare part that is used as a component in certain machines. The demand for the spare part can be expected to increase when the machines containing the part as a component are getting old. It is therefore reasonable to look for dependencies between the demand for the spare part and previous sales of the machines. As another example we can assume that the demand during a certain month will increase with the advertising expenditure the previous month. Such dependencies could be determined from historical data by regression analysis. (See Section 2.7.) Applications of such techniques are, however, very limited.

2.2 Demand models

Extrapolation of historical data is, as mentioned, the most common approach when forecasting demand in connection with inventory control. To determine a suitable technique, we need to have some idea of how to model the stochastic demand. In principle, we should try to determine the model from analysis of historical data. In practice this is very seldom done. With many thousands of items, this initial work does not seem to be worth the effort in many situations. In other situations there are not enough historical data. A model for the demand structure is instead determined intuitively. In general, the assumptions are very simple.

2.2.1 Constant model

The simplest possible model means that the demands in different periods are represented by independent random deviations from an average that is assumed to be relatively stable over time compared to the random deviations. Let us introduce the notation:

x_t = demand in period t,
a = average demand per period (assumed to vary slowly),
ε_t = independent random deviation with mean zero.

A constant model means that we assume that the demand in period t can be represented as

$$x_t = a + \varepsilon_t . \tag{2.1}$$

Many products can be represented well by a constant model, especially products that are in a mature stage of a product life cycle and are used regularly. Examples are consumer products like toothpaste, many standard tools, and various spare parts. In fact, if we do not expect a trend or a seasonal pattern, it is in most cases reasonable to assume a constant model.

2.2.2 Trend model

If the demand can be assumed to increase or decrease systematically, it is possible to extend the model by also considering a linear trend. Let

a = average demand in period 0,
b = trend, that is the systematic increase or decrease per period (assumed to vary slowly).

A trend model means that the demand is modeled as:

$$x_t = a + bt + \varepsilon_t. \tag{2.2}$$

During a product life cycle there is an initial growth stage and a phase-out stage at the end of the cycle. During these stages it is natural to assume that the demand follows a trend model with a positive trend in the growth stage and a negative trend in the phase-out stage.

2.2.3 Trend-seasonal model

Let

F_t = seasonal index in period t (assumed to vary slowly).

If, for example, $F_t = 1.2$, this means that the demand in period t is expected to be 20 percent higher due to seasonal variations. If there are T periods in one year, we must require that for any T consecutive periods $\sum_{k=1}^{T} F_{t+k} = T$. When using a multiplicative trend-seasonal demand model it is assumed that the demand can be expressed as

$$x_t = (a + bt)F_t + \varepsilon_t. \tag{2.3}$$

By setting $b = 0$ in (2.3) we obtain a constant-seasonal model.

In (2.3) it is assumed that the seasonal variations increase and decrease proportionally with increases and decreases in the level of the demand series. In most cases this is a reasonable assumption. An alternative assumption could be that the seasonal variations are additive.

Many products have seasonal demand variations. For example the demand for ice cream is much larger during the summer than in the winter. Some products, like various Christmas decorations, are only sold during a very short period of the year. Still, the number of items with seasonal demand variations is usually very small compared to the total number of items. A seasonal model is only meaningful if the demand follows essentially the same pattern year after year.

2.2.4 Choosing demand model

When looking at the three demand models considered, it is obvious that (2.2) is more general than (2.1), and that (2.3) is more general than (2.2). It may then appear that it should be most advantageous to use the most general model (2.3). This is, however, not true. A more general demand model covers a wider class of demands, but on the other hand, we need to estimate more parameters. Especially if the independent deviations are large, it may be very difficult to determine accurate estimates of the parameters, and it can therefore be much more efficient to use a simple demand model with few parameters. A more general model should be avoided unless there is some evidence that the generality will give certain advantages.

It is important to understand that the independent deviations ε_t cannot be forecasted, or in other words, the best forecast for ε_t is always zero. Consequently, if the independent deviations are large there is no possibility to avoid large forecast errors. Consider the constant model (2.1). It is obvious that the best forecast is simply our best estimate of a. In (2.2) the best forecast for the demand in period t is similarly our best estimate of $a + bt$, and in (2.3) our best forecast is the estimate of $(a + bt)F_t$.

In some situations it may be interesting to use more general demand models than (2.1) - (2.3). (See Section 2.9.) This would, however, require a detailed statistical analysis of the demand structure. In practice this is rarely done in connection with inventory control.

One practical problem is that it is quite often difficult to measure demand, since only sales are recorded. If historical sales, instead of historical demands, are used for forecasting demand, considerable errors may occur in situations where a relatively large portion of the total demand is lost due to shortages. (See Section 2.10.5.)

2.3 Moving average

Assume that the underlying demand structure is described by the constant model (2.1). Since the independent deviations ε_t cannot be predicted, we simply want to estimate the constant a. If a were completely constant the best estimate would be to take the average of all observations of x_t. But a can be expected to vary slowly. This means that we need to focus on the most recent values of x_t. The idea of the moving average technique is to take the average over the N most recent values. Let

$\hat{a}_t$ = estimate of a after observing the demand in period t,

$\hat{x}_{t,\tau}$ = forecast for period $\tau > t$ after observing the demand in period t.

