Roma Tre University
Ph.D in Operations Research
Year 2019

Rail Yield Management. Trenitalia Case

Alessandra Berto

Docenti Guida (Advisors): Prof. Andrea Pacifici
Prof. Gaia Nicosia

Abstract

The Yield Management System (YMS) described in this thesis has been developed by a joint group from IBM and Trenitalia, that I have joined since its inception. It has been implemented gradually to most trains 'Le Frecce' at Trenitalia, main Italian and 3^{rd} European railway undertaking, with 10 Million passengers and more than 260 High Speed trains offered daily, on average, in the first three months of 2018.

The operating YMS aims at maximising revenues through a two-stage stochastic optimization model which forecasts the unconstrained demand, optimizes the capacity allocations per Origin & Destination (O&D) and fare cluster, sets the protection levels using a nesting technique, develops the constrained forecasts and simulates the results.

During these years many and continuous improvements to the system have been designed, tested and deployed. Among the others: (i) a proportional correction to potential demand forecast, which is here a fundamental part of the overall optimization; (ii) the development of a methodological framework for assessing the YMS performance over time, consisting of a set of Key Performance Indicators, a Monitoring module developed from post-departure computation of the Revenue Opportunity; (iii) the design, run and analysis of a live test of a new prototype compared to the incumbent algorithm. They are subjects of this thesis, too.

Since 2005 the YMS has forecasted and optimized approximately

4 Million model instances: nearly 120 Billion decisions on combinations of fares and O&Ds, per each train, class and departure date. This automatic optimization tool has provided a powerful support to the Revenue Management team, with a high degree of productivity and solid results, even in a period of major changes in the mobility landscape, with the beginning of open track competition in Italian High Speed rail services.

Acknowledgements

I thank IBM team and in particular Dr. Emanuela Valle, for the help in collecting the results, Ing. Alberto Maria Marchetti and Dr. Stefano Gliozzi, the main inventors of this method, and all the Revenue Management (RM) team at Trenitalia and specifically Dr. Giovanna Militello, Head of RM, who provided precious insight and expertise during the research.

Contents

1 Introduction **1**

2 RM and YM for Passenger Transport Applications **9**
2.1 The Revenue Management (RM) Problem 9
2.2 The Yield Management (YM) Problem 12
2.3 Demand Forecast Models 13
2.3.1 Untruncating Demand 15
2.3.2 Micro-Level Forecasting 18
2.3.3 Macro-level Forecasting 22
2.4 Optimization Models 26
2.4.1 Independent Demand 26
2.4.2 Dependent Demand 33
2.5 Pricing and Price Optimization 41
2.6 Role of the RM team 43

3 A Comprehensive Methodological Approach to Assess RM Performance **47**
3.1 Pre and Post Departure Indicators 52
3.2 Approaches to Monitoring. The Revenue Opportunity (RO) Model 56
3.3 Simulation Models 60
3.4 Design of Experiments (DoE) Methodology 66

4 The Passenger Rail YM Problem **75**
4.1 Main Specificities of the Railway Sector 76
4.1.1 Multi-leg 'Line' Topology 78
4.1.2 Fare Basis Structure 81
4.1.3 Seasonality and Outliers 82
4.2 Main Theoretical Approaches and Implementations 83
4.2.1 Origin-Destination and Fare Inventory Control 86
4.2.2 Customer Choice 88
4.2.3 Price Elasticities 91

5 Building a Yield Management System (YMS) at Trenitalia 95
5.1 Background 95
5.1.1 Italian Competitive Landscape 97
5.1.2 Social Aspects 100
5.1.3 Multi-leg Topology and Bottlenecks 102
5.1.4 A Fare Family Clustered Structure 104
5.1.5 Demand Data Sparsity 107
5.1.6 Decisions Volume 109
5.1.7 Inventory Properties 111
5.2 YM Models and System Architecture 114
5.3 Potential Demand Forecast 117
5.3.1 Estimating the Unconstrained Demand . . . 122
5.3.2 A Multiplicative Correction 124
5.3.3 The Weights of Scenarios 127
5.4 Optimal Demand Selection 129
5.4.1 The Combined Use of Route-Topology and Fare Levers 133
5.5 Constrained Demand Forecast 141
5.6 Role of the YM Team 142

6 Assessing YMS Performance Through Live Testing **145**
6.1 Definition of a Set of KPIs 146
6.1.1 Monitoring Module and Revenue Opportunity 147
6.2 Assessing the Performance of a New YM Model . . . 156
6.2.1 Experiment Methodology and Overview . . . 161

6.2.2 Key Facts and Results Analysis 175
6.2.3 Modeling the Influence of Main External Factors on Performance 182

7 Conclusion and Future research **189**

References **195**

List of Figures

1.1 Revenue Management 2

2.1 Elements of Airline RM 14
2.2 Developments in RM forecast approaches 16
2.3 Alternative approaches for solving RM problems . . 28
2.4 O&D control early researches 30
2.5 RM System Alternatives 34

3.1 Main pre-departure Indicators 54
3.2 Main Post-departure standard indicators 55
3.3 Revenue Opportunity concepts and components . . 56
3.4 How PODS works 65
3.5 General model of a process or system 71

4.1 Summary of Railway Passenger RM Models 76
4.2 Comparison of Air versus Rail transport 79
4.3 Basic relations of nodes, legs and segments 80

5.1 The train topology 103
5.2 The Fare Basis structure 106
5.3 Overview on YMS models 116
5.4 A schematic figure of the scenario based chain of models 117
5.5 Example of booking profiles 120
5.6 Example of stacked booking curves 121

5.7 Sketch of the unconstrained forecast algorithm . . . 124
5.8 Example of Multiplicative and Additive forecasts . . 126
5.9 Example of scenario weights 129
5.10 Case study: a Milan-Naples train PAX average 134
5.11 Case study: a Milan-Naples train PAX discounts . . . 136
5.12 Partial Nesting 137
5.13 Analyst guidelines 144

6.1 Main KPIs identified per phase 147
6.2 YMS Monitoring process 148
6.3 An example of application of X Algorithm for an ODF 151
6.4 Computed difference between Actual and Optimal Revenue 154
6.5 Components of the Revenue Opportunity 156
6.6 A DoE approach fit to our case 174
6.7 Post departure KPIs: dispersion prototype vs. incumbent 179
6.8 Distributions of delta Revenue – Paired T-Test 179
6.9 Distributions of delta Revenue – Paired T-Test q/o outliers 180
6.10 Analysis of closures 183
6.11 Components of a time series 186

List of Tables

5.1 Volume of decisions taken 111
5.2 The role of YM levers 135
5.3 Practical example of availability computation 141

6.1 An example of elaboration of X Algorithm for a train/date 153
6.2 Computation of the values for the Monitoring indicators 155
6.3 Key test facts - Post departure 163
6.4 Key test facts - Pre departure 164
6.5 KPIs variation: prototype vs. incumbent 177

1 Introduction

Revenue management (RM) is an umbrella term for a set of strategies that enable capacity-constrained service industries to realize optimal revenue from operations, through selling the right product to the right customers, at the right price, at the right time (paraphrasing and integrating [Smith et al. (1992)]). Revenue management, or revenue optimization, is at the heart of a Company's strategy and provides the best results when deeply embedded in the Company's culture; RM is also related to its systems, people and processes, and relies on the availability of data in adequate quantity and quality. RM comprises several functions including Pricing Optimization, Capacity (or Inventory) Planning an Yield Management, as detailed below and illustrated in Figure 1.1 on p. 2. Main components of RM are:

(i) Pricing Optimization: monitoring and evaluating market segmentation and fares as well as setting a fare structure, defined per target of customers;
(ii) Yield Management: reacting to demand tactically and shortly in order to maximize revenue, e.g. selecting the demand with a higher value in case of (expected) contention;
(iii) Inventory or Capacity Planning: strategies to define the supply in medium and long term.

In this view, Yield Management (YM) can be described as the tactical and inventory-focused branch of RM, which consists of a set of methods, techniques and applications used to maximize a selected profitability parameter (mostly often the revenue) that can be obtained from a given supplied (perishable) set of resources under a stochastic demand willing to pay different prices [Gliozzi and Marchetti (2008)]. By fact, in practice in many cases the two terms RM and YM are used indifferently.

Since its inception in the United States in the 1970s, where it has been initially applied to the airline sector, several researches on RM and YM have been done. A common approach in transport YM consists in forecasting the demand and setting booking limits for the

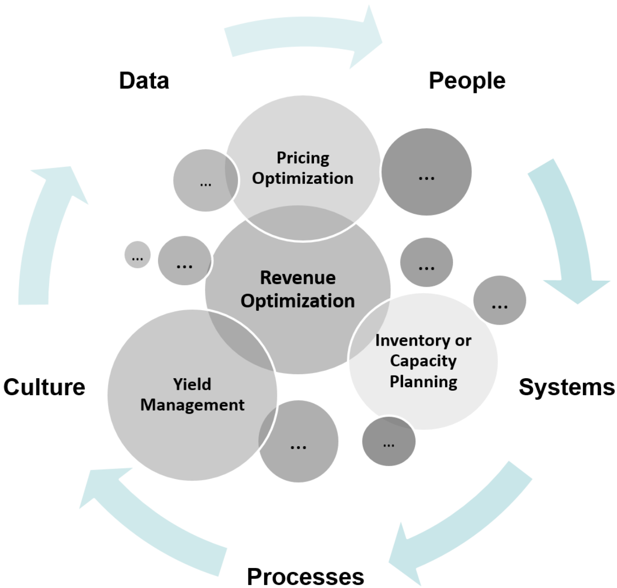

Figure 1.1: Revenue Management

different fare levels, which will determine whether any reservation request will be accepted or denied. Under this approach, starting from the potential demand an optimization is performed to set protection levels by fare and leg, segment or network. In this thesis other different approaches will be presented.

Within the Transportation YM problem, the Rail YM differs slightly from the Airline YM, in particular because in the first the demand is more related to a multileg topology with a large number of O&Ds, while the perishable and fixed offer is constrained by leg. Moreover, given the social impact of rail transportation system, the fares availability must be perceived fair by customers and prices fluctuations need to be carefully managed.

Since 2005 Trenitalia, main Italian and third European railway undertaking, operates a YM System (YMS) developed through a cooperation between IBM and the Demand Analysis & Revenue Management team at Trenitalia. Starting from a defined set of sound business rules, the YMS is able to optimize the capacity allocation per O&D based on a fixed fare family clustered structure, as it:

(i) provides the forecast of the potential demand—using an additive method with un-constraining, and a multiplicative correction—at each point of the load curve;
(ii) optimizes the capacity allocation per O&D during the booking horizon, based on a fare family clustered structure, and a set of defined business rules, through a stochastic optimization approach;
(iii) simulates the effects of the new set of inventory controls, resilient to distinct orders of arrival, in the context of partial nesting among O&D and fares;
(iv) monitors the presence of typical YM phenomena, e.g. spillage, spoilage and stifling, and the results achieved both from the YMS and the analysts, through key performance indicators and a revenue opportunity estimation.

The two-stage, scenario-based stochastic model at the very core of the system is fairly simple in its logic, and is represented as a linear program, taking into consideration O&Ds, fares, scenarios, and legs.

The adopted approach for the YMS was chosen based on its ability to fit to the Rail YM Problem, with its multi-leg topology and the presence of a fare family structure. Our approach, which exploits those factors as levers, aims to optimize the revenues by optimally managing the seats availability per Origin-Destination (O&D) or per leg at each price level through the booking horizon, on each given combination of train and date of departure. As detailed later on, the YMS uses the forecast to determine a partition of the train capacity and sets 'protection levels' against dilution; to be resilient to the variability of demand and the consequent risk of unsold the YMS adopts a partial nesting technique, which uses a variable nesting order, computed using the opportunity costs from the stochastic optimization instance.

A special mention deserves the scarcity of scientific literature on rail YM or RM. As noted by [Armstrong and Meissner (2010)] on p. 1: "While the airline and hotel industries have received their fair share of attention, the passenger rail and freight rail industries have been overlooked. There is still little published research on these industries. The lack of attention on the part of academic researchers is inexplicable.". In this same paper (p. 17) they cite only 5 papers dealing with Rail Yield Management for passenger applications up to 2010. This may be related to the recent rise of intra-modal competition; however, it is by far less than one should expect. Nowadays the main European Rail undertakings are routinely using Yield Management systems, in Germany, France, Spain, Italy, not to mention UK, where the National Reservation System provides Application Program Interfaces (APIs) to allow each company to implement their own YM.

One of the motivations of this thesis is therefore to describe the outcome of a research in this field that has led to a real implementation, with a focus on few major developments which took place in recent years, outlining main improvements towards the state-of-the art through a literature review, as detailed below.

Chapter 2 introduces the general background of Revenue and

Yield Management for passenger transport applications. This survey is centered on relevant topics for the current research which have been extensively explored in other sectors, especially aviation, and omits by purpose others which are not applied in our developments, such as overbooking practices.

Chapter 3 outlines main methodologies used across diverse industries to assess RM performances, encompassing pre and post departure measures, the Revenue Opportunity model, simulation. In addition, it provides some background information on methodological approaches to Experimental Design.

Chapter 4 presents main features for the Passenger Rail YM Problem and main literature approaches, integrated with references to methods from proprietary implementations. This builds upon outdated surveys and tries to overcome the scarce scientific literature available, therefore the Chapter includes materials from a patent and a technical white paper taken from other implemented systems.

Chapter 5 provides an in-depth description on how the YMS models work and the challenges faced while implementing and improving the YMS at Trenitalia. It also describes the main specificities of Italian railways in terms of competitive landscape, background and social aspects.

Chapter 6 details the methodological framework which has been built over time at Trenitalia to evaluate the YMS performance and determine any corrective actions, such as models calibration, user parameters improvements, or to evaluate the adoption of algorithmic improvements. This aims at creating an overall methodological framework and comprehends the definition of a set of Key Performance Indicators (KPIs), the development of a Monitoring tool based on the Revenue Opportunity estimation, the evaluation of the Experimental Design approach for planning, running and evaluating tests on algorithmic changes. It also presents how such well-rounded approach has been applied to a pilot live test on a major models change in 2018, not disclosed here for Trenitalia's choice. The confidence in the methodology, together with the encouraging results of initial tests, led to the extensive implementation of the new prototype after only 61% completion, allowing for cost saving and anticipated revenue gain.

Chapter 7 presents the conclusions and outlines some possibilities of future research.

It should be noted that parts of Chapters 4 and 5 resume and deepen the analysis from [Berto and Gliozzi (2018)] and [Berto and Gliozzi (2018b)], which have been respectively published and accepted for publication, as well as from Berto2019 which was submitted, in their first development and elaboration phases. The same work has been presented in [Berto (2018c)], [Berto (2018d)], [Berto (2018e)], [Berto and Gliozzi (2017)], [Berto and Gliozzi (2015)].

This thesis summarizes the main outcomes of many years of work within the Yield Management team at Trenitalia and dedicated working groups, where the writer has been following and participating in projects for the analysis, design, testing and commissioning of the Yield Management System and subsequent developments, for aspects concerning Operations Research.

Main activities aimed initially at building a theoretical and methodological background and defining the specificities of our problem. They have been the analysis of:

(i) the state-of-the art of methodologies and applications of transport RM, which is summarized in Chapters 2, 3, with a focus on the practices in the rail industry, presented in Chapter 4.2;
(ii) the background and specificities of the rail sector for what concerned our application, to be able to tailor it to the industry, presented in Chapters 4, 4.1, 5.1, 5.6.

It then followed the 'hands-on' activities of research and experimentation, development and calibration that are presented in Chapters 5 (more specifically in Chapters from 5.3 to 5.5) and 6. In particular, among these, the main ones were:

(i) modification of the forecasting model, presented in Chapter 5.3;

(ii) study, test and implementation of a Monitoring module with definition of a set of performance indicators, as detailed in Chapters 6.1 and 6.1.1;
(iii) creation of a live pilot test aimed at evaluating a new prototype and related analysis; this is outlined in Chapter 5.3.

In these areas the specific contribution of the author focused on the following activities:

(i) Research of new methodologies, starting from the analysis of scientific literature of reference and comparison with other proprietary implementations, as presented in Chapters 2, 3, 4.2;
(ii) Preliminary analysis of industry data and definition of business requirements, outlined in Chapter 5.1;
(iii) Definition of technical requirements and application perimeter, as presented in Chapter 5.1;
(iv) Proposals, revisions and modifications of algorithms and formulas, as detailed in Chapters 5.3, 5.4, 5.5;
(v) Testing, also on prototypes, with calibration of the models and initial setting of the user and system parameters, in one case a live pilot test; this is explained in particular in Chapters 5.3, 5.4, 5.5, 6;
(vi) Definition of the methodological framework for performance assessment and test in operation after deployment in YMS for validation or possible corrections, which are the subject of Chapter 6.2.

Such individual and autonomous contribution has led to significant innovations compared to past practices; in particular, reference is made to the evolution of demand forecasting models, the Monitoring models and the evaluation of the performance of modified algorithms.

2 Revenue and Yield Management for Passenger Transport Applications

2.1 The Revenue Management (RM) Problem

Revenue management (RM) can be defined as a set of methods and tools to maximize a certain measure of profitability, optimizing the extraction of customer value (e.g. through the price paid) and matching demand and supply. Lieberman (1991) and Skugge (2004), both mentioned in [Temath (2010)], suggest a potential revenue growth of 3 to 7 percent through the use of RM with respect to a non-RM case, generally associated to the 'First Come, First Served' approach.

The concept of RM was generated at US Airlines, following the liberalization of US aviation industry since 1970s and the consequent increase of competition. RM has been applied in several industries, mostly in periods of increased competition when the revenue performance was of vital importance: initially in airlines, then to other means of transportation (railways starting from late 1990s), later in services, manufacturing and industrial sectors, and the list is increasing (examples in: [Gliozzi and Marchetti (2003), Zatta (2007), Zeni (2001)]).

Following [Gliozzi and Marchetti (2003)] RM optimization can be fruitfully applied to any business with the following characteristics:

(i) Several products (or services) and prices which compete over one—or more—resources;

(ii) Stochastic demand which will pay different prices for consuming the same resource, and can be subdivided into different ‘segments’;

(iii) Products/services are scarce, which means that at least in some cases the demand exceeds the supply or capacity available at certain conditions;

(iv) Products/services are perishable, i.e. cannot be sold or used anymore after a determined moment;

(v) The offer is considered almost fixed in the short term;

(vi) Clients make reservations for the products and expect immediate response to their reservation request.

Under this “classic” approach, customers are usually subdivided into segments, i.e. groups of passengers that are relatively homogeneous for characteristics and behaviors. The customer value extraction is mainly obtained through managing price levels and/or available capacity of products or services in relation to specific target segments, with the application of restrictions like the ‘fencing rules’. The introduction of additional fare classes with the use of price discrimination allows for revenue increase. The presence in the market of the Low Cost Carriers, which often have a simplified fare structure and scarce use of fences, impacted the usage of this approach to pricing structures.

One of major RM challenges is related to the demand forecast, which is part of RM overall optimization framework and has a great impact on RM performances. In particular, a major issue is related to the estimation of the part of the demand which was not satisfied due to the unavailability of the preferred product or service.

Among the others, RM is generally based on a smart use of data analytics, e.g. to understand demand and in particular predict the customer needs, choice rationals and willingness to pay. The performance of RM is greatly related to the availability of data in adequate quality and quantity. Such data may come from the Company (demand, customer and business intelligence data, normally in dedicated data-bases and data-warehouses), external data providers (statistics and market researches, competitor data), or it can be web-based data (e.g. weather, trends and events, sentiment analysis, competitor intelligence).

Finally and as anticipated earlier, while the borders between RM and YM are often fuzzy and the two disciplines can overlap, here an enlarged vision of RM is adopted. Therefore, RM is intended as a superset of YM, including not only the tactical decisions on pricing and/or inventory controls typical of the YM, but also the operational to strategic and medium to long term activities. Therefore, under this approach the Revenue Management Department can embrace many functions of a Company, comprehending:

(i) Yield Management - reacting to demand tactically and shortly, e.g. selecting the demand with higher value in case of (expected) contention in order to maximize revenue;
(ii) Inventory or Capacity Planning strategies - rightsizing supply across different time horizons, from the short and medium to the long term, according to demand;
(iii) Pricing - defining optimal fare levels and defining a clustered fare structure with its set of rules, defined per target of customers;
(iv) Reporting - providing input for targeted actions from Strategy, Marketing and Communications, Customer Relationship Management, and other Departments.

Several research has been performed on RM topics across diverse industries, from both a theoretical and a practical point of view. An overview on the state-of-the-art of RM researches and approaches is provided by [McGill and Van Ryzin (1999)], [Talluri and Van Ryzin (2004b)] and [Chiang et al. (2007)]. Among main RM surveys and introductory researches describing RM and its applications, the following are also mentioned in [Temath (2010)] and [Zeni (2001)]: Cross (1995) providing an introductory work on the topic; Weatherford and Bodily (1992) describing the problems where RM is applicable; furthermore, the following works: Cross

(1997), Klein and Steinhardt (2008) and Phillips (2005), Yeoman and McMahon-Beattie (2004), Kimms and Klein (2005).

2.2 The Yield Management (YM) Problem

Yield Management has been already described as the tactical and inventory-focused part of Revenue Management. It consists of a set of methods, techniques and applications used to maximize a selected parameter that can be obtained from a given supplied (*perishable*) resource under a stochastic demand which will pay different prices for the utilization of the same resource. In the followings, revenue is considered as the optimized parameter, but several other measures, such as profit, yield, load factor, measures of market share could be considered instead ([Gliozzi and Marchetti (2008)]).

YM typically uses either tactical controls of inventory or dynamic pricing to select the most valuable demand when the capacity is insufficient ([Zatta (2007)]). The main approaches to YM are based on understanding, anticipating and influencing the consumer behavior in order to maximize a measure of the revenue (or other parameters), and comprehend the use the price lever e.g. to move part of the low-pay demand from peak to off-peak times and services, freeing up space for the high-pay demand ([Gliozzi and Marchetti (2003)]).

One of most common approaches in transport YM consists in forecasting the demand and setting — through a determined optimization approach — booking limits for the different fare levels. This in turn will determine whether any reservation request will be accepted or denied.

The reservation system will track only the accepted bookings, which will form the historical base used in main approaches to forecast the future demand, alone or with other factors, taking into account similarities of product or services. For instance, for trains or flights they can be the day and hour of departure and seasonality.

2.3 Demand Forecast Models

According to common definitions of YM, it is greatly related to an understanding of customer behaviors, demand levels and perception of product value from the market to consequently align prices, placement and availability of products and services, for every segment of target customers. Therefore, the timely availability of data in adequate quality and quantity is paramount, as well as the choice of the forecasting method, that should be adequate and fit to the complex and real-time nature of RM problem as much as possible. Forecast is a key point in YM and a fundamental part of the overall optimization, as it provides key input to the optimization models together with the actual bookings and the data on inventory and fares. Based on [Zeni (2001)], demand forecast is crucial as in YM inaccurate forecast can easily lead to inadequate inventory controls and suboptimal decisions on inventory controls and poor revenue performance.

[Belobaba (2002)] presents how RM is sensitive to improvements in forecast, in relation to a set of aviation RM models compared using PODS simulator. [Lee et al.(1990)] outlines how a 10% improvement in forecast accuracy can result in a revenue increase of 0.5 to 3%. According to [Temath (2010)], further researches on the improvement of revenue performances after an increase in forecast and un-constraining accuracy were done by Weatherford and Polt (2002), [Lee et al.(1990)] and Weatherford and Belobaba (2002).

According to [McGill and Van Ryzin (1999)], forecasting demand can be a difficult activity due to the presence of several complicating factors, that are involved in the systematic process, such as: seasonal factors, complex business and social environments, changes in fare prices and competition. Among complicating factor for predicting demand one could mention: seasonality, day-of-week and time-of-day variations, special events, price sensitivity, demand of fare classes, group purchases, cancellations, censoring of historical data, delayed services, recapture and no-shows. Figure 2.1 on p. 14 presents the main elements involved for Airline RM; by fact, most of them are related to customer behavior and demand forecasting.