We obtain:

$$\hat{x}_{t,\tau} = \hat{a}_t = (x_t + x_{t-1} + x_{t-2} + \ldots + x_{t-N+1}) / N . \tag{2.4}$$

Note that the forecast demand is the same for any value of $\tau > t$. This is, of course, because we are assuming a constant demand model.

The value of N should depend on how slowly we think that a is varying, and on the size of the stochastic deviations ε_t. If a is varying more slowly and the stochastic deviations are larger, we should use a larger value of N. This will limit the influence of the stochastic deviations. On the other hand, if a is varying more rapidly and the stochastic variations are small, we should prefer a small value of N, which will allow us to follow the variations in a in a better way.

If we use one month as our period length and set N = 12, the forecast is the average over the preceding year. This may be an advantage if we want to prevent seasonal variations from affecting the forecast.

2.4 Exponential smoothing

2.4.1 Updating procedure

When using exponential smoothing instead of a moving average, the forecast is updated differently. The result is in many ways similar, though. We are again assuming a constant demand model, and we wish to estimate the parameter a. To update the forecast in period t, we use a linear combination of the previous forecast and the most recent demand x_t,

$$\hat{x}_{t,\tau} = \hat{a}_t = (1-\alpha)\hat{a}_{t-1} + \alpha x_t, \tag{2.5}$$

where $\tau > t$ and

α = smoothing constant ($0 < \alpha < 1$).

Due to the constant demand model the forecast is again the same for any future period.

Note that the updating procedure can also be expressed as

$$\hat{x}_{t,\tau} = \hat{a}_t = \hat{a}_{t-1} + \alpha(x_t - \hat{a}_{t-1}). \tag{2.6}$$

We have assumed that $0 < \alpha < 1$ although it is also possible to use $\alpha = 0$ and $\alpha = 1$. The value $\alpha = 0$ means simply that we do not update the forecast, while $\alpha = 1$ means that we choose the most recent demand as our forecast.

2.4.2 Comparing exponential smoothing to a moving average

To be able to compare exponential smoothing to a moving average, we can express the forecast in the following way:

$$\hat{a}_t = (1-\alpha)\hat{a}_{t-1} + \alpha x_t = (1-\alpha)((1-\alpha)\hat{a}_{t-2} + \alpha x_{t-1}) + \alpha x_t$$

$$= \alpha x_t + \alpha(1-\alpha)x_{t-1} + (1-\alpha)^2 \hat{a}_{t-2} = ... = \alpha x_t + \alpha(1-\alpha)x_{t-1}$$

$$+ \alpha(1-\alpha)^2 x_{t-2} + ... + \alpha(1-\alpha)^n x_{t-n} + (1-\alpha)^{n+1} \hat{a}_{t-n-1}. \quad (2.7)$$

Let us now compare (2.7) to (2.4). In (2.4) the N last period demands all have the weight $1/N$. In (2.7) we have, in principle, positive weights for all previous demands, but the weights are decreasing exponentially as we go backwards in time. This is the reason for the name exponential smoothing. The sum of the weights is still unity.[1] When using a moving average, a larger value of N means that we put relatively more emphasis on old values of demand. When applying exponential smoothing, a small value of α will give essentially the same effect.

When using a moving average according to (2.4) the forecast is based on the demands in periods t, t - 1, ..., t - N + 1. The ages of these data are respectively 0, 1, ..., and N - 1 periods. The weights are all equal to $1/N$. The average age is therefore (N - 1)/2 periods. To be able to compare the parameter N to the smoothing constant α, we shall also determine the average age of the data when using exponential smoothing according to (2.5), or equivalently (2.7). We obtain:[2]

$$\alpha 0 + \alpha(1-\alpha)1 + \alpha(1-\alpha)^2 2 + ... = \alpha(1-\alpha)S'(1-\alpha) = (1-\alpha)/\alpha, \quad (2.8)$$

---

[1] Let $0 \le x < 1$ and consider the infinite geometric sum $S(x) = 1 + x + x^2 + x^3 ...$ Note that $S(x) = 1 + x \cdot S(x)$. This implies that $S(x) = 1/(1-x)$. The sum of the weights in (2.7) is $\alpha \cdot S(1-\alpha) = 1$.

[2] Let $0 \le x < 1$ and consider the infinite geometric sum $S'(x) = 1 + 2x + 3x^2 .. = 1 + x + x^2 + ... + x(1 + 2x + 3x^2 ...) = S(x) + x \cdot S'(x)$. This implies that $S'(x) = S(x)/(1-x) = 1/(1-x)^2$.

and we can conclude that the forecasts are based on data of the "same average age" if

$$(1-\alpha)/\alpha=(N-1)/2, \tag{2.9}$$

or equivalently when

$$\alpha=2/(N+1). \tag{2.10}$$

Consider, for example, a moving average that is updated monthly with $N = 12$. This means that each month in the preceding year has weight 1/12. Consider then an exponential smoothing forecast that is also updated monthly. A value of α "corresponding" to $N = 12$ is according to (2.10) obtained as $\alpha = 2/(12 + 1) = 2/13 \approx 0.15$.