In addition, the use of historical data in the process of forecasting may bring errors, e.g. due to the presence of missing or incorrect data. Furthermore, the demand forecast should be provided with a sufficient degree of disaggregation to feed the models. Finally, a typical issue is the truncation, or the incapacity to observe the real demand but only a part of it, which will be described in the followings.

Elements of Airline Revenue Management

Customer Behavior and Demand Forecasting	Revenue Factors
Demand volatility	Fare values
Seasonality, day-of-week variation	Uncertainty of fare value
Special events	Frequent flyer redemptions
Sensitivity to pricing actions	Company or travel agent special vouchers
Demand dependencies between booking classes	*Cancellation* penalties or *restrictions*
Return itineraries	Variable Cost Factors
Batch bookings	Marginal costs per passenger
Cancellations	*Denied boarding* penalties
Censorship of historical demand data	*Goodwill costs*
Defections from delayed flights	Fare Products
Diversions	Number of products
Go-Shows	*Fences* (restrictions)
Group bookings	Problem Scale
Interspersed arrivals	Large airline or airline alliance; e.g., United/Lufthansa/SAS ORION System: 4,000 flights and 350,000 passenger itineraries/day [see GARVEY (1997), BOYD (1998)]
No-shows	Problem Interfaces
Recapture	Market strategy
Upgrades	*Code-sharing* alliances
Control System	Routing
Booking lead time (often 300 days or more)	Gate acquisition and schedule planning
Number of *controllable booking classes*	*Fleet assignment*
Leg-based, segment-based, or full *ODF* control	
Distinct *buckets, parallel nesting*, or *full nesting*	
Reservations systems *connectivity*	
Frequency of control updates	
Overbooking	

Figure 2.1: Elements of Airline Revenue Management, from [McGill and Van Ryzin (1999)], page 235.

According to [McGill and Van Ryzin (1999)], the first attempts to model demand distributions and components started very early with statistical descriptions in Beckmann and Bobkowsi (1958), Lyle (1970), while Martinez and Sanchez (1970) analyzed data from Iberia in relation demand of the airline sector. Belobaba (1987a) and Shlifer and Vardi (1975) performed empirical studies on the efficacy of the normal probability distribution to provide a good continuous approximation and aggregate airline demand distributions. Other studies instead presented the inappropriateness of the method when applied to more and more disaggregated levels of demand - the levels which are necessary for RM implementations. The above mentioned researchers by fact used a model for the stochastic process of arrival of individual demand.

In the followings, the main forecasting approaches in use are outlined, as presented by [Zeni (2001)]. They are summarized here below in their progression:

(i) Deterministic demand, considered in the initial research works;
(ii) Micro-level forecasting, provided through the use of statistical measures to represent demand distributions and their unconstrained values;
(iii) Macro-level forecasting, encompassing customer centric and market aware models.

Initially the research focused on single legs, while at a later time [Williamson (1992)] studied on the vast amount of possible itineraries across airline networks, characterized by very different demand levels, so that e.g. statistical probability to traverse the ones with lower demand was very low while the contended ones were associated with higher prices; it also suggested the aggregation of demand in groups and the use of their averages. At the same time, it is worth mentioning the distinction by Ratliff et al. (2008) of single-class, multiple-class and multiple-flight models, which were historically grounded on the average or statistical measures of historical bookings and departures of a similar flight. Figure 2.2 on p. 16 illustrates major development factors in RM forecast approaches over time.

2.3.1 Untruncating Demand

This section presents a set of forecasting methods for transport revenue management (mainly airlines) in relation to the issue of censored demand data. By fact, the total demand for a certain combination of market and fare cannot be observed, but only the booked part which found some availability: during the booking horizon,

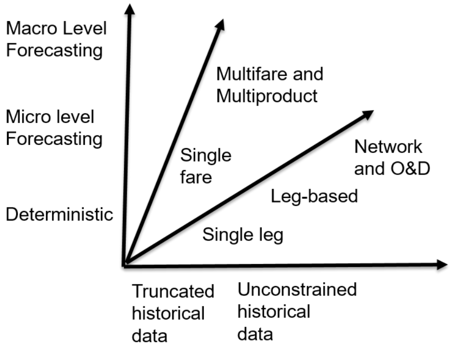

Figure 2.2: Developments in RM forecast approaches

reservations are accepted until the booking limits are reached. From this moment onwards the system stops accepting reservation requests and collecting data on that specific demand.

As the reservation system keeps only track of the accepted bookings, they will form the historical base to forecast the future demand. This leads to the presence of censored, or truncated, data and the consequent necessity to apply proper techniques to understand the real customer behavior, which is forecasted from the available data (booking profiles) and the current status of the reservations without the constraints of offered capacity and inventory controls. Collecting directly the data on the missing reservations is not a viable option ([Zeni (2001)]). However, trying to record or get proxies for the denials (i.e., customer demand didn't find an adequate price, so the client denied the booking) and regrets (i.e., customer demand is not accepted because of capacity issues) could be feasible in specific markets (e.g., freight related) where the booking process is more complex and requires a real—and recorded—formal negotiation process.

Using only the constrained (or truncated) demand would create negatively biased forecasts; as a consequence, the resulting inventory controls would tend to reserve few seats to high-pay customers; this outcome is commonly defined high spill. It has been estimated that the related dilution can cause up to 3% of potential revenue loss ([Zeni (2001)]); dilution can be defined as the sale of discounted rates to a customer likely to purchase at higher rates) Weatherford and Polt (2002) were able to measure and demonstrate the revenue increase related to better un-truncation ([Temath (2010)]). This provides evidence to the necessity to apply proper techniques to estimate the real customer behavior, not constrained by capacity limits, and improve forecast accuracy.

[McGill and Van Ryzin (1999)] outline how soon arose the problem of un-truncating demand, i.e. estimating the potential demand which could have been accepted without capacity constraints: Swan (1995) and [Lee et al.(1990)] were among the first researchers to focus on this topic, with McGill (1995) who developed a method based on multivariate multiple regression. [Taneja (1978), Sa (1987)] and Botimer (1997) developed methodologies to aggregate and disaggregate forecasted demand, mostly based on regression techniques, and proved that the latter worked better in comparison to analysis of time series or historical averages. Regression and time series techniques proved to be useful for aggregated demand but not for the disaggregated one.

Also, main researches were done on how to obtain forecast from historical data. According to [McGill and Van Ryzin (1999)], the early researches from [Harris and Marucci (1983)], [L'heureux (1986)], Adams and Michael (1987), and [Smith et al. (1992)], pointed out the importance of the recent history of bookings and their actual levels for the same or similar flights to retrieve key information on potential bookings for a certain flight; and incorporated partial reservation data from similar flights using smoothing techniques. In-depth introductory works on un-constraining and approaches for demand forecasting are provided by [Zeni (2001)] and Polt (2000), while other main researches on the topic comprehend Zeni (2003), Zeni and Lawrence (2004), Talluri and van Ryzin (2004a) and Weatherford and Polt (2002), Lee (1990) and [Talluri and Van Ryzin (2004b)] as stated by [Temath (2010)].

According to [Zeni (2001)], few research has been done in RM to handle incomplete historical data, much more in other applications. There, current practices and forecasting methods, aiming at predicting the future demand given the historical data (booking profiles) and the current status of the reservations, comprehend Micro and Macro Level Forecasting which are described in the followings. The classification into Independent and Dependent Demand Models is followed, as provided in [Talluri and Van Ryzin (2004b)] and [Temath (2010)].

2.3.2 Micro-Level Forecasting

One of the key underlying assumptions of early researches was the independence of demand among diverse fares clusters, described in [Talluri and Van Ryzin (2004b)] and summarized also in [Temath (2010)]. It assumes that demand for different fares are not related; this also applies, in general, to alternative services. This, together with the Low Before High (LBH) assumption on the arrival order of passengers in relation to their possibility to spend, has been mostly related to the use of restrictions such as fenced fares.

According to [McGill and Van Ryzin (1999)], various approaches have been proposed as the basis of forecasting demand with censored data([Maddala (1983)]); this provides a systematic way of estimating censored and truncated models where normally distributed data, features of the estimators, extensions and other distributions are outlined ([Schneider (1986)]). [Zeni (2001)] provides an overview of major methods for un-truncation, that comprehend booking profile analysis, moving average, linear regression, mean imputation, projection de-truncation, expectation maximization, additive and multiplicative pickup, exponential and double exponential smoothing, the latter developed by Crystal et al. (2007). In practice, quite often simple models based on moving average and smoothing techniques paired with analysis of booking profiles are still practiced; in particular, exponential smoothing has been widely used for its robustness, good quality of results and easiness of utilization.

[Taneja (1978)] has successfully incorporated traditional regression techniques and combined statistical methods to forecast demand in the airline industry. The research relies upon regression model, multi-equation analysis, test model assumptions, rigorous model specification and selection as the basis to forecast demand. According to [Taneja (1978)], causal models result appropriate compared to time-series approach, which is able to predict demand accurately but fails to explain motivations behind demand levels.

[Sa (1987)] favors a regression model as it can improve the performance of the implemented RM system better than time series. In relation to this, [Kanafani (1983)] has effectively pointed out a set of relevant aggregate measures for air travel encompassing number of passengers, aircraft operations, revenue and distance traveled. Such measures can be classified based on purpose, temporary stratification, origin and destination, trip duration and offered services. [Sa (1987)] attempts to use a regression model for bookings to come using an autoregressive integrated moving average (ARIMA) time series in a single cabin class but fails to obtain good results. A new experiment is then provided, with bookings-to-come being the dependent variable and current bookings as explanatory variable. The results of the latter are positive, though facing limitations as they did not incorporate booking limits as part of the statistical analysis.

[Brummer et al. (1988)] attempted to compute mean and standard deviation using an uncensored log-normal distribution with data from constrained observations. This work was advanced to provide three models: flight specific, class-specific and reservation forecasting by [Ben-Akiva and Lerman (1985)]. However, the findings of the study demonstrated that a combined model was far better than historical and advanced bookings. The limitation of this strategy is that the researchers never had enough data for validating the results within the estimated models on the predicted future flights. Secondly, they relied upon monthly data as opposed to daily data that can properly be achieved through microlevel forecasting of the choice underlying booking process. Consequently, the booking limits that were incorporated on the fare classes resulted to censored data.

Booking profiles can be effectively used to forecast short-term demand, as presented in [Harris and Marucci (1983)]. This can be achieved through the use of booking profiles, based on a set of five consequent snapshots of the bookings prior to the expected date of departure for the individual flights, as outlined in [Lee et al.(1990)]. A booking profile is a representation of all existing booking history. The model provides passenger demand forecast that is a function of current bookings at a specified time period [L'heureux (1986)]. The basic principle here is to forecast demand at departure time for all available seats, using historical data and current state of bookings; the latter is denoted by the number of reservations in place at a certain moment before departure, when the forecast is performed.

As such, according to [Zeni (2001)] there are five main micro-level forecasting models that can be implemented, all relying on historical booking data and under the assumption of demand independence. This classification include Exponential Smoothing, Moving Average, Linear Regression and (Additive and Multiplicative) Pickup Models outlined in [Weatherford (1999)]:

(i) Exponential Smoothing, a time series forecasting method that incorporates decreasing weight to make observations as they increase;
(ii) Moving Average, which determines the forecast for future demand through undertaking an average number of recent historical observations;
(iii) Linear Regression, which relies on the prevailing assumption that there exists a linear trend between final bookings at departure and at several days prior to departure;
(iv) Additive and Multiplicative Pickup Models, that add historical incremental bookings to the ones at departure and current reservations, prior to departure to determine the demand forecast, so that the final bookings level will be computed as the sum of current bookings and expected incremental reserva-

tions at departure;

(v) Sell-up Model, the most advanced, which incorporates some demand dependency among fares. Even though not included by the abovementioned works, it still falls within the Micro-level Forecasting according to [Temath (2010)].

The following part provide an overview on pickup models, then outlines the sell-up model.

Pickup models

Classical pickup models rely on historical bookings data as the basis of forecasting demand ([Duncanson (1974)]) and don't consider other available information that can be important in forecasting the level of demand. As such, they lack causal variables like population, employment, income and other important economic activities, as well as information on customer behavior or the competition landscape. As a second shortcoming, any weakness in the data will be reflected in the accuracy of the forecast.

[L'heureux (1986)] stresses the importance of observing the behavior of bookings and specifically the paramount importance of recent history in the forecasting process. Furthermore, through exploiting such data, it is likely that the model would improve its ability in handling abnormal events, like departures falling into holiday periods for instance.

Under the Additive Pickup approach the average pickup is obtained from historical data, subtracting the average bookings of a certain day from the average bookings at departure. Differently, the Multiplicative Pickup models utilize historical pickup data to forecast bookings, but multiplies current bookings by the average pickup ratio. The data used in the calculation of pickup ratio is determined by the selected methodology, so e.g. it can include only recent history. They are presented in two major forms, classical and advanced, in [Wickham and Richard (1995)].

Sell-up models

They estimate and incorporate in RM models the probability of sell-up, i.e. the purchase of a higher booking class by a customer in the case a certain fare level is closed to sales. Based on [Temath (2010)], Sell-up models have been included among the Independent Demand forecasting approaches. It shall be noted that other models computing price elasticities are described later within the Macro-level models, as they consider a wider range of factors and data to compute the probability of recapture and selling up, not at train or flight level, but at macro level.

2.3.3 Macro-level Forecasting

This paragraph presents an overview of the approaches to Macro-level forecasting, that overcome the underlying assumption of independent demand from classical models falling within the micro-level approach. Macro-level forecasting appears more representative of the real world, where the customer is able to compare prices and offers from several carriers (e.g. through web searches) and choose according to his own preferences and (or) Willingness to Pay (WTP) for a certain service, and the diverse options are characterized by a certain degree of substitutability. This approach is able to take into account customer choices and behaviors to estimate demand elasticities, as well as cross-elasticities.

Furthermore, it allows to consider external environment and market and phenomena, like recapture and up-sell, which are connected to elasticities. By fact, demand dependencies between fare classes and services become evident when passengers are unable to book, due to unavailabilities, capacity or travel restrictions, service cancellations or delays, or other reasons. This can happen, for instance, in case a flight cancellation compels passengers to travel on flights that are not of their preference. In such cases, it is reasonable to assume that a certain part of the demand will shift to other available fares or services. *Recapture* takes place when the customer, that changed her mind or wasn't able to proceed with the purchase at a certain point of the booking process, proceeds with other services from the same company. *Up-sell* refers to the case of a customer booking a more expensive fare of the same or similar travel service, due to the unavailability of the (lower) fare which was of her preference. Other typical phenomena related to the substitutability of travel services, together with the cancellations, are *no-show* (i.e., booked passengers that decide not to travel but keep the reservations) and *go-show* (i.e., passengers that get on-board without any previous reservation or booking).

According to [Temath (2010)], Ratliff e al. (2008) and [Cleophas et al. (2009a)] provide comprehensive overviews on main researches on Dependent Demand approaches, or Macro-level forecasting. The main models falling under Macro-level forecasting approach are Hybrid Demand and Customer Choice Models, that are presented in the following part of this paragraph.

Hybrid demand models

[Talluri and Van Ryzin (2004b)], mentioned also in [Temath (2010)], present the hybrid demand approach, which aims to model the contemporary presence and mix of product-oriented and price-oriented customers:

(i) The *yieldable* (or product-oriented) demand is independent, but relates to up-sell up or down-sell from other booking classes. This is related to fenced pricing.
(ii) The *priceable* (or price-oriented) demand is characterized by a certain willingness-to-pay, according to which customers opt for the cheapest booking class available. This is related to restriction-free-pricing.

According to [Temath (2010)], main descriptions of hybrid demand models can be found in Fiig and Isler (2004), Boyd and Kallesen (2004), Walczak et al. (2010) and Fiig et al. (2010), which can be seen as special cases of the model by Winter (2010) providing an extension to buy-down estimation.

Customer-choice models

The availability of diverse means of transport together with an increasingly high price transparency pose challenges to the classical RM approaches. As discussed earlier, traditional RM models are based on historical data and work at their best when the contention is relatively high, furthermore they assume demand independence among classes and fares. In particular, classical models do not consider the process and behavior of choice from the customer as well as the market situation at the moment of the booking, with particular regard to availability of alternative services and related prices; for instance, they do not account for the possibility to either travel by high-speed train, by car or booking a flight on the same itinerary.

Customer-choice models, instead, explicitly account for the availability of diverse possibilities to the customers, that choose among a set of alternative modes of transport, Companies and services. Specifically, they model the customer behavior of choice following multiple criteria, based on their preferences and the set of utilities associated to service features such as price, travel time and others. Customer data feeding the model may include socioeconomic and demographic factors as well as behavioral ones like traveling preferences and frequency, advance of purchase, reaction to certain events as well as policies or actions from the considered Company. For this purpose, customers can be subdivided in clusters or segments which should be reasonably homogeneous, numerous and populated.

According to [Zeni (2001)] and [Lee et al.(1990)], this approach considers a passenger as provided with several options, having to select a suitable one. In the context of air travel, main options can be listed as follows: route, airport, airline and fare type. The passenger choice is modeled through a multinomial logit method in [Ben-Akiva and Lerman (1985)]. [Temath (2010)] outlines how researches on discrete choice behavior modelling, primarily focused on multinomial logit estimations, were done mainly from [Ben-Akiva and Lerman (1985)] and Hopperstad (1997), while Gallego (2009) focuses on the Expected Marginal Seat Revenue (EMSR) model, which will be described later, incorporating diversion and recapture of passengers in different classes for single leg RM. Weatherford (2002) instead treated diversion within a stochastic model of arrival for two classes. [Temath (2010)] and [Cleophas (2009)] provides an overview on main approaches, which comprehend in particular the works from Kimms and Muller-Bungart (2006) and [Talluri and Van Ryzin (2004b)].

Further researches

[Temath (2010)] provides a wide overview on main researches related to dependent demand models, comprehending Ratliff et al. (2008), Mishra (2003), McGill (1995), Skwarek (1996). Cleaz-Savoyen (2005), Belobaba and Hopperstad (2004), Gorin and Belobaba (2004), Kambour et al. (2001) and Reyes (2006) researched on Q-forecasting, used to compute the priceable demand component in hybrid models, e.g. through the use of PODS simulator which will be described later in section 3.3. Ratliff et al. (2008), Stefanescu (2009) and Stefanescu et al. (2004) focused on unconstrained demand estimations, considering recapture and applicable to multiple flights; Ja et al. (2001) considered also connected flights. Talluri and van Ryzin (2004), Vulcano et al. (2010) and Vulcano et al. (2009) researched on the Expectation—Maximization approach.

This paragraph has presented how the consideration of customer and market data in a forecasting model improves its accuracy, with respect to traditional models that are based on historical data only and assume demand independence. Furthermore, given the total demand for transport in a certain market, the modal market shares can be improved or lowered by determined actions by a certain company and its competitors and consequent customer reactions. So, companies can expand their market share leveraging key factors for customer choice, such as: price and immediate availability, service quality and customer experience, frequency and punctuality. On the other hand, competitors can offer alternative services within the same mode (intra-modal competition) or another mode (inter-modal competition). In addition, in a long term perspective the overall transport demand shouldn't be considered fixed; contrarily, it can be influenced by the abovementioned market dynamics.

2.4 Optimization Models

2.4.1 Independent Demand

According to [McGill and Van Ryzin (1999), Chiang et al. (2007)], several approaches and methods have been developed across the last 40 years on RM optimization methods. The development of research activities displayed a progression under many factors. As described earlier for the demand forecast, the early researches assumed that the demand for a specific market and fare was independent from demand on other cabin classes; such assumption was overcome from newer models. Furthermore, initial works started with fairly simple cases, analyzing a single leg and being later able to encompass more ones, then assess the effects of a reservation over an entire network (or part of it). Another factor of progression was related to the improvement the estimation of 'displacement costs' of offering a certain service on a segment which traverses a set of adjacent legs. They are related to the cost structure of many industries under the transport umbrella, characterized by the presence of huge fixed costs and low variable costs associated to a certain service.

As reported in [Temath (2010)] the main underlying assumptions for those models were the following:

(i) demand is independent (therefore phenomena like sell-up and recapture are not included);
(ii) booking order follows a Low-Before-High (LBH) order (this is negligible with subsequent dynamic programming methods);
(iii) booking classes are sequential (this allows for the fruitful

application of the nesting principle, that will be described later).

As reported in [Zeni (2001)] the early optimization approaches to maximize revenue, mainly used in aviation RM, were based on:

(i) The Littlewood's rule, which sets the protections level taking into account the revenue of the upper fare, weighted by the probability of realizing it (value of holding the seat) compared to the actual revenue of the lower fare;

(ii) The Expected Marginal Seat Revenue (EMSR) approach from [Belobaba (1989)], which generalizes the Littlewood's rule for more than two fare classes, so that the EMSR represents the expected value of a certain seat and can be computed as its price multiplied by the probability of selling it as additional seat;

(iii) Network Formulations, which consider the value of a seat in relation to the overall network effect of the sale of that seat;

(iv) Deterministic Linear Program, that substitutes the expected demand with its deterministic value;

(v) Probabilistic Nonlinear Program, considering the non-deterministic nature of demand for nonlinear optimization.

[Chiang et al. (2007)] illustrate different alternative approaches for the solution of RM problems, also displayed in Figure 2.3 on p. 28 from [Chiang et al. (2007)]. In the followings these methods are illustrated: Leg-based control, Network-based control, Bid Price.

Leg-based control

[McGill and Van Ryzin (1999)] stated that most of the initial works focused on the single-leg seat inventory control across diverse fares, with a progression from the Littlewood's rule applied to two classes to EMSR control for multiple classes, optimal booking limits for single-leg flights, segment control and Origin-Destination and Fare (ODF) control, also thanks to the improvement of displacement costs

Table 5 Approaches used for solving revenue management problems

Linear programming	Cooper (2002), Möller et al. (2004), Pölt (2004)
Integer programming	Bertsimas and Shioda (2003)
Scenario tree	Möller et al. (2004)
Markov model	Brumelle and Walczak (2003), El-Haber and El-Taha (2004), Lindemann et al. (2004), Aviv and Pazgal (2005), Feng and Gallego (2000)
Dynamic programming	Bertsimas and Popescu (2003), Bertsimas and Shioda (2003), Brumelle and Walczak (2003), El-Haber and El-Taha (2004), Bertsimas and de Boer (2005a), Savin et al. (2005)
Bid-price methods	Kraft et al. (2000)
Stochastic gradient algorithm	Karaesmen and van Ryzin (2004), Bertsimas and de Boer (2005a)
Stochastic programming	Lai and Ng (2005)
Simulation	Oliveira (2003), Anjos et al. (2004), Kimes and Thompson (2004), Zhang and Cooper (2005), Bertsimas and de Boer (2005a)
Polyhedral graph theory approach	Kuyumcu and Garcia-Diaz (2000)
Reinforcement learning	Gosavi et al. (2002)
Adaptive algorithm	van Ryzin and McGill (2000)
Machine learning algorithm	Neuling et al. (2004)
Hierarchical mathematical optimisation	Cote et al. (2003)
Fixed point algorithm	Friesz et al. (2005)
Continuous time approach	Kosten (1960)
Robust optimisation	Koide and Ishii (2005), Lai and Ng (2005)

Figure 2.3: Alternative approaches for solving RM problems from [Chiang et al. (2007)], page 115.

evaluation. Secondly, first researches present preliminary analysis with competition among two distinct flights, possibility of upgrades to higher price classes.

According to [McGill and Van Ryzin (1999)], the initial studies were done with the underlying assumption that demand is independent and distributed uniformly for every category of fare classes. A milestone has been the Littlewood's rule (Littlewood, 1972) for optimization models of leg-based seat inventory controls, which applied to two fares class problems. It was further extended to multiple booking classes by Belobaba (1987) and [Belobaba (1989)] that also developed EMSRa and EMSRb heuristics to determine booking limits for seat inventory control. As presented in [Temath (2010)], EMSR and EMSRb are still in use due to their simple implementation and ability to provide good results in comparison to other methods; further studies have developed methods for the optimal booking limits with extended distributions and dependencies: Curry (1990), Wollmer (1992), Brumelle and McGill (1991), Brumelle et al. (1990).

Network-based control

As reported in [Williamson (1992)], there is a very large numbers of possible itineraries in large hub-and-spoke networks. Any decision on requested itineraries impacts on many legs, therefore very soon the interest of research incorporated the network effects, accounting for dependencies among the diverse legs of an airline networks. Therefore, and often due to the limited possibilities of the Global Distribution Systems (GDS), optimal solutions were conveniently grouped in a small number of controllable booking classes.

According to [Temath (2010)], overviews on network-based optimization approaches are provided by Barnhart et al. (2003), [McGill and Van Ryzin (1999)] and [Talluri and Van Ryzin (2004b)]. In particular, [McGill and Van Ryzin (1999)] outlines main researches listed in Figure 2.4 on p. 30.

[McGill and Van Ryzin (1999)] explained how first researches

TABLE VII

Origin–Destination Control 1982–1999

Year	Reference	Year	Reference
1982	D'Sylva	1990	Vinod
1982	Glover et al.	1990	Wong
1983	Wang	1991	Phillips, Boyd, and Grossman
1985	Simpson	1991	Vinod
1986a,b	Wollmer	1992	Williamson
1987a,b	Belobaba	1993	Talluri
1988	Dror, Trudeau and Ladany	1993	Wong, Koppelman, and Daskin
1988	Smith and Penn	1994a,b	Talluri
1988	Williamson	1995	Vinod
1988	Wysong	1996	Talluri and van Ryzin
1989	Simpson	1997	Garcia-Diaz and Kuyumcu
1989	Vinod	1999	Ciancimino et al.
1990	Curry	1999a,b	Talluri and van Ryzin
1990	Vinod and Ratliff		

Figure 2.4: O&D control early researches, from [McGill and Van Ryzin (1999)], page 242.

over Origin-destination control started from 1982 and led to the development of methods based mainly on Mathematical Programming Formulations, Segment Control, Bid Price Methods. Glover et al. (1982) provided the pioneer network formulation that were later applied optimal booking limits by Curry (1990). According to [McGill and Van Ryzin (1999)], among Mathematical Programming formulations main researches included: minimum cost network flow formulation per O&D from Glover et al. (1982), with a focus on network effect and deterministic demand; a Linear Program network formulation with stochastic demand, incorporating in the objective function the expected marginal seat values, from Wong, Koppelman, and Daskin(1993). These formulations had the potential to be incorporated in bid price approaches but the allocations they determined were non-nested. In addition, the large size of the problem was a challenge. Curry (1990) presented a combined formulation of mathematical programming and marginal analysis able to provide distinct bucket allocations per O&D, which were then nested, finally each nest was optimized separately per single-leg and nested booking limits.

According to [McGill and Van Ryzin (1999)], through further research activities, Smith and Penn (1988) and Simpson (1989) incorporated a bid price concept to be used alongside the network revenue management. The research on bid-price controls, which will be detailed in the followings, led to the development of an adaptive scheme to manage protection levels, created according to the frequency of soldout events and offering solutions for optimal conditions. The studies used historical bookings as a basis for determining the rate of occurrence of sold-out events. In those approaches, assumptions on distributions and un-censoring were necessary. To overcome this limitation, novel stochastic network models using Markov Decision Process were developed to provide different types of approximations (van Ryzin and Talluri, 2003). Later on, Markov decision process was combined with mathematical programming (Cooper and Homem de-Mello, 2003).

As outlined in [McGill and Van Ryzin (1999)] researches based on segment control, or partial O&D control, were initially performed by Smith and Penn (1988) and Vinod (1995) and allowed for the estimation of the value of a multi-leg itinerary (not comprehending different flights) and exploited the information on booking closures from the reservation systems. [Temath (2010)] reports on how several other works tried to use adjusted leg-based optimization to take into consideration network effects, obtaining booking limits over the network with leg-based capacity constraints. These works included: Williamson (1992, 1988), Smith et al. (1992), Vinod (2005) and Talluri and van Ryzin (2004b). They developed approaches encompassing prorated EMSR heuristic, virtual nesting, a deterministic linear program (DLP) considering mean demand which was extended by Talluri and van Ryzin (1999) with the randomized linear program (RLP). The latter evidenced how considering the demand variance allowed for improved bid prices.

According to [Temath (2010)], initial works Virtual Nesting from Belobaba (1987a), Smith and Penn (1988), Williamson (1998), [Williamson (1992)] and Vinod (1989, 1995) aimed at accommodating demand on the few booking classes available on reservation systems through techniques to achieve something similar to network control through clustering ODFs into single leg booking classes. Such clustering, or indexing, was reached through diverse techniques based on total value as well as leg value, and (most interesting for us, and dominating) by estimated net value allowance for displacement effects obtained as dual prices from a deterministic network LP; the latter was studied from Smith and Penn (1988) and Williamson (1988).

[Belobaba (2002)] outlines the main features of Network Bid Price Control, relying a lot on frequent re-optimizations for its effectiveness but at the same time being simple to implement in its control mechanism. On the other side, the Value buckets (or “virtual nesting”) approach requires costly changes to inventory and needs ODFs and buckets to be mapped off-line. The author also presents major network optimization methods: Deterministic Linear Programming (LP), Dynamic Programming (DP), Nested Probabilistic Network Convergence (MIT).

Bid-price

[McGill and Van Ryzin (1999)] present how Bid-price methods were dominating in the aviation industry, also for being simple in comparison to booking control and nesting, as pointed out by Vinod (1995). According to [Temath (2010)], however, the use of bid prices implies the need of a regular update of a set of information over the booking horizon. They were initially developed by Smith and Penn (1988), Simpson (1989) [Williamson (1992)]. [Talluri and Van Ryzin (2004b)] described the displacement adjustment virtual nesting (DAVN) and another approach providing an appropriate bid price for each number of remaining seats to be sold. Other simulation-based approaches didn’t make use of decomposition, in particular: Klein (2007), Bertsimas and de Boer (2001), van Ryzin and Vulcano (2008).

Also [Talluri and Van Ryzin (2004b)] describe the use of bid prices, computed per leg and summed up by segment, as the base to set network-based controls. Such values are incorporated into a cutoff value, leading to accept-deny decisions, related to an approximate displacement cost. The dual prices are used to compute the marginal values of incremental seats on different legs in an airline network, and are summed up per leg. Random demand is here replaced by expected demand used as constraints in the Linear Program (LP).

A vast research was performed on this approach, comprehending Bertsimas and Popescu (2003) on 'certainty equivalent control' and the application of Dynamic Pricing approach. Here the application of decomposition was aimed to mitigate the computational issues. In virtual nesting control the network problem was subdivided into to multiple leg problems.

Figure 2.5 on p. 34 from [Belobaba (2002)] summarizes main alternative approaches on RM systems. [Belobaba (2002)] presents how O&D control can be defined as the ability to respond to booking requests for different O&Ds with a different seat availability based on network revenue value. The author also outlines different approaches to O&D control, in particular: EMSR heuristic bid price (HBP), Displacement Adjusted Virtual Nesting (DAVN), Network Probabilistic Bid Price control (PROBP). In particular, DAVN uses deterministic LP for optimization while PROBP a probabilistic network model.

In aviation, based on a set of cases and approaches and the use of PODS simulator, the revenue gain related to the adoption of an O&D approach in a RM system was estimated as 1 to 2% higher than leg-based RM. In the same presentation it was underlined how an improvement in forecasting and un-truncation led to greater possibilities of revenue gain, and a lower sensitivity to the use of diverse optimization models.

2.4.2 Dependent Demand

Most of early studies, reported in Figure 2.4 on p. 30, were based on the underlying assumption that the demand for any product,

RM System Alternatives

RM System	Data and Forecasts	Optimization Model	Control Mechanism
FCYM Base	Leg/class	Leg EMSR	Leg/class Limits
Heuristic Bid Price	Leg/bucket	Leg EMSR	Bid Price for Connex only
Disp. Adjust. Virt. Nesting	ODIF	Network LP + Leg EMSR	Leg/bucket Limits
Prob. Netwk. Bid Price	ODIF	Prob. Netwk. Convergence	O-D Bid Prices

Figure 2.5: RM System Alternatives in [Belobaba (2002)], slide 4.

service or fare was independent and not affected by the availability of others. In this situation, various product offerings were fenced off. In relation to the airline industry, various brackets of fares are here provided in relation to particular restrictions that are implemented as a strategy for appealing to a specific client segment.

Independent demand assumption is still applicable in quasi-monopolistic situations, but has become unrealistic. In reality, competition increased and many companies are charging different levels of fare, the presence of low-cost airlines and transport operators changed the competition landscape, furthermore price and service comparisons are easier for a traveler due to the presence of digital travel agencies and web services, for instance.

The following part of the paragraph will present studies on price elasticity, hybrid models, choice-based models and further researches.

Price elasticity and hybrid models

The studies on price elasticity of demand try to overcome the assumption on independence of demand, upon which the early RM models were based. They are based on the observation of the change in demand following a determined variation in the price (and related set of terms and conditions) of a product or service (or its subsititutes). Price easticity is measured as the percentage of change in demand towards the current state as a response to a price variation. Therefore, it is possible to estimate the customer reaction to a certain decision from a company (elasticity), or from a competitor or related to another departure time and day (cross elasticity). The estimation of price elasticity is related to the Willingness to Pay of a customer for a certain product or service. This allows for the estimation of the probability of sell-ups and recaptures in the case of closures of the lower fares. In particular, price elasticity models aim at figuring out the optimal price so that the overall revenue is maximized, taking into consideration the percentage of customers that is possible to move to buy a higher fare or book lower value services. Normally price elasticities are computed at different times to departure, as related values tend to decrease progressively and sensibly over time.

According to [Temath (2010)], main optimization models within hybrid or customer-choice demand models are provided in Weatherford and Ratliff (2010), that work on RM models with dependent demand focused on heuristics that were particularly designed for airline booking systems. On hybrid demand, main researches are from Fiig and Isler (2004), Boyd and Kallesen (2004), Cleaz-Savoyen (2005), Belobaba and Hopperstad (2004), Reyes (2006), Walczak et al. (2010) and Fiig et al. (2010). Those are based on the transformation of demand and fares from a dependent to an independent demand model, which allows for the utilization of the optimization systems already in use. Such transformation converts all hybrid demand to equivalent yield-able demand.

In the followings some other researches of interest on this topic are outlined.

[Morlotti et al. (2017)], on the real case of easyJet flights at Amsterdam Schiphol airport, demonstrates how price elasticity of demand, measured at market and route levels, varies in relation to seasonality (i.e.: month of the year, day of the week, time of the day), route and time to departure. Also the customer mix, with particular regard to leisure and business, changes accordingly; this impacts on revenues. An estimation is provided on price elasticities, e.g. from 0.535 (Hamburg route) to 1.915 (Split route).

While most literature focus on elasticity estimation at market and route level, [Mumbower et al. (2014)] provides an estimation at flight level aimed at predicting demand responses to price variations. Such forecast can also be used as input for optimization models. Another purpose of the work is supporting airlines with recommendations to design better discount policies which consider the competitor fares.

[Cizaire and Belobaba (2013)] develops an approach to treat jointly and simultaneously the problems of price optimization and inventory allocation, the first impacting on demand volumes and mix, and the second on accepted bookings. The aim is modeling their combined effect on demand. Here booking limits and fare levels are both decision variables of the revenue optimization problem, applied to two fares and two booking intervals.

[Fiig and Belobaba (2010)] study a method to transform and integrate the traditional RM approach, valid for static to dynamic optimization, through the transformation from a discrete customer-choice to an independent demand model. With this method, classical RM approaches can be used, moreover they can be applied to diverse fares structures and in particular the less fenced or structured ones.

Choice-based network models

As described earlier, heterogeneous customers have various degrees of preferences about the services provided and this is related to the set of utility values associated to any services and their willingness to pay. As a result, diverse customers are likely to make different choices on the products being offered, even though they can be grouped in clusters with similar behaviors.

When customers are provided with distinct alternatives, they can choose the preferred service(s). In particular, whenever the preferred product or service is not available in the market, it is likely that the customers will demonstrate a certain substitution behavior. For instance, they can decide to switch for other services or declining to purchase any product that is not their preference. Such behavior causes a "network RM" phenomenon effect whilst demand results to a "choice-based RM". Choice-based models can provide capacity control for obtaining the optimal customer mix, which suits the prevailing behavior for specific markets throughout the selling horizon ([Alptekinoglu and Semple (2015)]).

According to [Temath (2010)], main researches on customer-choice models can be summarized as follows. A method for seats allocation and stochastic, dependent demand is provided in Brumelle et al. (1990). Gallego et al. (2009) apply the EMSR heuristic to choice-based demand on single-leg. Bront et al. (2009) provide a column generation algorithm for network RM and choice-based demand model. [Van Ryzin (2008)] estimate booking limits in a network approach through virtual nesting. [Talluri and Van Ryzin (2004)] provide a novel optimization model under a general demand approach.

Based on [Alptekinoglu and Semple (2015)], an acceptable approach would be able to specify a suitable demand model and properly estimate factors involved from historical booking data and past experiences. The same work also proposed a stochastic dynamic program approach. This mathematical representation demonstrates an existing relation of choice-based and network RM effects in the market. Bellman equation is utilized in this model to define the availability control problem.

[Masatlioglu and Nakajima (2013)] contemplate that a dynamic search process should be undertaken where consumer choices follow an iterative process. [Wang et al. (2016)] select a logit model after an extensive research on the impact of costs on assortment planning and prices of products. [Jagabathula and Vulcano (2017)] study on non-parametric expectation-maximization model and solution applied to joint assortment and product price optimization. Lastly, [Jagabathula and Rusmevichientong (2016)] provides a two-stage nonparametric model where customers are represented as partial orders of preference. [Shen and Su (2007)] provide extensive information regarding customer behavior within broader RM and auctions. These studies provide a focus on inter-temporal substitution and on customer buying behavior in a real business environment.

The following part presents an overview on main other models proposed, with a focus on the multi-product substitution topic and contemporary developments occurring on choice-based RM approaches that have been developed after [Shen and Su (2007)] studies.

Parametric models

Parametric models are based on random utility theory. Consumers associate a certain utility with each alternative choice and evaluate a set of alternatives to select the one that best maximizes their own utilities. The presence of a random factor explains the cases in which customers seem not to choose according to their rational preferences.

Multinomial logit (MNL)

The MNL model was first introduced by [Talluri and Van Ryzin (2004)] as a basis for revenue management. The overlying assumption here is that the utility component is independent and can be identically distributed for random variables based on a Gumbel distribution. This approach is based in a strong assumption of independence among alternatives provided for two sets. MNL is a simple model that is powerful and flexible; its disadvantage consists on the difficulty involved in dealing with the possibilities of no-purchase, that are normally unobservable. [Vulcano et al. (2010)] has successfully applied a similar approach, that is also consistent to [Gallego et al. (2015)]. Here, the Expectation-Maximization (EM) model forms the basis of an estimation approach that was earlier postulated by [Vulcano et al. (2012)] and can be applied to non homogeneous Poisson arrival process. On his side, [Newman et al. (2014)] develop a method based on marginal-log-likelihood functions, related to a similar EM approach.

Finite-mixture logit

Under this approach it is possible to combine a finite number of MNL models in a unique choice model that will be referred to as a latent class model. With mild regularity conditions, a discrete choice model can be derived from a random utility maximization, that provides probabilities that are arbitrarily assumed by the MNL model ([McFadden and Train (2000)]). The finite-mixture logit model provides a choice behavior that outperforms MNL, but with a more challenging estimation parameter.

Nested Logit (NL)

In this model the alternatives are aggregated into nests, using a toe hold IIA approach for each one, then spread across all nests. A single nest will represent a set of substitutes of the product or service that a customer would choose. However, customers may choose an alternative out of the nest. According to [Train (2009)], any derivation and estimation should be based on the observation of sequences of decisions. On the other hand, NL parameter estimation in the context of maximum likelihood is supported ([Anderson et al. (2012)]).

Markov chain model

[Blanchet et al. (2016)] develop a choice model where customer choice can be denoted by a Markov chain. According to [Blanchet et al. (2016)] this model is reliable for approximating a discrete random utility under the conditions of a mild assumption. [Berbeglia (2016)] indicates that Markov chain model can be classified as a discrete choice model because of random utility. Recently, [Simsek and Topaloglu (2017)] have proposed an EM approach that can be used to estimate the Markov chain model functions.

Exponomial models

Exponomial models incorporate a linear function of exponential terms. [Alptekinoglu and Semple (2015)] proposed a new choice model that combines negative skewed distributions of consumer utilities. This approach is different from the MNL and NL models presented above, which assume customer willingness to pay distribution is a positively skewed factor. Here, choice probabilities are expressed as a linear combination of the exponential terms. This method can be applied where customers are likely to choose to pay more if they are properly informed about the products and prices.

Non—parametric models

They are founded on the assumption that choice behavior can be represented in a functional form. The approach is very general and consists of several common choice models like MNL. Those models can be regarded as special cases of a rank-based model ([Mahajan and van Ryzin (2001)]). The possibility of increasing types of potential customers provides a challenge to the models. This lies in specifying the probability mass function that is linked with the process of identifying customer types. According to [Van Ryzin (2015)], the difference of model specifications and estimation errors in non-parametric models can be observed in rank-based models. The main disadvantage provided by non-parametric models is that they cannot make predictions for completely new products or services, that have not been treated earlier.

Multi—stage choice models

Under this approach the choice process is represented as a single stage, where a certain market is characterized by the presence of a set of alternatives and the associated probabilities, for any purchase. These consideration sets can be either observable or unobservable; this represents a challenge in determining the relevant consideration sets. [Masatlioglu and Nakajima (2013)] contemplate that a dynamic search process should be undertaken where consumer sets are determined iteratively. [Wang and Sahin (2014)] select a logit model after an extensive research on the impact of costs on planning assortments and prices of products. Contrarily, [Jagabathula and Rusmevichientong (2016)] state that a non-parametric EM model is ideal for joint assortment and product price optimization. Lastly, [Jagabathula and Vulcano (2017)] define a two-stage nonparametric model where customers are represented with partial orders of preference.

2.5 Pricing and Price Optimization

Pricing can be considered one of the main aspects considered for customer choice and, on the other side, one of the key success factors of a company. In general, price decisions have effects on the Company that made them but also on the overall market, and both in the short and in the medium term; therefore, Companies are extremely careful on fare decisions. Furthermore, the set of available prices and especially the lowest ones build a 'price anchor' in the mindset of the customer base. While acceptance of a lower price from the customer is normally easy, raise prices would be much harder and risky (in terms of customer loss). This is also why, in the long term, the raise of competition may lead to spiraling down of prices.

At the tactic level, the definition of price points and fares as well as their availability management should find the sweet spot that maximizes revenue. For instance, aggressive discounts and their large availabilty may lead to profit loss related to the fact that the customer value extraction does not reflect their willingness to pay. Contrarily, a price which is too high, or the systematic unavailability of lower fares, can discourage bookings and encourage substitution behavior, leading to even higher profit losses.

On the other side, price information is easily available to customers, e.g. through a web search, therefore transparency and comparability are high and pricing actions from a company impact the overall market shortly. Very often the Companies know the competitor prices in real time using tools which collect, process, aggregate and present them. The collection can be either done directly from the Companies, as mostly done in aviation, or through 'webscrapers' for price tracking through the Companies websites, Travel Agencies or other public sources. Those data are checked regularly from the RM team and are often a direct input for the RM System (RMS), which can be 'market aware' and react to competition actions and/or 'customer-centric' and adjust to customer preferences accordingly.

In addition, each industry has some pricing specificities. For instance, in aviation the final prices are often a result of a complex sum of addenda. Railway operators commonly adopt a fare family (clustered) structure, in most cases well differentiated per O&D, and fares levels or intervals can be negotiated with the States. In other industries price differentiation is built upon other dimensions (e.g. upgrades, ancillary services, optionals, others) and bundling/unbundling techniques are commonly applied.

Furthermore, traditional carriers commonly have a fare family structure with relatively stable price levels, while in low cost carries, for instance:

(i) the pricing structure is fairly simple, with few price levels;
(ii) rates vary dynamically over time;
(iii) ancillary services account for a relevant percentage of revenues.

In certain cases a price optimizer tool runs separately, while in others, namely in dynamic pricing approaches, price optimization is a part of the overall optimization and provides input to the main RM system every time a reoptimization takes place. In both cases, the process is repeated over time during the booking horizon to provide input to RM optimization, but only in the second one the temporal dimension is explicitly taken in consideration.

The dynamic pricing problem is considered computationally complex; an example is [JDA (2009)] approach, that will be presented in 4.2.3. Often, in practice, although truly dynamic pricing is not implemented, the booking system is able to perform *de facto* a dynamic pricing policy, reoptimizing several times at predefined moments in the booking horizon and displaying to the customer only the best available price of each fare basket.

2.6 Role of the RM team

This part describes the main activities of a RM team and its corporate environment and processes. The level of experience and expertise showcased by RM teams are core in the organization revenue strategies and tactics and through the use of advanced Analytics and Operations Research applications the RM Department contributes to the implementation of the company's directions. It emerges how the overall optimization is only partially impacted by the RMS decisions, as they are mediated and influenced by human decisions, as well as the company's systems and processes. At the end, the success of a RM system will be achieved where a revenue management culture is strongly embedded. Main source of this subsection is [Legohérel et al. 2010]; few parts recall also [Vinod (2006)], while others report some extract from pretty detailed job posts[1].

A RMS can either work on automatic way or provide a basis for making a system decision based on the parameters and business rules. It normally does a combination of the two. Routine decisions are mostly performed automatically by the system, following the rules and parameters set by the users. On the other side, early detection and manual management of exceptions is necessary as they represent an important revenue share.

RM is an useful tool for analysts to take decisions and achieve company objectives. The level and degree of automation used in a RMS is increasing continuously. Such automation aims at managing routine and spotting exceptions in advance, so that they can be checked or managed directly by the RM team, for instance setting

[1] e.g.: `www.xotels.com/en/hotel-revenue-manager-job-description`, retrieved on October 20^{th}, 2018

limitations to the system choices. However, the role of users is still of paramount importance and has a great impact to the system performance. By fact, users interact with the system on a daily basis through the user interface, establishing parameters and rules that will promote the implementation and realization of the company's strategies and objectives. RM requires highly skilled personnel who are tasked with managing both routine activities and exceptions.

The acceptance of RMS decisions from the analysts and users can be related to their understandability. Therefore, for instance, complexity perceived shouldn't be excessive or discouraging. In worst cases, this could lead to misuse of the system or components, or to overriding of the RMS decisions.

The RM team has the important role to ensure maximization of revenue generated through sale of products and services. Within the organization, the RM unit acts as an anchor to optimization, management and operational functions deal dealing with market awareness, forecasting activities, budgetary reporting, price and quotations decisions. RM influence decisions for opening or closing fares and activities required for distribution network management. To do so, quite often they implement management strategies and processes with a direct contact with the company's clientele.

The Revenue Manger has a central role within the corporate and is perceived as a key person in the creation of strategic plans and within the decision making apparatus. The RM Department is usually positioned at high level in the Company's organization chart, and may be even higher than other functions of the company, such as Marketing, Finance, Human Resources and Sales Departments. In certain cases, the RM team reports directly to senior management, even though financial decisions are normally reserved to top management or executive committees, depending on the governance structure of the company. Furthermore, revenue managers are supposed to provide financial reporting reflecting the Company's earnings, market share and value of the products and services; they also analyze finance and business operations as a basis for making future estimates of revenues in consideration to general economic and market conditions. This provides senior management the guidance for undertaking important business decisions. The team of analysts, on the other side, is mainly concerned with the operational responsibilities and decisions.

Within this framework, and recalling the difference between RM and YM discussed earlier, Yield Managers are considered like “technicians” at one point in the life-cycle of the organization and might come, for instance, from the Data analysis or Pricing functions. Yield managers incorporate the responsibility of implementing strategies and are tasked with recommending commercial options to optimize the whole revenue system of a company. The objective here is generally to reach a suitable balance of occupancy rates and revenue generated from sales. The scope of Yield Managers is also inclusive of monitoring performances of YMS tools as well as commercial roles and distribution and partnership activities. Also, they shall work consistently to Sales and Reservations departments as a single team, promote distributions, reduce costs, prepare budgets, undertake periodical reviews and increase operational activities of the organization. This is achieved through implementing revenue administration procedures and practices. Maintaining a daily, weekly and monthly reporting to showcase progress of the system is another task.