2.4.3 *Practical considerations and an example*

If the period length is one month, it is common in practice to use a smoothing constant α between 0.1 and 0.3. Table 2.1 shows the weights for different previous demands for $\alpha = 0.1$ and $\alpha = 0.3$.

Table 2.1 Weights for demand data in exponential smoothing

Period	Weight	$\alpha = 0.1$	$\alpha = 0.3$
t	α	0.100	0.300
t-1	$\alpha(1 - \alpha)$	0.090	0.210
t-2	$\alpha(1 - \alpha)^2$	0.081	0.147
t-3	$\alpha(1 - \alpha)^3$	0.073	0.103
t-4	$\alpha(1 - \alpha)^4$	0.066	0.072

We can see from Table 2.1 that the forecasting system will react much faster if we use $\alpha = 0.3$. On the other hand, the stochastic deviations will affect the forecast more compared to when $\alpha = 0.1$. When choosing α we always have to compromise.

If the forecast is updated more often, for example each week, a smaller α should be used. To see how much smaller we can apply (2.10). Assume that we start with a monthly update and that we use the value of α "corresponding" to a moving average with $N = 12$, i.e., $\alpha \approx 0.15$. When changing to weekly forecasts it is natural to change N to 52. The "corresponding" value of α is obtained from (2.10) as $\alpha = 2/(52 + 1) \approx 0.04$.

When starting to forecast according to (2.5) in some period t, an initial forecast to be used as $\hat{a}_{t-1}$ is needed. We can use some simple estimate of the average period demand. If no such estimate is available, it is possible to start with $\hat{a}_{t-1} = 0$, since $\hat{a}_{t-1}$ will not affect the forecast in the long run, see (2.7). However, especially for small values of α, it can take a long time until the forecasts are reliable. If it is necessary to start with a very uncertain initial forecast, it may be a good idea to use a rather large value of α to begin with, since this will reduce the influence of the initial forecast.

Example 2.1 The demand for an item usually fluctuates considerably. A moving average or a forecast obtained by exponential smoothing gives essentially an average of more recent demands. The forecast cannot, as we have emphasized before, predict the independent stochastic deviations. Table 2.2 shows some typical demand data and the corresponding exponential smoothing forecasts with $\alpha = 0.2$. It is assumed that the forecast after period 2 is $\hat{a}_2 = 100$.

Table 2.2 Forecasts obtained by exponential smoothing with α =0.2. Initial forecast $\hat{a}_2$ = 100.

Period	Demand in period t, x_t	Forecast at the end of period t, $\hat{a}_t$
3	72	94
4	170	110
5	67	101
6	95	100
7	130	106

In Table 2.2 the forecast immediately after period 3 is obtained by applying (2.5).

$$\hat{a}_3 = 0.8 \cdot 100 + 0.2 \cdot 72 = 94.4,$$

which is rounded to 94 in Table 2.2. Note that when determining $\hat{a}_3$ the demands in future periods are not known. Therefore at this stage, $\hat{a}_3$ serves as our forecast for any future period. After period 4 the forecast is again updated

$$\hat{a}_4 = 0.8 \cdot 94.4 + 0.2 \cdot 170 = 109.52.$$

If we compare exponential smoothing to a moving average there are some obvious but minor advantages with exponential smoothing. Because the average a, which we wish to estimate is assumed to vary slowly, it is reasonable to use larger weights for the most recent demands as is done in exponential smoothing. As we have discussed before, however, a moving average over a full year may be advantageous if we want to eliminate the influence of seasonal variations on the forecast. It is also interesting to note that with exponential smoothing we only need to keep track of the previous forecast and the most recent demand.

In practice, exponential smoothing (or possibly a moving average) is, in general, a suitable technique for most items. But there is also usually a need for other methods for relatively small groups of items for which it is feasible and interesting to follow up trends and/or seasonal variations.

2.5 Exponential smoothing with trend

2.5.1 Updating procedure

Let us now instead assume that the demand follows a trend model according to (2.2). To forecast demand we need to estimate the two parameters a and b, compared to only a in case of a constant model. As before, we cannot predict the independent deviations ε_t. There are different techniques for estimating a and b. We shall here consider a method that was first suggested by Holt (1957). (Another technique based on linear regression is described in Section 2.7.) Estimates of a and b are successively updated according to (2.11) and (2.12).

$$\hat{a}_t = (1-\alpha)(\hat{a}_{t-1} + \hat{b}_{t-1}) + \alpha x_t, \tag{2.11}$$

$$\hat{b}_t = (1-\beta)\hat{b}_{t-1} + \beta(\hat{a}_t - \hat{a}_{t-1}), \tag{2.12}$$

where α and β are smoothing constants between 0 and 1.

The "average" $\hat{a}_t$ corresponds to period t, i.e., the period for which we have just observed the demand. The forecast for a future period, $t + k$ is obtained as

$$\hat{x}_{t,t+k} = \hat{a}_t + k \cdot \hat{b}_t. \tag{2.13}$$