[Legohérel et al. 2010] provides an example from IdTGV company, a subsidiary of SNCF that works in the passenger rail transport sector. Its RM unit has a direct reporting mandate to senior management. The persons in charge of RM and Pricing have the responsibility of promoting the general objectives of the company as well as subsidiaries, if any. The mangers responsible for pricing, planning and transport should collaborate with the other Departments of the Company and, in some cases, with other external analysts to facilitate the implementation of the RM strategy.

Summarizing, a RM team should ensure: revenue maximization; revenue management and distribution strategies that are properly anchored to the routine day to day activities; routine analysis and adjustment activities for reporting. Also, performing competitive benchmarks will be required to understand the market trends and boost strategies for targeting success. Another task is to maintain an annual rolling calendar of demand forecast. Undertaking a weekly dynamic forecast of expected results is a task that will show whether the company is performing well. RM team regularly checks input and quality of data and customer segmentation. A performance review will form the basis of implementing a tactical action for achieving the corporate objectives. Furthermore, the managers are tasked with making competitor and business analysis, market and distribution modeling and assessments, demand mix control, pricing control and optimization following the best practice standards.

Finally, a critical area to be considered relates to RM system maintenance related to routine management issues. Here, a proper technology and process management is required for adequate data and revenue collection. These activities require well-rounded managers and analysts, with both extensive accounting to commercial and analytical skills and technological background.

3 A Comprehensive Methodological Approach to Assess RM Performance

This section focuses on how to assess RM performances and provide possible corrective actions. This should be done through a multi-dimensional approach, encompassing different points of observation and methodologies, due to the intrinsic complexity of RM. By fact, transport revenue management is a combination of processes, people and effective technologies that are combined to provide revenue maximization. RM works on the demand side to: improve customer value extraction, balance demand and supply, influence customer decisions by pricing actions, incentives and promotions. On the other hand, it applies to supply factors involving capacity imbalances, demand in traditional markets served and creation of new ones. Due to its transversal nature in incorporating many functions, it should be undertaken within the corporate level. Integration of business processes will be required to ensure that it acts as the basis of decision making for the organization. Performance indicators should be monitored at corporate silos because it has important implications to many stakeholders.

Revenue management has a central role in promoting the performance of the organization, both in the short and the long-term. The short-term objective is always profit maximization based on a particular schedule and fare structure. In the long-term, revenue management can provide huge financial impacts through: marketing plan, sales practices, the existing distribution channels, ancillary services, frequent fliers programs as well as advertising strategies. In general, whereas the decision to implement a RM system is fundamental for an organization in increasing revenues, at the same time a high level of revenue collection is obtained through management of topics comprehending: overbooking controls, discount mix optimization, pricing and groups controls.

[Vinod (2006)] provides an extensive description of best practices to measure performances of a RMS in aviation. Main parts of the followings are taken from his work. In his view, it is important to proceed with a certain order to ensure a successful implementation of RM strategy, system and processes:

(i) Firstly, the need for RM adoption should be perceived from top management;
(ii) Secondly, an evaluation of alternative RM systems should be undertaken to determine their suitability;
(iii) Thirdly, the Company should proceed with the formulation of RM strategies within the steady state operating business environment;
(iv) Lastly, a RM system should be evaluated from its inception to a reliable steady state.

Commitment from senior management in making the decision to implement a RM system and the following steps is fundamental for the achievement of targeted goals and objectives. Undertaking RM strategy requires the establishment of a proper organization structure and ecosystem, with dedicated skilled people. A paradigm shift is needed to properly manage demand, overcoming forecasting or other issues that may arise and providing timely corrective actions for achieving success. A continuous support and monitoring at all levels is also necessary to prevent, among the others, the following issues:

(i) lack of accurate forecast of demand for products and services;
(ii) suboptimal allocation of capacity, resulting in poor customer mix and revenue realization;

(iii) suboptimal pricing decisions, determining revenue and passengers loss;
(iv) increasing levels of spoilage, e.g. a discount allocation spoilage can be spotted in presence of empty seats occurring from premature closure of discount classes in presence of high demand;
(v) ineffective strategies for retaining customers with the risk of revenue dilution;
(vi) failure in maintaining quality at acceptable levels;
(vii) insufficient ability to adjust the offer for fitting demand.

In a steady state RM has a pivotal role of ensuring that the plan is finally executed. An in-depth costs analysis, which cannot be incorporated into RM systems, should be performed; activities for monitoring and ensuring delivery of cost targets should be achieved through adopting an optimized cost approach. Also, dedicated management of demand in case of special events is key and should be supported by the use of triggers and alerts from RM performance indices. Furthermore, there is a need for making a proper definition of RM performance indicators at every level, performing continuous monitoring and offering corrective actions whenever a problem is detected.

A larger transport organization is expected to obtain a high payback to the initial investment; [Vinod (2006)] reports that the time of payback of a RM system can be estimated in one year. In addition, incremental costs involved in undertaking revenue management are less heavy for a large organization, in proportion, due to their scalability. However, as explained earlier, focusing on revenue management system alone will not guarantee an increase in the revenue performance unless other factors are involved, such as supporting top level management, proper processes and tools and skilled resources. Commitment to establishing a working RM system with clear roles, responsibilities and accountability is paramount. Similarly, continuous monitoring of pre-departure and post-departure indicators can be used to obtain feedback and provide corrective actions.

According to [Vinod (2006)], during the decisional phase, the senior management will be required to define business needs and challenges that are specific for increasing the level of revenue maximization. A simulation model can be used for this stage, incorporating three major components: simulation environment, revenue management, a model for passenger decision. It can use passenger historical data, augmented with spilled passengers. Others use Monte Carlo simulation techniques. Furthermore, it is important that the passenger decision module relies on a well defined inventory control structure that will be backed through an established transport reservation system.

The outcome of the simulation model can be measured through performance indicators, including: network revenue, individual passenger revenue, load factor, spill, spoilage, stifling[1]. In particular, we will use the following definitions: *high spill*: selling seats at low fares while there was a part of the accepted demand willing to pay a higher or full price; *stifling*: not selling all seats available, in presence of higher demand for discounted prices; *spoilage*: selling a suboptimal mix of segments and categories, in presence of under capacity on at least one leg.

In addition, the simulation model can be used to test enhancements and algorithmic modifications that are performed over time, to ensure realization of maximum returns for investments. A dedicated part on simulations will follow in Chapter 3.3.

The organization needs to validate and monitor on continuous basis its identified performance indicators. They can be broadly divided into pre-departure and post-departure measures. Both are further described in Chapter 3.1.

On the other side, the Monitoring information serves to fulfill two distinct purposes. Firstly, it is applied as a yardstick to make comparisons of progressive measures on daily, monthly and yearly

[1] The concepts of Spill and associated phenomena, well known to the YM practitioners, are explained in the unpublished technical paper Spill Model Technical paper at `http://cyberswans.com/AirlineIndustryPubs/Spill_Model/Spill_Model_Technical_paper.docon` , September 20th, 2018

basis towards the previous period. Secondly, it is adopted to detect weaknesses occurring in the RM system and confer corrective measures that need to be implemented immediately to maximize profits. A detailed part on this topic will follow in Chapter 3.2.

Based on [Talluri and Van Ryzin (2004b)] and [Vinod (2006)], main related benefits include: to obtain the maximum benefits from an operating RMS, monitoring and measuring continuously the performances, fine-tuning the working system not only at aggregated level but also in its components (Curry, 1992), measuring results against previous periods and over time to refine the analysis and measure the effect of RMS decisions (Polt, 2001). According to [Temath (2010)], main approaches on performance measurement of RM performances are described in [Chiang et al. (2007)]. As reported in [Temath (2010)] and [Vinod (2006)] and displayed in Figure 3.1 on p. 54 and Figure 3.2 on p. 55, assessing RM performances encompasses the integrated use of:

(i) basic indicators calculated from inventory and booking data, such as Revenue, Passengers, measures of Load Factor or LF[2], Revenue per Available Seat Km or RASK[3], and others[4];
(ii) other indicators which foresee the comparison of performances over diverse time intervals;
(iii) simulation, mostly used to compare new RM algorithms before implementation;

[2] Load Factor or LF can be defined as the percentage of booked seats on the total offered. It can be computed as the sum of booked seats, each multiplied by the length of the trip (i.e. kilometers traveled), divided by the sum of the offered seats, each multiplied by the total length of the route (i.e. kilometers offered). The numerator is a measure of the demand and each passenger is multiplied by the kilometers of his own reservation, while the denominator is a measure of the offer calculated as the number of seats multiplied by the total length of the train route expressed in kilometers. It is mostly used in transport industry.

[3] RASK, or Revenue per Available SeatKm, a measure which compares revenue and supply. It is computed as the as a ratio of revenue gained to the offer of seats multiplied, expressed in SeatKm

[4] in all the thesis we measure distances in Km following European notation, but Miles could be used indifferently

(iv) monitoring systems based on the Revenue Opportunity (RO) Model, that will be the subject of Subsection 3.2.

In the next paragraphs such tools and measures will be further discussed.

3.1 Pre and Post Departure Indicators

This part is dedicated to provide an overview on main indicators of performance used in RM. As anticipated previously, a first classification of RM performance indicators which best fit the transport industry is between pre and post departure indicators. The second can be further subdivided in standard indicators and measures stemming from the Revenue Opportunity estimation. Again, most parts are taken from [Vinod (2006)] that focuses on aviation industry.

Pre-departure performance indicators

Pre-departure indicators are key to provide a structured basis to offer corrective actions where problems have been identified, minimizing latency and revenue loss. Pre-departure performance analysis aims at detecting, for instance, abnormal situations as well as new demand trends, and provide triggers and alerts to the RM team during the booking horizon, so that there is still room for corrective actions.

The importance of pre-departure indicators is also related to the fact that forecast accuracy is key for the quality of models decisions as well as for corporate performance: where forecasts are accurate, performance issues can be spotted and quantified precisely. As described earlier, one of main issues of forecasting is related to the aggregation level at which they are provided: a more aggregated level can ensure more reliability, but a disaggregated level provides atomic informations able to support punctual decisions, and, most of all, properly feed the RM models. For instance, forecasts can be accurate at the macro level, but may become inaccurate at itinerary and class levels and even misleading for the purposes of seat inventory control from RM. Consequently, this results in the transmission of errors to sales and revenue operations, determining in suboptimality and, in general, a failure to achieve targets planned. For instance, this is the case of the sale of too many discounted seats, leading to the spillage of high-pay passengers.

Finally, given the high volatility of competitive transport environment it is difficult to have accurate, stable and reliable forecasts. This can be mitigated through progressive observation and evaluation of demand to account for variability and setting proper corrective actions and business levers to adjust to the plan, balancing current and future goals. Latency can be defined as the interval of time between the arrival of input information and the availability of the output (here its implementation in effective RM decisions). It is a problem that is created by slow propagation of information from excessive cycle times within the broader decision-making structures. A RM will provide transparency in a company and leverage revenue performance indicators to assist revenue plans that minimize latency in the broader decision-making processes.

[Vinod (2006)] presents several key measures that can be involved in the continuous monitoring of processes and models performance over time; a key distinction is presented between Corporate and Model measures. Corporate measures include: booked and expected load factor, expected yield and revenue, expected Revenue per Available Seat Kilometer (RASK), spill and closing rates. On the other side there are the Model measures, with a focus on forecast errors. Figure 3.1 on p. 54 presents main pre-departure indicators used in aviation for continuous monitoring of RM pre-departure performances, as presented by [Vinod (2006)].

Post-departure performance indicators

Historical performance of a transport service, on the other side, is calculated following the departure. In this regard, post-departure performance indicators can be subdivided in standard performance

Pre-Departure Performance Measures	
Corporate measures	**Description**
Booked load factor	The ratio of onboard traffic to available seats expressed as a percentage. An alternate definition is the ratio of revenue passenger miles to the available seat miles expressed as a percentage.
Expected load factor	The expected load factor of a flight at departure based on current bookings, forecast of remaining demand and inventory controls.
Expected revenue	Expected revenue of a flight at departure load based on current bookings, forecast of remaining demand and inventory controls.
Expected yield	Expected passenger revenue per revenue passenger mile/kilometers.
Expected revenue per available seat miles/kilometers	The ratio of expected passenger revenue to available seat miles/kilometers.
Spill rate by class	Ratio of spilled passengers to unconstrained demand to date for a future departure. This statistic must be computed by leg class or service class depending on inventory control method. If the spill rates are not ascending from the highest valued class to the lowest, it identifies a problem with how discount allocations are set.
Closing rate	Probability that demand for a booking class exceeds the available seats in the class (pre-departure closing rates).
Model measures	**Description**
Forecast errors	Errors associated with demand forecasting, cancellation rate forecasting by reading day interval. Common measures include mean absolute deviation, standard error, bias and mean squared error.

Figure 3.1: Main pre-departure Indicators. From [Vinod (2006)], page 16.

and revenue opportunity measures. The first ones can be retrieved from simple historical or RM data, the second are the results of Monitoring models which reconstruct the Revenue Opportunity. For the latter, a dedicated part in 3.2 will provide further details.

The standard post departure measures can be provided on a certain aggregation level, for instance per period of time or geographic areas, according to the information set which is needed for the specific purpose. Among the Corporate measures are considered: Load Factor (LF), Yield, RASK, Market Share (MS), Denied boarding and related costs, Spoilage, LF on closed flights, Closing rate, Employee efficiency. On the other side the Model measures, as in pre-departure indicators, focus on forecast errors. Figure 3.2 on p. 55 summarizes the main standard indicators used in aviation for post-departure performance assessment, as presented by [Vinod (2006)].

Post-Departure Performance Measures

Corporate measures	Description
Load factor	The ratio of onboard traffic to available seats expressed as a percentage. An alternate definition is the ratio of revenue passenger miles to the available seat miles expressed as a percentage.
Yield	Passenger revenue per revenue passenger mile/kilometers.
Revenue per available seat miles/kilometers	This is considered the single most important measure and the ratio of passenger revenue to available seat miles/kilometers.
Market share	Represents an estimate of the proportion of total traffic in a market. Can be estimated from market information data tapes and passenger shopping data.
Oversale (denied boarding) rate	Calculated separately for voluntary and involuntary denied boardings. It is the ratio of number of denied boardings to the number of passengers boarded expressed in denied boardings per 10,000 passengers boarded.
Spoilage	Represents the number of empty seats on closed flights.
Spoilage rate	Ratio of number of spoiled seats to the number of passengers boarded expressed as a percentage.
Over-sale cost	Cost of customers who were denied boarding. Should be tracked by airport. Components of denied boarding costs are voucher costs, meals, ground transportation and goodwill. Expressed in unit of currency per person by airport.
Load factor on closed flights	This is the correction factor for incorrect overbooking. Measurement is based on open/close status throughout the flight by booking class and calculated as a weighted average by flight leg and base compartment.
Closing rate	Probability that demand for a booking class exceeds the available seats in the class.
Employee productivity ■ Average number of employees ■ Seat capacity per employee (seat miles/kilometers) ■ Passenger load per employees (ton miles/kilometers) ■ Revenue per employee	Standard measures of employee productivity and usually reported year over year. Productivity statistics are important since employee compensation constitutes a significant part of the operating costs.
Model measures	**Description**
Forecast errors	Errors associated with demand forecasting, cancellation rate forecasting, boarding rate forecasting. Common measures include mean absolute deviation, standard error, bias, weighted mean absolute percent error and mean squared error.

Figure 3.2: Main Post-departure standard indicators. From [Vinod (2006)], page 16.

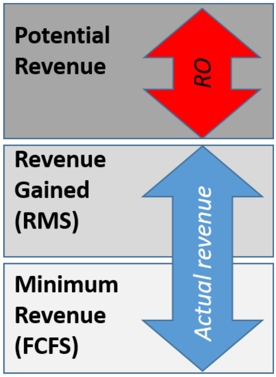

Figure 3.3: Revenue Opportunity (RO) main concepts and components

3.2 Approaches to Monitoring. The Revenue Opportunity (RO) Model

One of the main challenges in assessing revenue performances, realized from a working revenue management system, and in determining related incremental revenues, consists in isolating RM results from external influences; these external factors include pricing actions, business and competitive environment and in particular the possibility of low-cost competitors.

Measuring direct results of RM and supporting the definition of corrective actions are the main purposes of Monitoring tools, mostly based on Revenue Opportunity Estimation. They are the subject of this subsection.

The Revenue Opportunity (RO) can be described as the potential revenue achievable, given the potential demand and optimally managing the available capacity. According to [Vinod (2006)], RO can be estimated determining the optimal mix from the exact count of boarded passengers. Figure 3.3 on p. 56 displays the concept of RO and its main components.

Post-departure measures related to RO determine the level of incremental revenue that is obtained as a percentage of the total RO realized through perfect RM system decisions. The potential revenue is subject to changes because of the influence of external factors; conversely, the proportion realized from the organization is consistent to revenue management performance. Therefore, an increase in the revenue opportunity measures observed over a period of time is an indication of improved performance. As explained in [Vinod (2006)], the RO can also be used for other purposes and specifically to monitor RM in its steady state, as a term of comparison for achieved results to provide punctual corrective actions.

[Talluri and Van Ryzin (2004b)] outline main methods for RO estimation, comprehending:

(i) setting an upper bound to achievable revenue, to be compared to actual revenue or the FCFS representing the situation without RM;
(ii) using simulation techniques and tools.

The RO estimation can be used to quantify the performance of a new RM implementation: in this case, to assess its net effect, the part of the RO achieved through the use of RM is usually compared to the 'First Come, First Served' (FCFS) case, assuming it represents the initial situation without RM. As anticipated earlier, the Monitoring tools based on the estimation of the Revenue Opportunity are frequently used not only for the continuous measurement of RM system performances, but also to isolate them from external factors which may impact on revenues. In particular, seasonality and special events or circumstances causing demand shocks or capacity variations are quite common in transport industry, and services (e.g. flights) with higher demand are mostly related to higher revenues and performances. The RO is able to take this into account through the use of FCFS revenue as baseline to measure potential performance improvements.

[Temath (2010)] in his dissertation presents an overview on Network-based RO Models (ROM) approaches for Airline RM; below some major works are reported. First researches from Kempka

(1991) and [Smith et al. (1992)] explore the opportunity to apply "perfect" controls from RM and calculate the optimal revenue; initial applications were for single legs. Other works focus on the estimation of the baseline or "no RM" values; Daudel and Vialle (1992) suggest the comparison to actual revenues, while Rannou and Melli (2003) and Smith et al. (1992) use FCFS revenue to compute the baseline situation without RM. Researches from Adler (1993), Chandler and Ja (2007), [Talluri and Van Ryzin (2004b)], Temath et al. (2009) and [Vinod (2006)] aim at extending ROM from a leg to a network-based model. Adler (1993) and Chandler and Ja (2007) research on ROM model-related limitations and Temath et al. (2009) on the possibilities to overcome them through the use of dependent demand models. Curry (1992) explores the opportunities achievable in relation to the mitigation of forecast errors taking place in the course of the booking horizon. Polt (2001) researches on the validity of ROM in presence of errors in the unconstrained demand forecasts, stating that their effect is negligible and muffled at aggregated level. Adler (1993) outlines how the goodness of the ROM is related to the unconstrained forecast; they both investigate on relations between ROM and "classical" performance measures. Researches on the use of ROM to test the performances of single components of RM have been performed by Smith et al. (1992), Polt (2001) and Chandler and Ja (2007). The latter work in particular proposes to subdivide ROM outcomes into spoilage and dilution and perform focused analysis for single legs or markets. Other researches from Mak (1992) and Dar (2006) focus on the use of simulation to compare different RM systems or components, such as optimization approaches.

In his research, [Temath (2010)] observes how the main errors, which can affect ROM performance, can be either model-related or data-related. In particular, accuracy flaws of ROM estimation appear to be affected by the quality of input data, and mostly related to inaccurate unconstrained forecast (that affect estimations of potential revenue) and FCFS case estimates. However, no further researches have been done on the relation of un-constraining inaccuracy with ROM robustness. As for model-related errors, it is found out the following. Unconstrained forecast may be flawed by the aggregation of reservations data at points of observations (data collection point, or DCP) within the booking horizon and the assumption that demand comes in Low-Before-High (LBH) order. In general, in this view the Revenue Opportunity Model shouldn't be necessarily following the same approach of RM, but has to be as accurate and realistic as possible. For this reason, it needs to be updated progressively with new findings and developments in RMS, such as dependent demand structures. It is also pointed out the lack of literature on this topic. Finally, it is stated that the use of leg-based ROM leads to invalid and misleading outcomes in airlines with network-based RM.

Summarizing, [Temath (2010)] researches on the use of ROM for performance measurements in network-based RMS with independent and dependent demand, for a large network airline. For independent demand RM models, computational results display how the data-related errors have much more impact on ROM validity than model-related errors. For dependent demand RM models, a decrease in robustness is observed in relation to a lower degree of unconstrained forecast accuracy. For both cases, however, ROM demonstrates to be robust and suitable enough for real life applications.

[Temath (2010)] also investigates on the possible application of ROM to subproblems, such as part of the network or single parts of a RM System. ROM demonstrates to be robust enough also for leg-level analysis, nevertheless in this case the forecast errors tend to increase, especially in the dependent demand case. Also the ROM application to RMS components deliver encouraging results, in particular for upgrading and overbooking. In his work further researches are suggested on the utilization of ROM as training tool for revenue managers, as well as for pre-departure measures.

3.3 Simulation Models

It was previously explained how simulators can be used for diverse purposes, such as the evaluation of a whole RM system (prototype, newly implemented system or already in its steady state), or its single components. Other purposes can embrace, for instance, testing diverse models for demand forecast, optimization and booking controls based on strategic goals to allow for a comparison.

In general, the performance of a RM system can be evaluated based on decision theory as depicted by [Granger and Pesaran (2000)], that offers an approach that does not conclusively determine whether a single part of a RM system plays a decisive factor to providing success. Contrarily, a simulation-based approach is focused on evaluating a complete system with its ability to maximize revenue. According to Mayer (1976), provided with correct forecasts, a RM system can work optimally. Unfortunately, this may not be useful in real practice as it is based on stringent assumptions; here, the ones considered in such mathematical proofs can be simplified to reflect a low fare demand preceding a high fare demand (LBH). To overcome these shortcomings, an ideal approach for evaluating and comparing RM systems is through focusing on the outcome. This was also supported by [Talluri and Van Ryzin (2004)]: the actual results of a RM system shall be evaluated based on the number of bookings, monetary implication and collected revenue. A simulation approach using a broad-based knowledge on demand behavior can be properly applied into a RM system. In particular, through a simulation approach it is possible to estimate booking requests requests that were turned down for any reasons; this is the basis for untruncating demand forecasts. The choice of a simulation model should consider the consistency of the levels of of complexity of a RM system compared to the one it is supposed to manage in the real world.

The use of simulators provides several advantages. The main one is determined by the possibility to analyze outcomes and replicate processes, limiting the influences of real world dynamics on results.

By fact, in reality results can be altered due to external factors, such as economic trends, and this may lead to false demonstrations of performance. As such, artificially generated demand can be used to apply a certain level of volatility in the process, as well as modeling demand to encompass flexible choices that allow customers to minimize costs through selecting the most affordable booking class. Since data is provided on bookings, restrictions and price acceptance, it is possible to make a comparison of actual bookings and potential demand through the use of simulation. Another important advantage of simulations is the possibility to repeat processes and tests, while at the same time blocking single components; both are key concepts of Experimental design, as will be detailed in 3.4. Finally, long term consequences of actions and variations can be properly analyzed.

Simulation can also be used to compare different RM systems and approaches. A reservation approach can be used for this purpose, with the application of different inventory controls to evaluate the booking requests recorded within a period; the results obtained in one approach can be compared with the other methods used in the evaluation. Furthermore, it is possible to evaluate single forecasting components of RM systems though comparing forecast performances of different systems in diverse periods. Thirdly, a straight-forward approach can be applied to evaluate predicted demand and actual recorded bookings. Forecasting requires the transformation of historical information and recognition of patterns [Armstrong and Meissner (2010)]; this facilitates the understanding of historical demand and making predictions for the future as a time series component. Unfortunately, approaches for predicting demand based on time series don't provide constrained data; instead, a simulation approach included in the processes of evaluating a time series will provide un-constrained data from inventory. In this model, it is also possible to ideally set capacities to infinity and overcome the issue of demand truncation, as stated in [Armstrong and Meissner (2010)].

The use of simulation to measure and compare performances of RM systems and components is widely spread. According to [Temath (2010)], main works on the use of simulation for evaluation of single RM components are from Weatherford (2004b, 2002, 2004a), Belobaba and Weatherford (1996), Weatherford and Belobaba (2002) and Weatherford and Polt (2002). Oliveira (2003) and Eguchi and Belobaba (2004) research on applications to evaluate the impacts of the usage of diverse RM approaches on specific markets. [Cleophas (2009)] and [Cleophas et al. (2009b)] focus on the evaluation of forecast performances. Basumallick and Singh (2009) research on the applications for strategic decisions, while Gerlach and Frank (2010) introduce the use of simulations for experts training (ReMaTE simulator), as it evidences the effect of revenue manager choices on the outcomes and allows to simulate competitive environments. Williams (1995), Jain and Bowman (2005) as well as Lieberman and Raskin (2005) research on the benefits of comparisons between time periods, developing approaches that are suitable for the effects of RM decisions at a global level, but aren't applicable for continuous monitoring. Finally, [Talluri et al. (2010)] develop a novel methodology, named 'sandbox testing', to assess the revenue potential of a new RM approach through parallel live testing of new RM prototypes against the incumbent algorithm on current data. This is an important component of the theoretical background for the live test described in Chapter 6.

[Temath (2010)] provides an overview on main researches on the simulation topic and key factors for a successful RM. Anderson and Blair (2002, 2004) deal with 'performance monitor' of RMS outcomes and focus on revenue opportunities which are lost, subdivided in their components. Other works from [Vinod (2006)], Polt (2001) and [Talluri and Van Ryzin (2004b)] focus on performance indicators computed from available RM data, split in pre and post departure measures by [Vinod (2006)] as detailed in Chapter 3.1. Load factor and RASK are identified as the most important ones, while Phillips (2005) points out how the RASK incorporates values both on revenues and seats offered. Armstrong (2001) researches on measures of forecast accuracy and RM performance. Lieberman

(1991), Skugge (2004) and Gerlach and Frank (2010) instead focus on the importance of simulation-based training as key factor for assessing RM outcomes as well as organizational requirements, also developed in Wishlinski (2006). In particular, Lieberman (2003) lists the criteria for successful RM, which include: performance measurement, support strategy and business processes, management support, accountability, integration, awareness of RM limits, career progressions. Finally, Crystal (2007) points out how RM success is related to technical and organizational capabilities.

PODS Experience at MIT

It has been outlined previously how the use of simulation is quite common to analyze the performances related to RM outcomes and approaches. One of the most common simulator used for aviation is the Passenger Origin/Destination Simulator (PODS); many researches have been done with the support of PODS realistic environment. [Temath (2010)] mentions in particular Skwarek (1996), Reyes (2006), Carrier (2003), Cleaz-Savoyen (2005), Gorin (2000), Zickus (1998) and Gorin and Belobaba (2004); such researches focused either on forecasting approaches, fare decisions or other topics.

According to PODS website[5], it was first developed to be used at Boeing aircraft in the 1990s by a consortium funded by a corporate member of the Massachusetts Institute of Technology (MIT). During 2016, the membership of the consortium was comprised of Boeing, Canada airlines, Air America, Delta, and other airlines from UAE, Qatar and Switzerland. The members provide a source of funding and guidance for PODS programming and research, as well as novel PODS research topics.

The main focus of PODS research undertaken at MIT involves determining the interactions between RM optimization approaches, forecasting and widely used estimation models. In addition, PODS has sufficiently explored the Willingness To Pay (WTP) from passen-

[5] `http://podsresearch.com`, retrieved on September 29^{th}, 2018

gers and the revenue impacts that occur as a result of sell-ups, that are determined through a combination of several forecasting and optimization models. Through enhanced capabilities, the PODS has allowed for simulations on larger networks that follow a different airline route, and established passenger flow characteristics. Recent research has focused on modifications of existing RM methods, creating new ones including MNL model as well as cancellation forecasting and cabin optimization. The aim for those developments is to keep RM models consistent with the changing airline fare structures, competition from low fare companies, the increasing importance of ancillary revenues and growing levels of code-share associations.

PODS simulator has been used from researches at MIT on the following topics:

(i) Network optimization of alliance revenues based on amelioration on seat allocation or bid price methods;
(ii) Improvements of choice-based models and WTP estimations;
(iii) Combined optimization of multiple aircraft cabins;
(iv) Price optimization and simulation for novel fare families;
(v) Optimization of ancillary revenues;
(vi) Cancellations forecasting;
(vii) Development of 'classless' RM and dynamic pricing models.

PODS designed to achieve a realistic simulation and understand the interactions between passenger choices and behaviors and the effects of RM decisions, in the context of a competitive airline market. PODS features are mostly used in forecasting demand and levels of seat availability in relation to RM optimization. It takes into consideration both business and leisure demand in relation to the preferred choices of flight and fare classes options. The simulator is able to provide insights in realistic cases of transport environment, with more than two competitors on two different routes and diverse airplane capacities, departure alternatives and fare levels.

Figure 3.4 on p. 65, taken from PODS website[6], illustrates how

[6] `http://podsresearch.com/pods.html`, retrieved on September 29^{th}, 2018

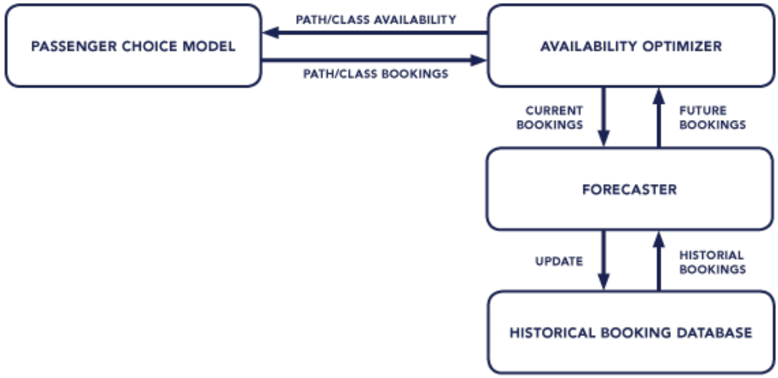

Figure 3.4: How PODS works. Source: http://podsresearch.com/pods.html, retrieved on September 29th, 2018

the simulator works through two separate interacting components, namely the Passenger Choice Model and the Revenue Management System (RMS) that, in turn, is composed by the Availability Optimizer, the Forecaster and the Historical Booking Database. All components exchange data on current and future bookings, availability and other real time information.

The RM system used by a single airline as modeled by the PODS comprises a historical booking database, a mechanism for forecasting and a seat availability optimizer. Within each specified Decision Control Points (DCPs) of the booking process, the airline forecaster provides an estimate of future demand levels utilising historical bookings. The forecasted values are input to the RM optimizer that uses them in combination with prevailing capacities per flight leg. This creates the possibility for determining a path and fare categories. In the simulation runs, the airline determines the price levels (provided with a set of restrictions) and the set of availabilities, as well as the number and type of seats and fares that should be available to passengers per cabin class, through the analysis of historical bookings. The levels of accepted bookings between consecutive DCPs are input to determine the booking limits, which can lead to leg class and path class closures.

A choice model which considers the passenger behaviors and preferences among several airlines is incorporated in PODS. Such model is able to consider travelers with distinct characteristics. These can be observed through their ability to pay certain fare levels, assign costs in relation to individual preferences over diverse schedules that are consistent to: preferred departure and arrival time, itineraries including additional stops, connections and restrictions provided for every class. For a single generated passenger, a complete path or class provided in the market will be assigned to those passengers with a certain willingness to meet that cost. Finally, remaining O&Ds and classes will be ranked in increasing order, based on the passenger evaluation of fare and related costs for every alternative solution provided. The simulated booking process will then proceed with the ordered passenger set, based on preference listed. This is recorded by the RM system and creates a useful historical data for future forecast.

3.4 Design of Experiments (DoE) Methodology

This part covers the basic principles of Design of Experiments (also referred to as DoE, or DoX) and has the purpose to provide a methodological background for the live test described in Chapter 6. The following description on the topic is based on a set of studies and manuals, in particular: [Seltman (2018)], [Montgomery (2001)], [Mason et al. (2003)], [Oehlert (2000)].

Experimental Design is a well known methodology, applicable in many fields to observe and analyze variations of factors under certain conditions and hypothesis, normally in association to experiments, or quasi-experiments (that comprehend the observation of a set of natural conditions impacting on the observed modifications)[7]. More specifically, DoE focuses on planning, conducting, evaluating controlled experiments and interpreting related data to derive conclusions and recommendations. It can also be seen as a systematic method to determine the causal relationships between factors affecting a process and the output of that process.

DoE can apply to tests related, for instance, to improvement and management of products and processes. In particular, for what can be of interest to our case, it can support a product or software validation and prototype testing to spot any defects or sensitive areas in advance. This mitigates the risk of failures, that can be higher in case of multiple design changes provided at the same time.

As stated in [Montgomery (2001)] and [Oehlert (2000)], engineering experiments provides several benefits, in particular to reduce the amount of time necessary to design and develop new products or processes, as well as to boost efficiency and results of existing processes (or products).While providing validity, reliability, robustness and repeatability to an experiment, DoE provides it with the statistical power to properly answering the research questions. It can also help to compare alternatives, evaluate components or subsystems, test sensitivity and tolerance for both products and processes. Finally, it ensures that experiments yield most information from the smallest amount of data collected, thereby minimizing costs and saving on time. In this sense, it is possible to state that DoE optimizes tests.

Broadly speaking, experiments are used for process characterization and optimization, as well as for the assessment of the properties but also design and development of products or single components and systems. [Montgomery (2001)] reports that all experiments are designed, but for some the design is made well, while for others it is done poorly. Therefore, there is always the opportunity for DoE. According to [Oehlert (2000)], a good DoE should prevent system-

[7] https://www.isixsigma.com/tools-templates/design-of-experiments-doe/design-experiments-%e2%90%93-primer/, retrieved on October 29th, 2018

atic errors, boost accuracy and error estimation, broaden validity of experimentation.

In its simplest form, an experiment aims at prediction of variations in the dependent variables after introducing a variation to one or more independent variables, while control variables are kept constant. An experiment consists on a test or a set of tests aimed at modifying the conditions to produce a certain impact on the outcomes and predict them. The treatments, experimental units and assignments as well as the observed outcomes are key features of an experiment ([Oehlert (2000)]).

[Oehlert (2000)] presents the key concepts and terms used in DoE, which among the others comprehend:

(i) Treatments: the procedures to be compared;
(ii) Experimental units: the subjects of the treatments;
(iii) Responses: the experiment results observed and measured after treatments are applied to experimental units;
(iv) Experimental errors: the random variability of tests outcomes;
(v) Measurement or response units: the subjects of outcomes analysis;
(vi) Control: (i) the assignment of treatments to experimental units, or (with a different meaning) (ii) a standard treatment which becomes the baseline;
(vii) Factors: the elements which are combined, at different levels, in the treatments.

[Montgomery (2001)] provides some suggestions to plan, run and analyze an experiment. In this view, pre-experimental planning and the respect of a strict sequentiality are key. Specifically, pre-test phases comprehend: (i) problem analysis; (ii) determination and characterization of relevant factors; (iii) selection of the response variable(s); (iv) selection of a proper DoE, fit for the purpose. The test phase follows, where experiment is run. The post-test phases comprehend (i) the analysis of the outcomes, supported by statistical methodology, and (ii) the presentation of results through conclusions and possibly recommendations. Further suggestions comprehend early statistical thinking, which should be complementary to the non-statistical knowledge or comprehension of the problem and its results. In addition, the data collection and analysis should fit the selected design and not the opposite. Finally, creating a clear and conclusive documentation of the experimental methodology is necessary to ensure that a research can be replicated ([Mack 2018]).

According to [Montgomery (2001)], diverse strategies for experimentation exist, comprehending:

(i) 'Best-guess' experiments, widely ans relatively successfully used, though having many disadvantages;
(ii) 'One-factor-at-a-time' (OFAT) experiments, in some cases associated with the engineering approach but sensitive to interactions with external factors and not properly efficient;
(iii) Statistically designed experiments, following the factorial concept from [Fisher (1971)].

Statistical methods and approaches are used in DoE to evaluate single or multiple changes to a process and offer a prediction of the output, under hypothetical conditions. In the case (mostly common in real life experimentations) that many factors affect the output results simultaneously in any design, a key challenge is to determine the individual and interactive effects, while addressing the higher risks of errors. According to [Mason et al. (2003)] and [Seltman (2018)], main statistical methods can be applied to DoE for selection of samples, parameters setting, mathematical and statistical modeling. In this view statistical inference and probability are widely used, as well as concepts like hypothesis testing and confidence interval, mean and deviation. In [Seltman (2018)], regression (and in particular linear regression) is considered the mostly used method for analyzing the relationship between quantitative independent and dependent variables. Based on [Mack 2018], in undertaking a regression design analysis six principles should be observed, including the consideration of capacity for primary and alternate models, minimum variance of all estimated coefficients and predicted values.

As explained in [Mason et al. (2003)], comparisons between treatments and against a baseline are used in case independent measurements of a quantifiable standard are not feasible. A major issue related to these multiple comparison procedures is represented by the high error rates for families of tests. One mitigation approach is smultaneous inference, that is aimed to control a specific error rate through a determined procedure. Other structured approaches (comprehending blinding and double-blinding) are used for the common issues of bias, false positives and so-called 'p-hacking' (unconscious data manipulation finalized to a certain result). According to [Mason et al. (2003)], other methods to mitigate the risk of biases comprehend the publication of the initial research question before the experimentation, the development of the experiment from an independent team, the use of the original data and the organization of the study in further phases. Furthermore, a way to perform many implicit tests is the practice of data snooping. It takes place when the null hypotheses is chosen based on a first observation of the data. The latter spots initial interesting aspects and does not consider others, which can be discarded for their behavior, either null or not of interest.

Another common issue is related to the elimination of spurious, intervening and predecessor variables as well as nuisance variables. The DOE methodology distinguishes among controllable and uncontrollable input factors and responses:

(i) Controllable (or controlled) input factors, or x factors are the ones that can be modified in an experiment or process;
(ii) Uncontrollable (or uncontrolled) input factors are those parameters that cannot be changed on purpose. These factors need to be recognized to understand how they may affect the response. They can be divided into factors which can be observed (u) or not (v). Furthermore, the observable factors can be either measurable or not;
(iii) Responses, or output measures (y) are the elements of the process outcome that produce the desired effect;
(iv) Nuisance inputs are not in the scope of the experiment but af-

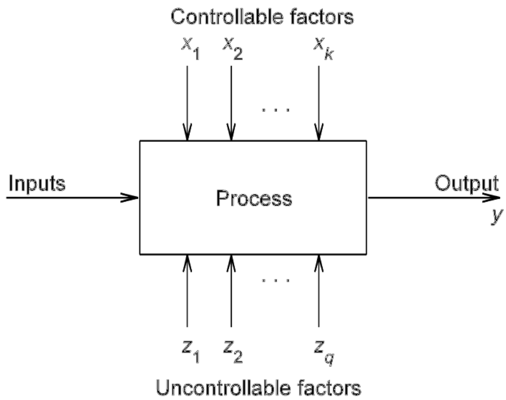

Figure 3.5: General model of a process or system ([Montgomery (2001)])

fect the outputs; they can be dealt with through implementing control checks and an additional measures in the experiment design.

The controllable input factors can be modified to optimize the output. A general model on on how a process or system works is provided in Figure 3.5 on p. 71 from [Montgomery (2001)].

According to [Montgomery (2001)], the 3 basic principles of DoE consist on:

(i) Randomization;
(ii) Replication;
(iii) Blocking.

Those methods are detailed in the followings.

Randomization

According to [Oehlert (2000)], randomization refers to the utilization of a known and well understood probabilistic scheme to one or more aspects of an experiment, such as the assignment of treatments to relevant experiment units. The selection of an adequate sample, comprehending a sufficient number of observations, can mitigate the risks associated to randomization (e.g. imbalance among test sets). Randomization can mitigate confounding, which takes place when experiment outcomes are impacted by factors not included in the treatment, i.e. balancing the effects of "lurking" variables. Randomization can also be used as a basis for inference.

'Contrasts' is a common method to assess the results of treatments and their related odds, measured as differences (or averages) among means. [Oehlert (2000)] states that contrasts allow a focused analysis of specific features. This is an advantage but at the same time a flaw, as it doesn't provide an overview. Therefore, several contrasts can be used to provide a picture on a set of determined characteristics, and may suggest to move the focus of the analysis accordingly. Based also on [Seltman (2018)], related null hypothesis is on differences of means, or combinations (e.g. averages) of means, of the population.

Also, the use of 'Factorial experiments' is common to test the outcomes and interactions of diverse independent variables, and in general the possible combinations of variable levels ([Montgomery (2001)]). Specifically, in 'Orthogonal' factorial design, orthogonality relates to the fact that contrast variables are uncorrelated and independent, therefore any treatment results in a different set of information.

[Mason et al. (2003)] states that performing a comparison of two samples needs an analysis on whether the experimental units were either paired (or with analogue preconditions) or there was independence between the results of the sets. According to [Oehlert (2000)], to compare the averages and variances of two samples, the 'Paired t-test' can be used. In this case, the outcomes for the two samples are dependent for the double measurement of units, one per set. The null hypothesis to be tested here is that the differences of the averages of the two sets are zero. On the other side, according to [Oehlert (2000)], the 'Two-sample t-test' is the standard method for testing whether the averages of two samples are the same in the case they are receiving two different treatments. It consists on testing the null hypothesis that the averages of the two samples is the same.

Statistical replication

Replication aims at mitigating the issue of variability and errors of measurement ([Montgomery (2001)]). Under this approach, not only the measurements can be repeated, but also the experiment itself, in order to:

(i) spot the factors that produce the modification of the dependent variables;
(ii) refine the measurements, or estimations, of the tests outcomes;
(iii) estimate the background noise or error;
(iv) improve validity, reliability and soundness of results.

This technique should include a sufficient sample size.

Blocking

Blocking consists on the classification of experimental units into subgroups, also called lots, or blocks, based on similarities. This approach allows an accurate identification of the factors that produce variations and improves the estimation of their effect; it also minimizes the effect of sources of variation among units which are known but considered not relevant, the so-called 'nuisance factors' ([Montgomery (2001)]).

In general, diverse DoE approaches should be used to design the experiment(s) that fit the research study. For instance, referring to what has been described earlier, the following methods can be used to deal with diverse types of inputs[8]:

(i) Controllable Input factors (x) could be managed through a Variation and Replication or repeat approach;
(ii) Observable and Uncontrollable Input factors (u) through either (i) Blocking, where every block has a predetermined value of u, or (ii) Covariance analysis, to estimate and subtract the value of the outcomes impacted from u;
(iii) Unobservable and Uncontrollable Input factors (v) by Randomization.

Software testing

The following part, aimed at providing further details on the use of DoE for software testing, is based on the works from [Phadke (2013)]. Here, DoE can be fruitfully applied to minimize the software utilization and test the overall operating domain in a statistically proven way. In particular, it can be used to detect the faults in a software with the minimum number of test cases. This is valid particularly for region faults, for which there are systematic errors, than for isolated faults that can be for instance related to specific combinations of parameter levels. For the latest ones, there isn't full assurance to spot them unless testing all possible combinations. For region faults spotting, a commonly used method is orthogonal array, which can be further developed through regression. This was demonstrated to be useful in all phases of software testing; productivity of test planned though the use of orthogonal array was greatly improved (by a factor of 2). Another one is OA-based test cases, supported for instance by a geometric picture of the test cases.

[8] https://en.wikipedia.org/wiki/Design_of_experiments , retrieved on September 29th, 2018

4 The Passenger Rail YM Problem

Even though RM has been largely explored across industries, the presence of literature related to rail RM is inexplicably low, as already pointed out by [Ciancimino et al. (1999)] and [Armstrong and Meissner (2010)]. The reasons why rail has been neglected were indicated by the above mentioned authors as the following: the dominant presence of RM systems that are either proprietary or taken from the airlines sector, as well as the minor use of rail in the US context where RM discipline was born in the 1970s.

Other main factors can be added on top of existing literature, based on the European experience; they focus on the mostly recent rise of liberalization, privatization and competition in the rail sector. In the absence of those compelling factors, the industry didn't develop earlier RM culture, systems and methodologies. We refer in particular to:

(i) the presence of the competition, both intra-modal (diverse railway Companies serving the same market) and inter-modal (across different means of transportation), exacerbated by the price transparency and comparability supported by the web, with the price as a major factor of customer choice; the competition at the same time is expected to lead to more mature and transparent markets in the long run;

(ii) the process of liberalization and privatization of railway industry across Europe, together with a global economical stagna-

Reference	MF	MP	ML	DP	CA	Opt	DE
Ciancimino et al. (1999)	1	×	n	×	√	√	×
Hood (2000)	1	√	×	×	×	×	√
Sibdari et al. (2008)	1	√	×	√	×	√	√
Bharill and Rangaraj (2008)	1	√	×	×	×	√	√
You (2008)	2	×	n	×	√	√	×

MF: Multi-Fare, **MP:** Multi-Product, **ML:** Multi-Leg, **DP:** Dynamic Pricing, **CA:** Capacity Allocation, **Opt:** Revenue Optimisation, **DE:** Demand Estimation

Figure 4.1: Summary of Railway Passenger RM Models - Source: [Armstrong and Meissner (2010)], page 17.

tion, that are pushing railways to give more and more attention to revenues and margins.

Up to 2010 the research works, summarized in [Armstrong and Meissner (2010)], were few and barely applicable to a real world railway context, which is multi-leg, multi-fare and strictly regulated (therefore, without the real possibility of bumping discounted passengers to other trains as suggested by [You (2008)] for instance). Figure 4.1 on p. 76 from [Armstrong and Meissner (2010)] summarizes Railway Passenger RM Models developed up to 2010. After 2010 other researches have been done, mostly towards customer centricity and market awareness.

4.1 Main Specificities of the Railway Sector

The development of many High Speed (HS) lines and their increased performances boosted a direct competition against airlines in domestic markets. By fact, in many cases (e.g. TGV in France, Shinkansen in Japan, ICE in Germany, Le Frecce in Italy) there are many similarities between HS railways and airlines as well as several diversities. [ExPretio (2009)] has properly pointed out the important differences, that are also consistent to earlier findings by [Armstrong and Meissner (2010)] and [Mitev (1998)]. In the followings main similarities and differences of rail and airlines sectoers are listed, for what is relevant to the RM problem; they comprehend the above mentioned findings and novel points.

The two industries appear to share the following similarities:

(i) they have a complex mix of business and leisure customers,
(ii) demand levels vary and a cyclical seasonality can be identified in relation to month of the year, day of the week, time of the day; demand variations are also influenced by external events such as holidays, special events, strikes and heavy weather conditions;
(iii) they manage a perishable inventory, for which the value of any (unsold) seat after the departure lowers to zero;
(iv) they share a similar cost structure, with high fixed costs and low variable ones, so that the marginal cost of accepting on board an additional passenger is low;
(v) the booking process is managed in similar manners, broadly speaking, with mandatory reservations, pre-assigned seats, fares differentiated through restrictions and fences.

Nevertheless in rail, differently from airlines:

(i) the rolling stock assigned to the routes is fixed and predetermined, even if it may be coupled in certain cases; this is different from the aviation sector, where there is some flexibility as the fleet is composed by diverse airlplanes with different capacities and can be scheduled according to customer demand;
(ii) the level and maturity of competition in the two sectors is pretty different;
(iii) the line network topology with a vast presence of O&Ds is a key characteristics of rail;
(iv) there is a high interdependency of demand across legs, segments and O&Ds;
(v) the average Load Factor is much lower in railways, with very few trains surpassing 90% — a level from which classical RM

approaches work at their best — this is also related to train topology and unbalanced demand per O&D;

(vi) higher inter-modal competition in rail, with massive presence of alternative and interchangeable means of transportation, especially for shorter O&Ds and in certain circumstances; this implies a high presence of go-show, no-show and walk-up phenomena;

(vii) overbooking (which is one of the key points for airline RM) is rarely practiced in rail;

(viii) rail is more involved than airlines in the societal challenge to offer accessible, inclusive and affordable transport for all, also on HS trains — this is valid especially for National operators.

Generally speaking, rail is more complex than air, as pointed out in [Mitev (1998)]. Figure 4.2 on p. 79 from [Mitev (1998)] illustrates main points on rail and airline industries with a difference between European and US markets.

4.1.1 Multi-leg 'Line' Topology

A specific characteristic of the railway industry is the number of stops during a train route, or multi-leg topology, which leads to the presence of several segments. Here a segment is defined as any O&D served by a single train, and leg as any route between two adjacent stations (also called 'nodes'). While demand shows up by segment, the capacity supply is defined by leg. Let n be the number of nodes, the number of legs will be $n-1$ and the number of segments $\frac{n(n-1)}{2}$; the number of segments increases quadratically with the legs ([Gliozzi et al. (2014)]).

The multi-leg topology is a complicating factor for RM. This may be the underlying reason for the choice from some RM providers to implement in the rail industry RM systems which are leg-based. If such choice is easily understandable, at the same time it doesn't exploit one of the key feature of the industry. As pointed out by [Belobaba (2002)] for the airline industry, the use of O&D control

Table 1. Comparison of US and European air and rail transport

	US	Europe
Trans-port market	Long distances Unique liberalisation regime	Short/medium distances Densely populated National liberalisation regimes
Air	Intramodal compe-tition Concentration of operators Hubs and spokes	Intramodal pan-European competition Many European operators Intermodal national competition between air, rail and road
Rail	Little rail passenger transport No intermodal air/ rail competition	High speed and tradi-tional trains Many stops Costly infrastructures National intra and intermodal competition Little pan-European competition

Figure 4.2: Comparison of Air versus Rail transport for European and US markets, as reported in [Mitev (1998)], page 3.

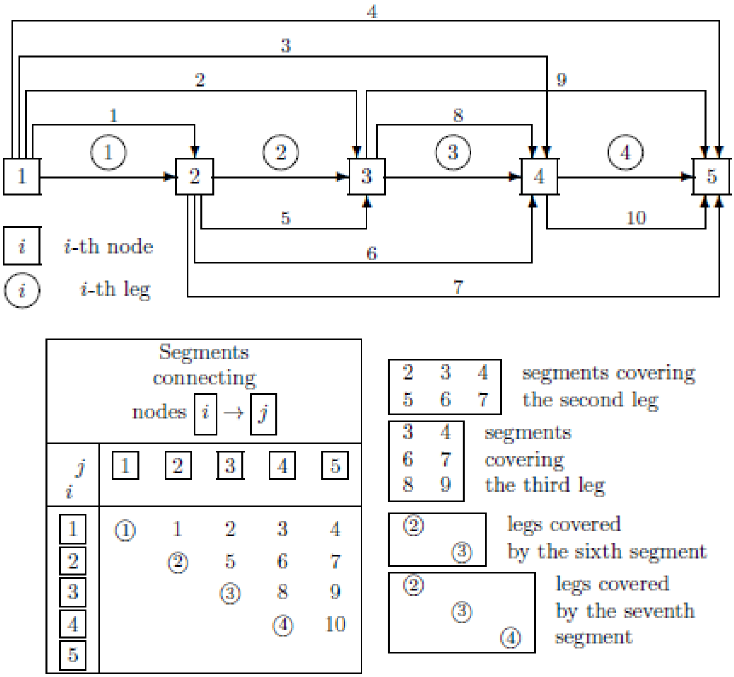

Figure 4.3: Basic relations of nodes, legs and segments, with n= 5, from [Gliozzi et al. (2014)], page 4.

can improve network revenues by 1-2%. This seems even more valid for the railway industry, that is much more impacted by the presence of a large amount of O&Ds than airlines.

Finally, the presence of this topology, together with the fact that some segments are much more requested than others, give space to the presence of bottlenecks. This may be the case, for instance, of the segments traversing the central legs of the overall itinerary. Therefore, there are some legs which are under-capacitated ('bottlenecks'), while others are not. RM can be an useful means to correct such imbalance, e.g. 'selecting' the high—pay demand for segments traversing bottlenecks or adjusting price levels, to maximize revenue. On the other side, a company has the opportunity to use the data on demand and acting with other levers, also in medium and long term, e.g. communication or planning.

4.1.2 Fare Basis Structure

Most railway companies have a fare basis (clustered) structure. This means that for each Service Level (also referred to as Class, and in aviation as Cabin) a certain number of different fares can be available to the customers. Each service level has a distinct physical capacity, so that a passenger in class A and a passenger in class B are never competing over a seat (unless free upgrades are offered, or similar cases). Therefore, this can be treated as a multi-product problem.

Customers are commonly subdivided in segments through a clustering process that aims at obtaining a reasonable number of groups, consistent with the managing possibilities of the RM and Reservation systems, populated and stable enough. Each segment should be relatively homogeneous, compared to the others, for customer behavior and preferences (and other variables which may be considered by a RMS).

Finally, it is assumed that within the same fare the prices respect the 'triangular inequality' within the same fare basis. Considering a route with the ordered nodes A, B, C and the price V for the segment AC, traversing the legs AB and BC, defined as V_{AC}; the inequality

$V_{AC} \leq V_{AB} + V_{BC}$ is always satisfied for any given fare.

4.1.3 Seasonality and Outliers

As described for the general transport case, it is of paramount importance for RM forecast to understand seasonal fluctuations, trends and statistical variations of demand. Not only the amount of booked customers and the time when bookings materialize is unknown, but also the customer mix (and related revenue), the arrival order of bookings (LBH is not a given). A certain trend can be normally seen at high level for the overall demand for transport, but also for demand related to single modes and companies.

According to [Zeni (2001)], seasonality is one of main factors impacting on demand forecast. Those factors include namely:

(i) seasonality;
(ii) special events;
(iii) sensitivity to pricing actions;
(iv) demand dependencies among fare classes;
(v) group bookings;
(vi) cancellations;
(vii) censoring of historical data;
(viii) defections from delayed services;
(ix) no-shows and go-shows;
(x) recaptures.

Notably, this list include special events: those factors are not to be included in seasonality but are related to it, either directly or indirectly. In particular, they can be described as particular dates which perform differently from cyclical seasonal fluctuations and regular random noise that can be associated to normal demand behavior. For instance, Christmas takes place in different days of the week every year, so it is both difficult to compare the performances associated to that date to the ones of previous week, or make comparisons to previous year. Such special dates of demand shocks, which differ slightly from the normal seasonal behavior and performances, lead to 'outliers', here defined as observations with anomalous values. While seasonality per market can be related to month of the year, day of the week, hour of the day, also in rail industry, peaks are often concentrated in certain hours of the days (e.g.due to massive presence of commuters), days of the week (e.g. related to business travels), weeks and months of the year (e.g. groups can be concentrated in Spring).

A RM system should take into account demand fluctuations and deviations and react to them. It should be also calibrated periodically to adapt to different seasonalities and paths which can take place over time. On the other side, the same RMS should be able to detect as soon as possible outliers, promptly react and alert the RM team.

The real issue with rail RM is that, differently from aviation, here the supply is fixed with few exceptions, as described earlier. Therefore, during demand peaks there is an excess of demand which can be only partially reallocated to other trains, while the other part cannot be recaptured; this leads to loss of passengers and revenue. In such cases the price lever can ease the reallocation of price sensitive customers to off peaks trains or days. On the other side, in presence of low demand a YMS can offer lower prices to promote bookings. In this second case, discounted prices shall be communicated to diverse travelers as much as possible, in the attempt to sell the discounted seats to new customers and limit the risk of dilution. The latter takes place if customers, already willing to travel at a determined price, find lower prices available for equivalent services and buy the discounted tickets. Another form of dilution control relies on fare restriction policies, which are widely used in the transport industry.

4.2 Main Theoretical Approaches and Implementations

This part follows the same approach of the general section for the presentation of models, but here the focus is on rail specific researches and on implemented models.We have already seen main traditional approaches on rail RM and pointed out how the presence of scientific literature related to rail RM is inexplicably low, although RM has been largely explored across other industries. Among main European railway operators, a couple use proprietary implementations, one (Trenitalia) has implemented a custom solution with IBM, and others use (customized) commercial softwares.

Demand Forecast

[Milenković and Bojović (2016)] present an overview on rail forecast methods and the results of the application of these innovative approaches to Serbian railways. Qualitative methods to forecast demand are not included here as not relevant for our purposes: their results are not suitable to feed RM models. Therefore, in the followings the focus is on quantitative approaches, here divided into econometric (or causal) and time series. Based on [Milenković and Bojović (2016)], econometric Models analysis comprehend:

(i) Univariate and Multivariate Regression Analysis, Odgers and Schijndel (2011);
(ii) Co-integration and Error Correction Approach, Wijeweera et al. (2014);
(iii) The Engle and Granger co-integration method (1987);
(iv) Two time series regression models, Doi and Allen (1986);
(v) Co-integrated VAR methodology (for freight), Kulshreshtha et al. (2001);
(vi) Three econometric models for the forecast of rail and one private car demand, Profillidis and Botzoris (2006);
(vii) Six econometric time series models, based on annual time series, for road plus rail freight, including: OLS (Ordinary Least Squares) regression, PA (Partial Adjustment), reADLM (reduced Autoregressive Distributed Lag Model), unrestricted VAR (Vector Autoregressive), TVP (Time-Varying Parameter) model, and STSM (Structural Time Series Model), Shen et al. (2009);

(viii) Analysis of freight demand related to macroeconomic trends, Wijeweera et al. (2013);
(ix) Support vector regression method and BP (Back Propagation) neural network method, Xia et al. (2014)
(x) A set of methods prescribed and illustrated for UK in the Passenger Demand Forecasting Handbook by ATOC — Association of Train Operating Companies, (ATOC, 2005);

On the other side, as stated in [Milenković and Bojović (2016)], Time Series Models can be listed as follows:

(i) SARIMA models, Milenković et al. (2013);
(ii) ARIMA and Holt-Winters models (freight), Guo et al. (2010);
(iii) a combination of ARIMA and RBF (Radial Basis Function) neural network model (freight), Jiuran and Bingfeng (2013).

Besides these two broad classes of approaches, in [Milenković and Bojović (2016)] are also presented other approaches suitable to handle non linear data. They are namely:

(i) Kalman filtering;
(ii) Neural networks;
(iii) Fuzzy models;
(iv) State-space models.

Inventory Control

As outlined in [ExPretio (2009)], early approaches to RM based on booking limits were built on basic assumptions, namely:

(i) cabins subdivided into booking classes, mutually exclusive;
(ii) fares fixed and predetermined;
(iii) use of fences to limit dilution;
(iv) demand estimations based on historical data of similar trains;
(v) seats availability determined per leg, not exploiting the network aspects;

(vi) the order of arrival foresees early bookings of lower fares (LBH) and sequential closures of fares;
(vii) demand for different classes or services is independent; this excludes possibilities of sell-up and recaptures and in general of choice substitution patterns among services and fares.

More recently, other researches deployed network-based inventory control systems, optimizing the allocation based on the network value of a seat. Such studies comprehend the bid-price approach, solved for instance with a capacitated network LP (either deterministic or stochastic) and summing the Lagrange Multipliers, or 'shadow prices', associated with capacity constraints along that itinerary.

4.2.1 Origin-Destination and Fare Inventory Control

This paragraph summarizes the surveys on Origin-Destination and Fare (ODF) seats allocation techniques from [McGill and Van Ryzin (1999), Chiang et al. (2007), Armstrong and Meissner (2010)] and reviews further works, mostly focused in behavioral approaches. The aim is to provide a complete view even though most of them have limited potential of applications in the real world railway context, which is multi-leg and multi-fare and in most cases offers public transport services. This exclued, for instance, the real possibility of bumping discounted passengers to other trains as proposed by [You (2008)]. The choice of a specific RM model should take into consideration the context it should fit, with its main features and constraints; not considering the rail context, with its multileg topology and social issues, can lead to failures, such as the initial implementation of a leg-based RM used in aviation from SABRE at SNCF, described in [Mitev (1998)], and the issues at DB AG described by [Link (2004)]. Notably, the scarce research works on the topic up to 2010 (i) had limited presence of nesting (most common in the air transport sector) and (ii) included some forms of dinamicity with the suggestion to re-optimize trains during the booking horizon e.g. in [Ciancimino et al. (1999)], [You (2008)].

[Ciancimino et al. (1999)] developed a model for single fare, multi leg capacity allocation problem for seats allocation based on data from Ferrovie dello Stato Italiane (Italian State Railways Group). Both a deterministic model (LP) and a probabilistic formulation based on normal demand distribution were developed. The results suggested that the potential for revenue gain was related to the increase in the number of legs.

[Kraft et al. (2000)] explored common characteristics and critical differences between a variety of railroad revenue management problems, mainly O&D traffic management as well as overbooking and price discrimination. They are characterized by the importance of the network and study the booking arrival patterns, which are independent from fare class value and have very short booking lead times, except for long distance passenger service. The advantages of bid-price approach are explored towards traditional airline-style, leg-based EMSR approaches. The work is focused on traffic mix optimization, that is, determining the optimal combination of origin-destination fares across a route. A list of implementations is provided.

[You (2008)] developed a constrained nonlinear integer program for the determination of seat allocations, through an efficient heuristic approach on booking limits for potentially all ticket types in a railway network. This is applied to a rail booking system for two fares (full and discounted price). The work extended the one from [Ciancimino et al. (1999)] to a two-fare, multi-leg model. A hybrid optimization algorithm was used for solving the problem. Its underlying assumption on the possibility of bumping the low pay passengers to other trains appears as unrealistic. It was suggested to use multiple run of such algorithm at different points of the booking horizon, though not considering directly the multi-stage aspect of the booking horizon, as in [Ciancimino et al. (1999)].

[Hood (2000)] presents "MERLIN: a model to evaluate revenue and loadings for Intercity" on data from British Rail. The aim was to support timetabling and pricing decisions through a choice model and encourage passenger shift to off peak trains through the application of YM features in association to fenced fares.

[Bharill and Rangaraj (2008)] considers how RM can be applied to Rajdhani Express, a HS segment of Indian Railways. The work provides the estimation of cross-price elasticity of demand for three products, with a focus on how demand could react to cancellation fees and others; it also analyzes pricing strategies and, in general, is aimed at decisions support.

[Sibdari et al. (2007)] focuses on the development of a series of pricing policies for a multi-product RM problem at Amtrack Auto Train, where tickets cover a bundle price including both vehicle ad accommodation. The model they develop is then solved through different approaches and the dynamic program outperforms the others.

Among the other recent research works are the following. [Wang et al. (2016)] studied a stochastic seat allocation problem for passenger rail revenue management with discrete random demands through a variety of policies tested using simulation studies. The allocation is decided for each cabin class and train service. [Xiaoqiang et al. (2017)] focuses on dynamic pricing strategy for seat inventory control at China high-speed rail. The aim is to determine the optimal price for passenger groups and handle them as individual travelers. [Hetrakula and Cirillo (2014)] proposes an empirical study of railway revenue management using ticket reservation data, and the latent class choice model in the form of multinomial logit to account for passenger heterogeneous preferences. The proposed formulation allows for simultaneous optimization of pricing and seat allocation.

4.2.2 Customer Choice

As outlined in previous chapters, the Customer Choice-based model is one of the main alternatives to “classical” rail RM models based on inventory control and nesting; related researches started in 1970s mainly for macroeconomic applications. Among its advantages, it takes into account the current market situation and incorporates customer preferences. The underlying vision is of customer as actors with the possibility to choose and react to the decisions from the market operators, e.g. on pricing.

Among its disadvantages, it is computationally hard to solve. By fact, investing effort on the above mentioned factors forces in many case the choice of a simple modeling of network topology, e.g. leg and not O&D-based. Therefore, a possible risk associated to this model can be of not exploiting the line network multi-leg topology of trains, which is one of the main features of rail industry. In [ExPretio (2009)] the Choice model formulation develops a model per single O&Ds, here considered as Operational Constraints of the objective function. Other models instead explicitly take into consideration the presence of scenarios and O&Ds as a key part of the problem.

The key assumptions of the Choice-based model are related to the rational choice of the customers on a set of alternative choices based on their utilities (i.e. perceived value related to the attributes of a choice). Utilities are described as deterministic with a continuous random noise ("random utilities"). The random noise incorporates "the aggregate value of attributes that are suspected to have an impact on choice decision but cannot be measured, or even observed, by the modeler" [ExPretio (2009)]. Based on the probability distribution of the random term it is possible to identify different models, e.g. the LOGIT and PROBIT choice models (under Gambel and Gaussian probability distributions, respectively), the first with a pretty simple resulting formula for choice probabilities and the second being more realistic.

One of main implementations in rail, related to the choice-based (also called customer centric or behavioral) revenue optimization, has been developed by [ExPretio (2009)]. Most of the information in this paragraph are derived from its Technical White Paper (2009) ([ExPretio (2009)]).

Under Choice-based optimization, forecasts are performed at aggregated level, then subdivided into the choices actually offered based on the behavioral model, anticipating and stimulating substitution patterns and providing atomic forecast of demand. For instance, Appia Suite solution of ExPretio Technologies ([ExPretio (2009)]) forecasts at customer segment, O&D and time-window level. It includes element of dynamicity through considering the variation of behavior which occurs during the booking horizon, impacting also to the mix of passengers over time. Here demand elasticity is computed as the difference of demand between scenarios with diverse prices available. Competitor prices are collected through 'web-scraping' tools, able to extract price availabilities of competitors (e.g. on the web and travel agencies) to be incorporated in the models.

In [ExPretio (2009)], the forecasts feed an optimization framework based on bi-level programming, considering the railway operator as leader and customers as followers. The first decision variable of the model (the set of attribute values) is determined by the railway company, while the second (the traffic vector) is determined by the customer choice. The railway operator is willing to maximize revenues under commercial and operational constraints, while travelers maximize the utility related to the set of attributes of their choice. The bi-level Revenue Optimization problem is then formulated as follows:

$$\max_{v,X} \quad p(X)v \tag{4.1}$$

$$\text{s.t.} \quad X \in \chi \cap C \tag{4.2}$$

$$v \in V \tag{4.3}$$

$$v = \Phi(X|\beta) \tag{4.4}$$

Where:

$p(X)$ is the vector whose components are the price attributes of each alternative $i \in I$

v is the "traffic" vector representing the number of passengers expected to be present on the alternatives of set I; it is a collection of itinerary/product combinations

X is the market state, defined by a set of attribute vectors for the alternatives of a choice set I

χ is the superset of all possible market states

C is the set of commercial constraints, comprehending the fare structure, discounts and cards

V is the set of operational constraints, i.e. route structure, schedule and inventory

The last equation indicates that the traffic vector must be consistent to the demand model, with β as a weight for each attribute.

In this basic formulation are then included both the dynamicity d from the presence of multiple time periods within the booking horizon, and the customer segmentation s. It shall be noted that here the multi-leg topology is incorporated in the model within the operational constraints V.

Lastly, [ExPretio (2009)] pointed out how bi-level pricing models are "the hardest computational problems currently known" ([Roch et al. (2005)]). This formulation in particular is NP-Complete, for the presence of several regions of local optimality to be tackled with global optimization methods subdividing the solution space in subregions which shall be solved separately. Then, local optima (which can be exponential numbers) are compared to find the global optimum. The large size of real problems to be solved with this model represent a major challenge.

4.2.3 Price Elasticities

It was described previously how own and cross elasticities of demand can be incorporated in sell-up, hybrid and choice-based models. For the rail sector, among scientific literature one can mention the work from [Gama (2017)] which estimates the own and cross elasticitiy of demand to price and substitutability of domestic flights and passenger trains; here the approach from Berry (1994) is adopted to evaluate whether travelers would change transport mode. It was concluded that the substitutability between trains and flight is pretty low, and that internal substitution among trains is higher. In the following section the Italian experience will be presented, with some differences to conclusions of [Gama (2017)]; in our case, the development of high-speed lines improved inter-modal competition while allowing for comparable travel time between trains and flights. On substitution among trains, this can be particularly high whenever the frequency of the alternative offers is high.

A common approach, with some implementations in the rail industry worldwide, focuses on the search for the 'optimal' point of price to be suggested per each fare. This proprietary implementation is the Travel Price optimization (TPO) from [JDA (2009)]. This model, for which a US patent is pending, is derived from service implementations and approach that comes from manufacturing and service industries, then adapted to the rail sector.

Here, the demand is modeled as a function of a price difference ratio towards the market reference price. The latter is computed as a weighted average of competitors prices. Demand can change as a function of both own price and competitors' price.

[JDA (2009)] provieds the following definitions:

(i) Own-elasticity indicates how a modification in the monetary value of a Demand Forecasting Unit (DFU) is likely to impact on demand and is measured as percentage variation in the amount of demand for a DFU for a one percent variation in the rate;
(ii) Cross-elasticity quantifies how a rate variation of a DFU is expected to affect the sales amount for the relevant DFU and is measured as percentage modification in the demanded quantity per DFU for a one percent change in another DFU's price.

The Price Optimizer developed in [JDA (2009)] provides an optimal price profile, with regard to time to departure and considering demand forecasts, competitors price and supply, inventory and other data and parameters. It can maximize revenues, profits, market share. The Optimized Price Recommendation module determines the optimal price using a quadratic program, that includes in the objectives the maximization of revenue, modeled as a quadratic function of the price. The TPO recommends prices by dated-DFU, so the time dimension and specifically the moment in the booking cycle is considered. Competitor prices as well as demand elasticities are considered in both the forecasting and optimization models.

[JDA (2009)] includes a Constrained Forecast Evaluator (CFE), a post optimization discrete event simulator that allows comparison of the forecasted bookings and revenue at current price profiles and optimized price profiles provided by the TPO. It allows decisions on acceptance for bookings to come at each price profile for the available inventory and forecasts related results, as well as enforcement of constraints, previously fulfilled only in part.

User parameters allow, among the others, to: weight competitors importance, prioritize price recommendations, set the nearness of recommended prices to actual ones. Model parameters, on the other side, include: baseline demand forecast, competitive intelligence, relative weighting, inventory data parameters, data representing real-world objects, penalties and model configuration parameters.

5 Building a Yield Management System (YMS) at Trenitalia

5.1 Background

Ferrovie dello Stato Italiane Group (FSI), Italian national railways, is the the third in Europe for number of passengers. The total length of Italian railway network infrastructure is of 17,560 km, of which FSI part accounts for 16,787 kilometers, and 2,201 active passengers stations. Passengers transported yearly are 600 Million, and revenue of 2016 amounted to about 9 Billion Euros. The value created by the Group represents 1% of Italian GDP, and total investments foreseen up to 2026 account for 94 Billion Euros. The number of employees at the end of 2017 were 74,436. FSI is organized as a Holding Group, present in 60 Countries globally and comprehending: the National Rail Infrastructure Manager (e.g., Rete Ferroviaria Italiana), Train Operating Companies in Italy and abroad, bus companies, engineering companies for large railway projects, station managers and others. Among its subsidiaries, Trenitalia is the main Italian Rail Operating Company and former incumbent, providing both long distance and regional services[1].

The long distance, High Speed (HS) services ‘Le Frecce’ offered by Trenitalia are classified in ‘Frecciarossa’, ‘Frecciargento’, and

[1] Source: `https://www.fsitaliane.it/content/fsitaliane/it/il-gruppo-fs/la-holding-fsitaliane.html`, retrieved on October 29^{th}, 2018

'Frecciabianca'. In the first semester of 2017, more than 260 trains Le Frecce were offered per day on average, carrying 24 million passengers; in the first three months of 2018, the passengers have been nearly 10 Million.

Since the end of 2005 it was implemented a custom made Yield Management System (YMS), which is described in this section. The initial implementation covered a limited number of trains: part of the High Speed trains, at the time branded as 'Eurostar Italia' and the low cost trains 'TrenOK'. It followed an initial phase of analysis and test on prototypes, starting from 2004, which took about one year. The YMS presently optimizes most of Le Frecce services.

At the time the system was conceived, the status of pricing and demand selection at Trenitalia and the social framework were the following:

(i) the standard ticket price was equal for everybody, with a fixed amount of discounted tickets accessible to everybody. Those'logical contingencies', predetermined and fixed per train class and date, were based on the average Load Factor per train and day of the week and corrected for certain peak calendar dates. This approach is still in use for some type of trains for which the implementation of the YMS is not convenient.

(ii) Trenitalia was the monopolist and rail service was perceived as "universal and democratic" from the customer and social base. As a consequence, the fairness perceived on any decision taken on price and availability was paramount. This had a profound impact on the choices performed for first implementation of the YMS.

Part of the Yield Management approach described in this thesis was previously implemented by IBM in the Airline Sector at Alitalia Cargo, where it was characterized by a mono-dimensional environment and volume and weight optimization. The methodology implemented at Alitalia has been submitted by IBM for a US patent in 2001 ([Gliozzi and Marchetti (2008)]). It took more than one year of analysis for the joint working group Trenitalia-IBM to study how this solution could fit to the passenger rail industry, and specifically to the context described above. It was finally decided to exploit the multi-leg structure for seat optimization, while adopting a scenario-based stochastic linear programming approach. In this way, it was possible to take into account not only the stochastic nature of the demand, but also the correlations between the different demands for all combinations of segments and fares, without the need of trying an estimation of the covariance matrix, for instance.

5.1.1 Italian Competitive Landscape

At the time the YMS was being implemented, major changes were taking place within the European Railway market. In the followings they will be outlined, with a focus on the ignition of competition and its effects.

The First Railway Directive 91/440/EEC on rail competition paved the way to the competition among rail operators, followed by the European Directives 2001/12/CE, 2001/13/CE and 2001/14/CE. Their implementation into the Legislative Decree nr. 188 of 8^{th} of July, 2003 ignited in Italy the process of liberalization of the rail industry.

According to [Beria et al. (2018)], FSI was re-organised as a public holding company in relation to the need of un-bundling network and services (ART, 2014) and an independent regulator (Autorità di Regolazione dei Trasporti, ART) initiated its activity in 2013. Following this, any authorised rail company could have access to the national railway network through either open access or contracted services. Italy is actually one of the most liberalised countries in Europe for degree of liberalizaion and competition on rail market services, and one of few cases with open access competition (or on-track competition) on High Speed rail.

Nuovo Transporto Viaggiatori (NTV), which is considered the largest newcomer operator within on-track competition at European level, entered Italian HS market in 2012 with its brand ‘Italo’. Such entry impacted the market landscape, including both intra-modal and inter-modal competition conditions. Also Areanways entered the marke,t but failed after few months in operation. [Beria et al. (2018)], AGCM (2012) and [Bergantino (2015)] provide detailed information on the process. Similar cases are represented by Sweden, Czech Republic and Austria.

The effect of the liberalization was studied by few researches. [Beria et al. (2018)] reported works from [Bergantino et al. (2015)] and [Cascetta and Coppola (2014)] related to the competition on the route Milan-Rome, where it was particularly noticeable. [Beria et al. (2018)] presented some researches on intra-modal competition. [Cascetta et al. (2013)] and [Cascetta and Coppola (2015)] based their work on empirical evidence rather than simulations. In particular they performed a multi-year survey on customer behaviors for the HS route Milan-Rome. The latter study confirmed and enforced the results of the first ones on modal shifts, measured as share of PassengerKm (where each passenger is multiplied by the length of their trip in Kilometers). Outcomes displayed a reduction of modal share from 57% to 44% for cars and from 10.5% to 7.3% for air following the raise of competition.

According to [Beria et al. (2018)], few researches have been done on the effect of liberalization on fares. In particular, [Cascetta and Coppola (2015)], [Cascetta and Coppola (2014)] and European Commission (2013) stated an average price reduction of 31% in one year and of 34% in two years in the O&D Milano-Rome starting from 2011. Furthermore, [Bergantino et al. (2015)] studied on prices offered by Trenitalia, NTV and two airline carriers for the O&Ds Rome-Milan, Rome-Turin and Rome-Venice, finding evidence of strategic pricing decisions for both operators but not of predatory pricing behavior or price leadership from any of the two. Other researches have been done on intermodal competition and specifically between air and rail. [Beria et al. (2018)] reported in particular the works from Mancuso (2014), Yang and Zhang (2012), Bergantino and Capozza (2015), Hazledine (2011), Malighetti et al. (2009) and Alderighi et al. (2011).

The same [Beria et al. (2018)] studied the evolution of prices for the route Milano-Ancona in two periods of three months in 2013 (when Trenitalia was the only operator on the route) and 2014 (with the presence of both Trenitalia and NTV). The data collection for the experiment was performed through the use of a web-scraping application, which collected the prices available on the web from both Trenitalia and NTV at predefined time intervals before departure. Results displayed a reduction of average prices of the Economy Class by a 15% (from 10% to 20% in relation to the time before departure). Furthermore it was evidenced a different approach of the two operators over time during the booking horizon, with NTV struggling to be more convenient quite in advance before departure and rising prices over Trenitalia since few days before departure, and Trenitalia not responding to NTV price tactics. For this case, the newcomer in the short term acted as a price-taker. On the upper business classes the effect of competition was not observed, as Trenitalia maintained the same price for the flexible fare.

At national level, the analysis displayed how Trenitalia used a set of other levers in association with pricing, in particular the offer of an integrated network and pretty frequent services ([AISCAT 2010]). Also, in general there have been a major effort from all operators to improve customer experience related to booking and information, such as notifications in case of delays ([ENAC (2010)]). In other national cases, specifically in Czech Republic prices fell dramatically in response to the raise of competition, according to Tomes et al. (2014) mentioned in [Beria et al. (2018)]. (Bergantino, 2015) pointed out how a reason for newcomers to focus on price competition can be their lower costs and their lack of frequent services or integrated network. He also states how in other cases the newcomers aim at providing a new business service, thus the effect may be of service improvement rather than price fall, and relates this to the Italian case.

At the 10^{th} World Congress on High Speed Rail held in 2018 by UIC, the Global Railway Union, Eng. Mazzoncini (CEO at FSI at the time) stated that Trenitalia held around 70% of the passenger market for domestic high speed travels, according to [Giuricin (2018)].

Globally, high speed rail service was presented as a success story, as stated in [Givoni (2006)], while providing reliable and safe connections on busy routes based on [Cascetta and Coppola (2012)]. This has increased the number of passengers by 20% whilst the rate of passenger-kilometer raised to over 40% [Campos and de Rus (2009)]. The adoption of high speed rail in the world has increased up to 42,000 kilometers totally.

5.1.2 Social Aspects

As largely discussed, the main purpose of a YMS is to improve revenue performances of supplied services as much as possible. Nevertheless, it needs to satisfy customers through decisions on price and availability and related timing.Even more important is the acceptability of those decisions, i.e. the fairness perceived on the price discrimination operated. For instance, a customer would probably accept that the availability of discounted prices is limited and may end some days before the departure date; at the same time, however, she would consider unfair if another customer has the opportunity to pay less based on different individual characteristics or any other information. In addition, [Leary (2010)] states that social acceptance levels can be different in diverse sectors, and in particular they usually differ from the levels of other industries apart from rail transport that have already implemented revenue management and faced a certain degree of competition. Therefore, companies in less mature markets should act carefully while implementing RM.

The social exchange theory can be used to evaluate the extent of acceptance of a RM system from society. Under this theory, [Blau (1964)] specifically derived two important exchange patterns, social and economic, in dealing with customer demand. Social satisfaction is observed where the customers trust and appreciate the new system. On the other hand, economic satisfaction is provided where it meets the customer demand. Under this view, social acceptance of RM can be higher in the event that it is able to properly provide a correct forecast of demand and adjust its decisions accordingly.

This in turn enables the RMS to take decisions and management to set processes that will maximise revenues achieved from sales of passenger tickets and bookings, extracting value from high-pay customers while keeping availability of discounted fares where the presence of price sensitive demand is envisaged. At the same time, demand data shall be used to improve service planning.

According to [Kalyanaram Little (1994)], price sensitivity is an important pillar to be exploited for performance and revenue realisation of an organization. Monroe (1973) defined price sensitivity as the level of awareness and response that customer demand is likely to show in relation to the price of a product or service. In this view, customers are willing to accept a price up to a certain threshold or range, and reject the purchase when the price exceeds limits. Therefore, a RM system that causes an excessive increase in the offered prices will be responded by negative perceptions from the customers; contrarily, a system that causes a decrease in the price is likely to get a wide level of social acceptance, but will face revenue loss and other negative effects. A correct demand forecast, taking into consideration price sensitivity of demand, enables the organization to set available prices that are consistent to the level of present demand for the product or services. At the same time, the availability of diverse prices for different demand targets will increase the overall demand captured by the operator. This appears related to the underlying relation of customer demand and price of the demand curve. Finally, under this theory when the customer demand is high, social acceptance for an increase in price will be high, and vice versa.

For Trenitalia case, the core of the social issues are related to the fact that it is a national state company, willing to serve all citizens while providing adequate and affordable services; therefore, the stability and fairness of prices and seats allocation rules should be guaranteed. Hence, the introduction of RM has been careful and gradual, and was preceded by an adequate preparation: in 2005 the ‘Eurostar Italia’ HS brand was well-known, reservations were mandatory and some price differentiation and experiments had already been done. This was a solid starting point for the social acceptance of a YMS. By fact, careful RM choices together with service quality improvements helped Trenitalia to prevent negative customer reactions. On the other side, by fact the YMS adoption contributed to social purposes by letting low fares available for low-pay customers while extracting value from the high-pay market segment, while providing the national train operator with positive economic results over the years.

It shall be noted, however, that the YMS was implemented on the high-speed routes, for which the target demand is different from other services. This lower price sensitivity of the customer target of the initial implementation of the YMS at Trenitalia could have mitigated the customer reactions. This was considered in the initial study that preceded the implementation. The choiche of the right perimeter for the implementation, therefore, could have been one of the factors that allowed for a smooth implementation of RM from the customer side.

Different is the case of other sectors, where RM implementations that didn't respect the specific nature of railways resulted in a failure. As described in [Mitev (1998)] this happened, for instance, at SNCF (French national railways), where Sabre tried the implementation of a new Computer Reservations System (CRS) in use in the airline sector. Such distribution system included yield management system techniques which didn't take into consideration appropriately the specificities of the sector, resulting in a shock for travelers and consequent protests. As reported by [Link (2004)], something similar happened at DB (German national railways) when there has been an attempt to change dramatically the pricing system with the introduction of a new one, called PEP, designed as a yield management system similar to the one used in airlines.

5.1.3 Multi-leg Topology and Bottlenecks

It has been described previously how the multi-leg topology is a key feature for the rail sector. This is particularly true at Trenitalia where the average number of O&Ds of a single train service is 36. It was decided to explicitly model this characteristic, which acts as one of the levers for optimization, together with pricing.

Figure 5.1 on p. 103 shows the relationship and the numbering conventions used for YM route modeling at Trenitalia. The example refers to the hypotetical route Milan-Naples, characterized by the sequence of stops: Milan, Bologna, Florence, Rome, Naples. To simplify the examples, cities and stations coicide, while by fact major nodes comprehend a set of stations stops.

Here, the model decisions are performed per O&D, while capacity constraints are defined per leg. The solution of the dual problem provides the shadow prices, that are summed up per legs and used as input for the nesting. Therefore, while the bid price approach considers them directly in the computation of displacement costs, here the difference between price and shadow prices is incorporated in the nesting order. This will be described later on in the following part of current Chapter.

As for bottlenecks, here we define them as the legs where the demand of O&Ds traversing these legs exceeds the supply on systematic basis. The YMS provides detailed information on bottlenecks, that can be used internally as well as shared to other departments and companies to undertake at higher level corrective actions, e.g.

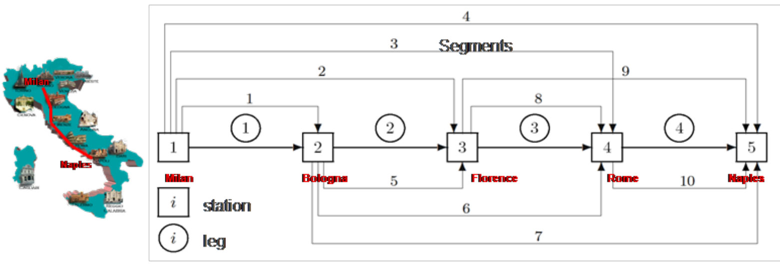

Figure 5.1: The train topology: relationship among stations, legs, segments, and their numbering convention.

on supply planning, to balance demand.

5.1.4 A Fare Family Clustered Structure

During the initial analysis phase, it was found out that the number of fares at Trenitalia was extremely large (more than 200). This was also one of the major factors of demand data sparsity, which will be discussed later in the dedicated Chapter 5.1.5. It was decided that the partition of demand into clusters—called ‘Categories’— was the best solution to overcome sparsity during both forecast and optimization. Therefore, one of the preliminary steps has been to group the fares into a certain number of categories for each Class. Categories are a partition of the fare set, or disjoint sets of possible ‘fares’ within a ‘Class’, such that there should not be any fare not belonging to any Category.

A stable and meaningful segmentation was implemented, based on behavioral analysis as well as similarities in price and fare restrictions, together with an adequate numerosity repartition. A clustering analysis (periodically redone) was at the basis of the division of the fares into Categories, actually 15. Also social issues have been taken into consideration to be able to grant availability to certain fares based on social purposes or constraints — and not only revenue-related. In details, due to commercial policy and to allow for a smooth introduction of price discrimination through demand selection, the possibility to ‘associate’ categories has been added, so that the associated categories will have the same nesting order of the highest one while preserving their own forecast.

A good clustering should:

(i) have products of *similar unit value* inside each category, that behave almost similarly and have a reasonable volume of bookings;
(ii) produce a *reasonable number* of categories, to obtain stable forecasts;
(iii) be *easily and immediately recognizable* by the operators and

within the business processes and properly 'selected' by the YMS.

Furthermore, in details, according to [Gliozzi et al. (2014)] the main factors considered were the following:

(i) Fare value similarity should be maximum within each Category and minimum between diverse categories, to ease the definition and estimation of the value per segment and category;
(ii) Number of observations shall be reasonably homogeneous across diverse categories, so that each set consists of a reasonable share of reservations over each segment and the variability of the forecasts is limited;
(iii) Priceable and fenced fares may not be mixed up within the same fare category for YMS purposes;
(iv) The classification should be consistent to the capability of the Sales or Reservation system.

Like most railway companies, Trenitalia has adopted a fare family clustered structure, which means that for each service level (also referred as 'Class', and in aviation as 'Cabin') a certain number of different fares can be available to the customers. Each service level has a distinct physical capacity, so that a passenger in class A and a passenger in class B are never competing over a seat (unless free upgrades are offered, or similar cases). Over each class, passengers from different fare Categories can be allocated. The 'Category' allowed the definition of a reasonable inventory (henceforth a related YM System) to be defined. For any class c and segment r, YMS decisions will treat equally all the passengers of category i (X_{cri}) on the same train-date. Here we refer to train-date, or TD, as a unit determined by a train leaving on a certain date, and with TDC the train-date-class, i.e. one class of a TD.

Fares were classified as Priceable and Fenced, as explained in the followings:

Priceable fares : only differentiated by the price, for a given class and segment, while they have same (or similar) terms & conditions (T&Cs);

Fenced Fares : differentiated either through access limited to specific customer groups (i.e. young, elderly) or with T&Cs that are substantially different from standard fare.

Figure 5.2 on p. 106 shows an example of the adopted family structure, as can be found on Trenitalia website. While priceable fares have almost the same T&Cs, and are open to any passengers, fenced fares are available only to specific customer groups, e.g. the owners of the frequent traveler card 'Carta Freccia', moreover some of the discounts require the passenger to be in a specific age class.

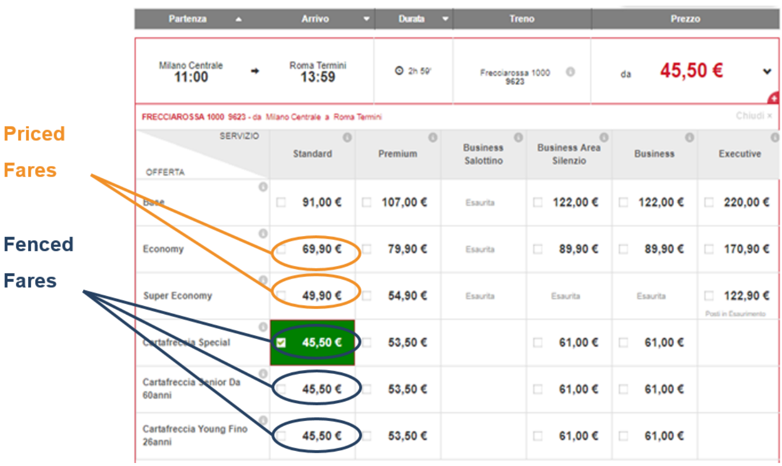

Figure 5.2: The Fare Basis structure: example of actual fares, and their possible classification as priceable or fenced fares. Source: `https://www.trenitalia.com`, retrieved on 28th September, 2018.

5.1.5 Demand Data Sparsity

This part has the purpose to explain the presence of a data set largely unpopulated, with the presence of many null values. Demand data sparsity is an issue to be treated for any of the purposes of a YMS. For the rail industry it is related to its multi-leg topology and the presence of a high number of fares, furthermore to the time dimension.It shall be noted that here sparsity does not denote a model with many null coefficients. Instead, it refers to data sparsity.

The topic was already pointed out in [Williamson (1992)] early study, stating that in a large hub & spoke airline network the number of potential O&Ds was very high, but by fact very few were requested. The author also noticed that the most requested segments were characterized by higher prices. Among the other studies, [Lewbel and Nesheim, 2010] though dealing with a non-transport problem, treat explicitly demand sparsity and provide a specific overview on the topic. Based on those works, for what concerns our case we can state the followings. Sparsity of demand is related to the fact that while the offer of segments is very high, as it increases exponentially with the number of legs offered by a train, customers typically book on a limited number of O&Ds. In addition, the vast majority of reservations are concentrated in certain fares and take place in the last days (or hours) before departure. This results in the presence of many data equal to zero, as customers book a null amount of the majority of offered segments.

In [Lewbel and Nesheim, 2010] the flaws of main (non-transport) demand forecasting methods, both continuous and discrete, is put in evidence when dealing with sparsity. In details, major models of reference are:

(i) continuous demand models from Deaton and Muellbauer (1980), a traditional approach able to consider joint purchases, or bundle of products, defined in continuous quantities;
(ii) multinomial choice models studied in Berry, Levinsohn, and Pakes (1995), that discretizes purchases, and defines each of them as independent decisions, modeled as a multinomial

choice.

For both, a major flaw is represented by the fact that their underlying assumptions can be violated, more likely, in empirical or real world applications and especially in presence of data sparsity. Furthermore, choice models often struggle to manage complementarity of products or services. Finally, the presence of demand sparsity makes either intractable or nonexistent both approaches to overcome assumptions, which were violated for discrete and continuous models.

[Lewbel and Nesheim, 2010] also define a novel model, which is able to deal with substitution patterns and remains tractable in those cases; it nests, as special cases, choice based and classic continuous demand models and traditional continuous and discrete demand mixed models. In this view, another possible response to data sparsity could be the aggregation of demand in groups. The mentioned work reports the effect of a variation in price of certain products, after the introduction of a tax with asymmetric effects on part of them, and demonstrates how it affects the diverse products in the market in different ways. It concludes that the elasticities of demand were different and, therefore, an aggregation of demand could be possibly biased and misleading, unless the aggregation follows strict conditions. Furthermore, considering bookings at aggregated level does not provide information on impacts at atomic level and on the heterogeneity and variability of disaggregated responses. Therefore, such aggregations won't provide the demand determinants at a sufficient detailed level, consistent to industry needs of the real world. In particular, for what concerns our case, it should be reminded that a YMS needs to receive input with a proper granularity.

As for Italian HS trains, they serve on average 36 O&Ds each, the offer is split in 2 to 4 classes, fares are grouped in 15 categories for each class, and reservations take place during a booking horizon of up to 180 days (here divided into intervals, delimited by snapshots or points of observations). The demand is concentrated in certain combinations of O&Ds and fares and takes place in certain periods of the booking window, more than in others. In addition, bookings level and path are affected by seasonality (comprehending month of the year, day of the week and time of the day). Those combined factors result in a high degree of sparsity, which can impact on accuracy and volatility of forecasts in less populated clusters or snapshots. Therefore, and taking into consideration the choice to use O&Ds as levers in the optimization models, sparsity arises as a complicating factor, specific to the rail industry and related to its multi-leg topology. Corrective actions that have been undertaken to mitigate this issue comprehend:

(i) clustering demand and subdivide it into categories; they are made to promote a stable and uniform repartition of demand, as well as respect business requirements and social constraints;
(ii) possibility to associate O&Ds into one if the origins, or destinations, are close (within a certain threshold in kilometers);
(iii) possibility to associate categories under certain conditions;
(iv) aggregation of the booking horizon time interval in points of observation, or snapshots.

By fact, no further aggregations are deemed convenient to ensure the forecast is at the granularity level which can properly feed the YMS optimizer module.

5.1.6 Decisions Volume

Among specificities of rail YM, a main point is represented by decisions volume which is much larger than in other business contexts, resulting in a computational challenge. This, along with the real time nature of the problem and the consequent need for quick results, has impacted the decision taken on the models to adopt.

The current number of inventory decisions to be managed each day on average for HS trains at Trenitalia can be estimated as follows: 260 daily trains $\times$ 180 days of booking horizon $\times$ 36 segments $\times$ 3.2 classes $times$ 14 fare basis per class, equal to $75,059,040$ daily.

Using the same measures, by contrast, the largest US Airline operator (American Airlines)[2] in 2017 offered 6,700 flights per day on average (including the regional partner American Eagle), with a booking horizon of 365 days and 21 fare basis shared by all the classes. This would bring to $51,355,500$ inventory decisions, 32% less than Trenitalia High Speed trains.

Therefore, it was considered the large number of decisions and other business characteristics, such as booking patterns per market over time and operational procedures. Consequently, the models were designed to ensure maximum speed and responsiveness and it was decided not to reoptimize all trains all times but to evaluate daily the ones that needed it most. For this reason, an algorithm has been implemented in order to prioritize and select which trains-classes-dates should be re-optimized each night through a batch process. Here the decision unit is the single 'train-class-date', in the followings TCD, a combination given by one class of a specific commercial service for a certain departure date. Higher priority is automatically assigned to trains-dates-classes which are closer to the departure, have never been optimized (i.e. a "new" TCD for which sales have to be opened the day after) or have been impacted by a schedule change. Furthermore, other algorithms have been implemented to allow for automatic re-optimization of certain sets of trains, e.g. leaving on the same day or during peak periods and weekends. A full parametrization of the re-optimization is available to the analysts.

Every night these algorithms run and up to 10,000 trains-dates-classes are selected and elaborated, then the new controls are automatically sent to the reservation system if the levels of the 'criticality' parameter, determined by another algorithm during the process, results below a threshold set by the analysts per train, class and departure date. The estimation of the quantity of decisions computed

---

[2] Details can be found in: American Airlines Group - About us - American Airlines. (n.d.). Retrieved December 29, 2017, from `https://www.aa.com/i18n/customer-service/about-us/american-airlines-group.jsp`

Table 5.1: Estimated volume of decisions taken during a 32 months observation period, 2013-2015

Period	32 months
Trains - dates	3,475,572
Trains - dates - classes	9,463,952
Trains - dates - classes - segments	310,977,500
Trains - dates - classes - segments - fares-basis	4,353,685,000

and actuated by the YMS during 32 months is shown in Table 5.1 on p. 111.

5.1.7 Inventory Properties

The reservation system of a travel company requires an adequate ‘inventory’ sub-system to work effectively ([Kouro et al. (2012)]) and provide a basis to manage availability of seats for bookings. It defines a set of keys and features that are required to be stored and processed by an existing inventory system, so that the yield management controls become effective. The functionality of an inventory system is to compute the ‘availability’ of seats on request, so that the inventory level can be updated, computing the seats that are available for sale ([Goyal, 1974]).

According to [Kouro et al. (2012)], the inventory system will be required to provide consistency so that:

(i) Available seats do not exceed capacity;
(ii) Availability computations must be resilient to alterations in decisions of seat reservations and not linked to the arrival order;
(iii) Availability computations must withstand changes in the capacity and not influenced by arrival of the capacity;
(iv) Availability can be determined or computed anytime and expressed as a number.

The careful understanding of fundamental concepts like ‘leg’, ‘segment’ and ‘category’ is the base for the design and implementation of an adequate Inventory system. By fact, they will determine the least number of attributes that will be required for a proper functioning of the Inventory system data base. In addition, within a transport YM system, it is important to properly select the topology of the model used, according to [Roundy (1985)].

Based on the same [Roundy (1985)], the network consists of nodes that are linked by legs. Supply is based on legs, that connect two nodes, departure and arrival, without intermediate stops. On the other side, demand comes up by nodes pairs within a preferred routing context.

Based on [Kouro et al. (2012)], the implementation of different inventory environments leads to diverse reservation systems. A ‘single-leg’ reservation system is the simplest form, founded on the physical availability to customers. It requires in any case an inventory system that is able to compute the difference of capacity and reservations. On the other side, with a multi-leg or Multi-fare environment, the reservation system becomes complex. In this second case, there is a need to deal with the topological structure occurring within the train-route, but also alongside fare classes where seat limitations can be imposed.

A better complexity, and concrete solution, can be achieved through a consideration of the whole network referred to as origin-destination (O&D) in [Roundy (1985)]. This O&D solution is capable of optimizing services because demand is on routing. An overlying O&D solution can support an advanced YM system in transportation industry. However, it is the most demanding for human and machine resources, particularly in a rail network where the O&Ds are expected to increase quadratically with legs.

A possible strategy is through considering a linear sub-network that is made up of nodes and legs related to a single routing of a particular transport system. This is a typical case of a multi-leg topology that can be used where the network is not complex like in long distance rail transport system. In this regard, the OD will correspond to a segment that provides the link between one node and another ([Parla (1988)]). Assuming that n is the number of nodes, the total number of legs will be provided by $n+1$. Similarly, the number of segments will be obtained from $(n+1) = 2$. This means that the number of elements is similar to that in the upper triangular matrix that shows all possible connections. Also, it is possible to use incidence matrix to calculate an existing relationship of legs within a segment through adopting the concept of 'proper' segment. This simply represents an O&D in a particular train route, which can be 'Inow', 'Outow', and 'Transit'.

A better segment will incorporate numbers $r = 1 \ldots n$. where n is the number of segments within a linear network. A Transit segment will assume the numbers $r = n+1 \ldots 2$, while 'Inow' segment will be represented through numbers $r = 2n+1 \ldots 3n$. An 'Outow' segment will be represented as $r = 3n+1 \ldots 4n$ to satisfy the s index observed in the incidence matrix ns so that S = 1 + (r + 1). In some situations, it will not be worthy keeping 'Inow' or 'Outow' segments to remain distinct and the 'Transit' notation to be part of the latter. Where the demand of all the segments is negligible within the linear network, it is worthy dropping this distinction so that the entire demand forecast is estimated as $r = 1 \ldots n$.

Based on the above understanding, one can properly define the elements that can be stored in an inventory system for a single train-date. Depending on the route that shows a list of stations where the train will stop, a set of legs and segments can be derived. It is possible to alter the route during times of 'Schedule Change' occasioned by a change in the timetable [Zhao and Atkins (2002)]. The only requirement is a re-initialization of the corresponding train date and class set available for passengers to buy.

Ordinarily, it is obvious that a class set will be rare because of alterations in the timetable. A class-set change can be effected through re-initialization on the inventory occurring for train-date and the capacity representing every class-leg. The latter represents total number of seats that can either be sold or reserved irrespective of its category [Muckstadt and Roundy (1988)]. Capacity changes can occur as a result of a schedule changes or a prevailing 'on-line' decision observed in an operational control center ([Florian and Klein, 1971]. As such, several types of controls can properly be proposed so that they can be used in computing 'availability' value [Zhao and Atkins (2002)].

Based on [Gliozzi et al. (2014)], the number of reservations X performed for each fare class C within segment r and category i can be computed. Such values must be continuously updated through data collected in the sales and reservation system to become a function of the inventory system. 'Controls' present a subset of parameters when combined with the others will enable the inventory system to subsequently determine the value of 'Availability' of seats for all the segments c, r and i achieved on the train-date. Here, the relative stability of the inventory across the booking window provided a good basis for YMS use.

5.2 YM Models and System Architecture

The YMS overall work flow is highly automated and the system is able to autonomously manage the routine, computing and implementing decisions. Every night batch processes update the information from the Reservation, Inventory and Pricing Systems through the YMS database. The data download procedures cover PNRs (reservation codes, here identifying the passenger and related itinerary), capacity, controls (defined later in this Chapter), schedules. Secondly, an algorithm sets a ranking of the combinations of trains and departure dates, that are open for bookings, prioritizing the ones:

a. where a schedule change recently impacted the offered capacity;
b. just released in the reservation system;
c. with a departure date close to the elaboration date;
d. not elaborated for long time;
e. queued for elaboration, on purpose, by the analysts;

f. other criteria (e.g. if the actual sales levels look very different from the last forecast).

The first 10,000 of the ordered list are selected by the system every night to perform the automatic re-optimization[3]. Then, as displayed in Figure 5.3 on p. 116, a chain of YMS models is operated on each selected train-class-date to compute the new availabilities. If a set of conditions is met, the YMS automatically sends to the Reservation System (RS) the controls which implement the authorization.The re-optimization can also be launched manually by the analysts in their daily activities.[4]

In both cases, three models are concatenated to compute and evaluate the effect of the 'controls' which determine the availability, as presented in Figure 5.4 on p. 117 and described below:

Unconstrained Forecasting Module : the potential or unconstrained demand is estimated per each combination of segment and category, based on historical data (booking profiles) and advance data (actual bookings) through a scenario-based stochastic model, not considering the train capacity as a constraint;

Optimization Module : the model partitions the capacity into '*protections*', allocating the unconstrained demand to maximize the weighted average of the 'utility' (a combination of revenue and LF). Moreover, from the duals, it defines a '*nesting order*'. The resulting protections and nesting orders, combined with capacity, determine the 'authorizations' and, subtracting the booked passengers, the availability: i.e. the number of available seats per segment and category;

---

[3] The upper limit of 10,000 trains has been set in relation to compatibly the hardware and software architecture, to ensure that the overall re-optimization process has a proper timing, allowing any other action from the system—from operations to analyst inquiry—to be finalizes during the working hours. In reality, the limit becomes effective only few times per year, when large schedule changes are run.

[4] The set of batch procedures that run overnight comprehends the download, transformation and elaboration of data on PNR, capacity, controls, schedules. Automatic controls are sent to the reservation system if the value of the 'criticality' parameter is below the user threshold. Business Rules are also applied.

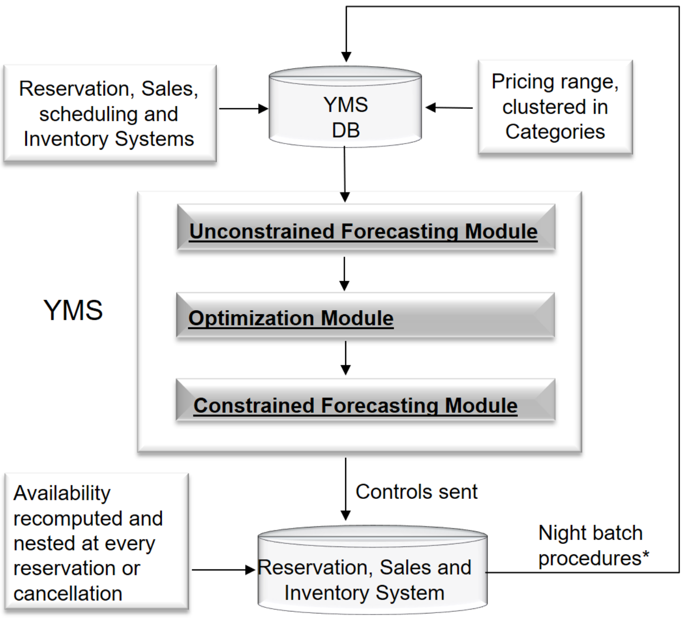

Figure 5.3: Overview on YMS models

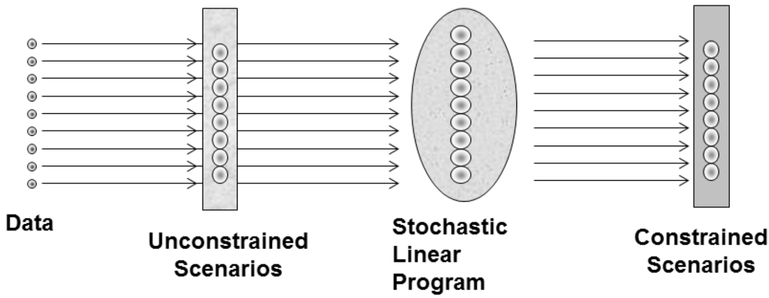

Figure 5.4: A schematic figure of the scenario based chain of models. In this case the history of 9 trains (historic scenario, single class) is processed in the un-constraining forecasting model, producing 9 sets of ODF forecasts for the class which are input to the optimization and simulation models.

Constrained Forecasting Module : the constrained demand forecast is computed simulating the unconstrained demand forecasts to be a queue and considering the optimal allocation within the availability computed by the Optimization model.

Summarizing, the unconstrained scenarios are the input for the scenario-based two stage stochastic linear program, which will determine the 'controls': protection and nesting order. The full information is then passed to a simulator, which transforms each unconstrained forecast into several potential booking queues, with order of arrival permutations, and acts using the same logic of the booking system in accepting the single elements of the queue (e.g. booking requests). This produces a forecast, that is 'constrained' by the decisions taken in the optimization model.

5.3 Potential Demand Forecast

Forecasting demand with a satisfactory degree of accuracy is of paramount importance for the quality of the YMS decisions, yet it is a challenging activity. The travel industry is a complex reality and the history of reference is made up by many repeated bookings with a relatively small single value. The difficulty to forecast demand can be related to: (i) demand-specific factors, (ii) data sparsity, (iii) data truncation, as explained in the followings.

Firstly, among demand-specific factors we can list:

(i) seasonal fluctuations (e.g. per day of the week and time of the day);
(ii) special events;
(iii) sensitivity to pricing actions;
(iv) demand dependencies between fare classes;
(v) group bookings;
(vi) cancellations;
(vii) defections of passengers from delayed services;
(viii) no-shows (booked passengers that didn't show up at departure, but didn't cancel their reservation);
(ix) go-shows (passengers that didn't book before the train departure but simply showed up for boarding);
(x) recapture.

Secondly, in the railway industry, specifically, for many classes-segments-categories historical data consist on few bookings and many scenarios are not even populated: such sparsity doesn't help predictions accuracy.

Thirdly, as explained in Chapter 3, one of the main problems in forecasting the actual demand is to overcome the 'truncation' or 'censoring' issue. In fact, during the booking horizon, reservations are accepted until booking limits are reached. Then, the system stops accepting reservation requests and collecting data on that specific demand.

Therefore, as explained in [Berto and Gliozzi (2018)], here to forecast the demand per each train, date of departure, class and combination of segment and category, the estimation process has been divided in:

(i) understanding possible demand behaviors from the reservations in past scenarios;

(ii) estimating the potential demand occurring during each period of the booking horizon that was closed to reservations as the demand for travel exceeded the booking limits (or capacity).

Here the demand is not deterministic, but its stochasticity is taken into consideration. It regards the advance of purchase per fares and classes, and last but not least the order of arrival of the distinct cluster passengers. Differently from other models, here the LBH is irrelevant and it is not a given that low-pay customers book earlier that the others.

As listed above, the first step is to understand possible demand behaviors. It starts by taking 'snapshots' (or 'readings', or 'points of observations') of historical reservations of an already departed train, at predefined moments. This allows to build, for a given ODF, on several departure dates, the so-called 'booking profile', that is the graphical representation of the sequence of reservations observed and summed up at determined Points of Observations (or snapshots) of the booking horizon, for a certain train.

Figure 5.5 on p. 120 shows an example of booking profile, one for each historical scenario on a single train and segment. In particular it displays the booking profiles of one train for the segment Milan – Naples, in 2^{nd} class, on Fridays. The two historical scenarios with the steepest curve before -40 days to departure represent the scenarios before Easter and Christmas. They become quickly fully booked, hence the second part of their booking curve is flat as either the capacity or the booking limit had been reached. In most of the other scenarios, instead, the booking profile curves become flat from about 5 days to departure. Here, in the forecast it is possible to keep exact track of the soldout phenomenon, as also presented in [Berto and Gliozzi (2018)].

A similar curve displaying, stacked, all the segments which traverse a given leg in a single date, is shown in in Figure 5.6 on p. 121. They are summed up per leg to display how the physical capacity,

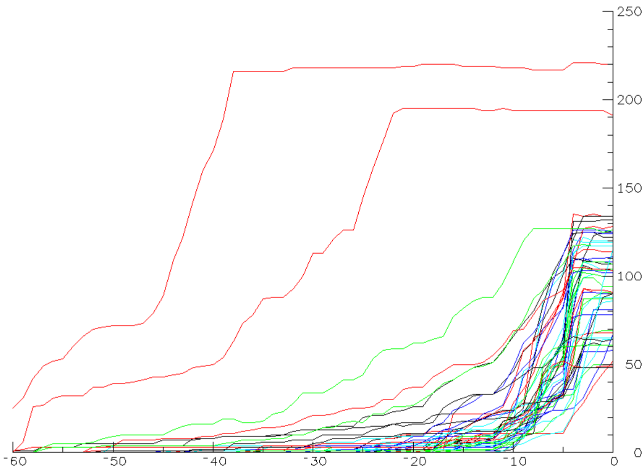

Figure 5.5: Example of booking profiles. Historical scenarios of one train, segment Milan – Naples, 2^{nd} class, on Fridays.

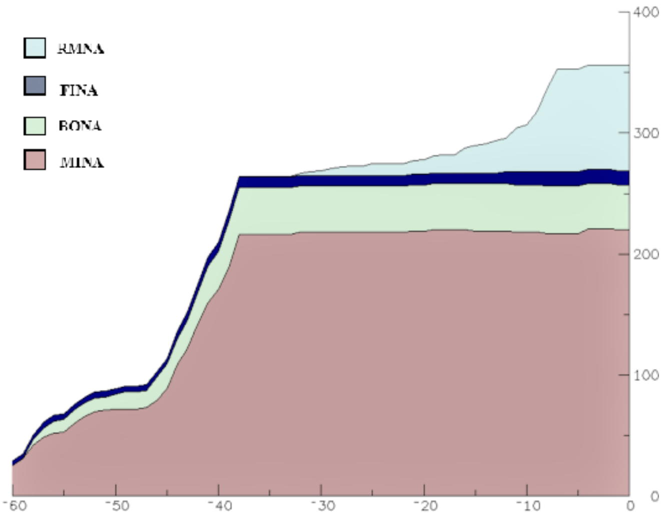

Figure 5.6: Example of booking profiles: stacked booking curves of the 4 segments traversing the Rome-Naples leg of one train (Milan – Naples, 2^{nd} class, Fridays).

offered per leg, is saturated by customers with diverse itineraries traversing that leg. Again when the curve becomes flat either all the segments have reached their booking limit or, as it was in this case, the physical capacity of the train is saturated.

Even if this is not represented in the graphs, it should be noted that the model includes snapshots after departure of the train (e.g. t=0.5, t=1). This allows to take into account the events—e.g. reservations and cancellations— taking place after the departure. Due to the multileg topology of the rail industry, arriving to the final destination may take some hours and stops (up to 10 hours and a dozen stops for longest trains). Therefore, in a route there are segments that can be booked, or cancelled, after the train departure and a high rate of reservations, cancellations, reimbursements, go-shows and no-shows, and other typical phenomena take usually place after departure.

The YMS includes also a computation of the actual 'mortality' per segment and category on each train, date and class using a similar criterion. Here mortality is defined as the number of reservations in place at a given snapshot that, most probably, will be canceled over the remaining period before departure. It is often used in aviation to manage overbooking; however, in our implementation overbooking is not used, due to the Company's policy in favor of the customer and the presence of safety restrictions.

5.3.1 Estimating the Unconstrained Demand

As explained in [Gliozzi (2006)], the forecasting model needs to estimate the potential (or 'unconstrained') demand in order to feed the optimization model with a more accurate estimation of the demand. The latter has already been defined as an estimate of the demand without the constraints of offered capacity and inventory controls on the train. The un-constraining process aims to estimate the real demand behavior from the available historical booking profiles, whenever a given segment-category was not open for sale and a customer was not able to book due to the unavailability of the desired service; in this case, the information on the real demand is lost.

In the implemented system, the un-constrained part of the demand, (also defined as 'emphasis') is computed as a function of the bookings that occurred in other historical scenarios, during the same period where the considered scenario was censored ('gradients') and added to historical values. The un-constraining method is based on the computation of the derivative of demand during the part of the booking window when there was some availability. The differences between consecutive photos (or snapshots) are not considered for this purpose, rather than the absolute values of demand. They are added to the booking profiles as potential finite differences.

The graphical representation of the logic behind the full method, also explained in [Berto and Gliozzi (2018)], is in Figure 5.7 on p. 124.

The abscissa represents the time from bookings opening up to departure, divided in ‘snapshots’, or ‘photographs’, while the ordinate is the demand X. The time of the departure, when the train leaves from the first station of its route, is d (usually zero), while the present time when the forecasts are computed is t. The dotted line represents the actual booking history of a train-date-class (TDC) on a specific ODF, or segment-category (that is going to be un-constrained); it stops at t which is the moment when the un-constrained forecast is performed, while its prosecution (the continuous line) is the unconstrained forecast from an historical scenario.

Let us define $\bar{X}$ the historical scenario, and X the result of our unconstrained forecasting. First of all, the estimate of potential demand at departure X_d adds to the component $X_t - \bar{X}_t$, which is already acquired. Secondly, it computes the ‘emphasis’ for the period(s) when the category was closed to sales—in our example from the following photograph (or snapshot) and time r. Then, the algorithm estimates the gradient of the demand path in each snapshot interval; this is done as follows. Let s identify a generic scenario, $\hat{s}$ the reference scenario, k a snapshot and $d_k s$ the duration of the closure interval interval from $k-1$ to k; the gradient $\bar{X}'_{ks}$ is then computed as follows:

$$\bar{X}'_{ks} = \max(0, \frac{(\bar{X}_{ks} - \bar{X}_{(k-1)s})}{d_{ks}}) \qquad \forall s \tag{5.1}$$

This will clearly provide positive values only for the snapshots with a positive gradient. Whenever a ‘closed for sale’ scenario is found in a snapshot, the quantity

$$E_{k\hat{s}} = \max_{\forall s}(\bar{X}'_{ks}) d_{k\hat{s}} \tag{5.2}$$

will be added as the emphasis for the snapshot. Other details on the general method, and possible pitfalls when there are cancellations or the closure duration has to be estimated, can be found in [Gliozzi and Marchetti (2003)] where it was applied to the air cargo context at Alitalia.

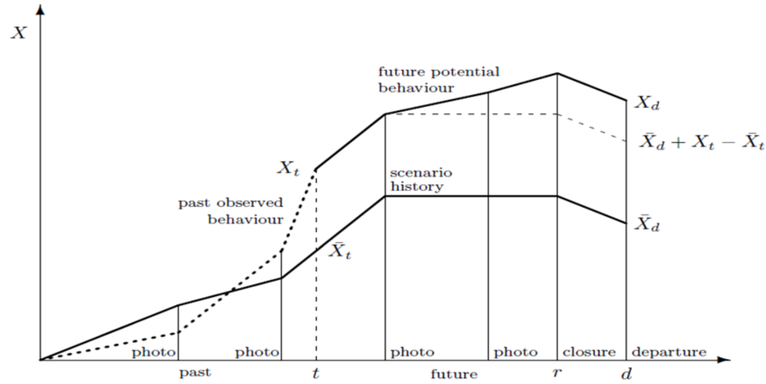

Figure 5.7: Sketch of the unconstrained forecast algorithm, at TDC and ODF level, for one scenario.

The final unconstrained forecast for the scenario will therefore be:

$$X_{d\hat{s}} = \bar{X}_{d\hat{s}} + \sum_k E_{k\hat{s}} + X_{t\hat{s}} - \bar{X}_{t\hat{s}} \tag{5.3}$$

5.3.2 A Multiplicative Correction

Initially the demand estimation implemented the purely additive method, described earlier, with satisfactory results. The method was then modify to address the necessity to detect and treat promptly anomalous demand behaviors. Outliers that can be defined as trains-dates (and in particular some of their segments-categories) characterized by actual bookings fairly different from their history. The early detection and management of outliers, as well as the possibility to alert YMS analysts to deal with them separately as exceptions, are paramount. In particular, it can reduce or avoid the occurrence of phenomena like dilution or high spill, spoilage, stifling. This was needed, for instance, at early stages of the booking horizon (i.e. up to 180 days before departure), when all promos and discounted fares allocated are available on each train and there is still the possibility to correct the decision of the system before they are sold. At the same time, in those early snapshots it is difficult to spot and weight properly the outliers as in many segments and categories few or no reservations are in place, not actually nor in the history.

Therefore, to improve the forecasts sensitivity to demand peaks, off peaks and abnormal fluctuations, a multiplicative correction has been applied to the additive method. Such correction is based on the ratio between booked passengers at departure and at time t. It proved to be a good correction rather than a method to be used *per se*, as the latter could lead to unstable and less usable results (e.g. infinite or indefinite values, divergent or highly oscillatory behavior over time). To limit any instability, an exponential smoothing lowers gradually to zero the multiplicative correction as departure approaches, so that at departure time ($t = 0$) additive and mixed forecasts coincide to, and equal, the advance (actual) bookings; this limits the risk associated to abnormal correction values.

As presented in [Berto and Gliozzi (2018)], the main formula for the multiplicative forecast is:

$$\ddot{X}_d = \frac{X'_d}{X'_t} X_t = \frac{X'_d}{X'_t} \bar{X}_t \tag{5.4}$$

where $\ddot{X}_d$ is the multiplicative forecast at departure; X' is the reference historical scenario, X_t and $\bar{X}_t$ are respectively the actual bookings at time t and the unconstrained booking forecast at time t, which are identical by definition.

Finally, a parameter (ρ) has been implemented to allow users to modify manually the weights of the additive and multiplicative components of the forecast, and analysis have been done in order to set its initial standard value(s).Let $\ddot{\bar{X}}$ be the combined prediction; it is computed using the following equation:

$$\ddot{\bar{X}}_k = \rho \ddot{X}_k + (1-\rho)\bar{X}_k \tag{5.5}$$

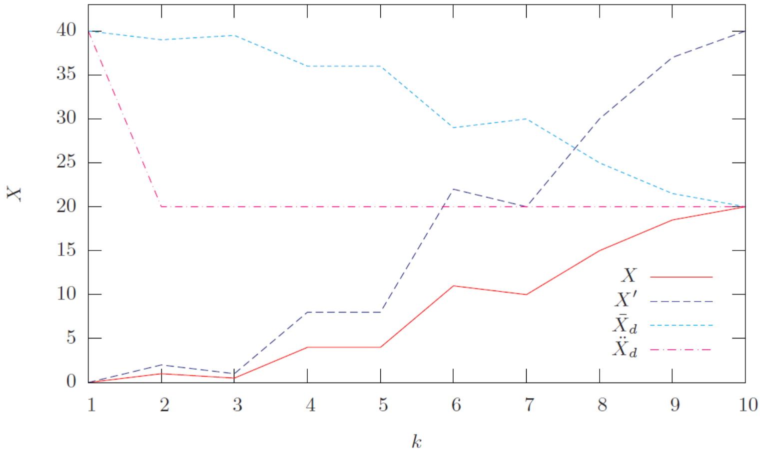

Figure 5.8: Example of Multiplicative and Additive forecasts at departure time at train-date-class and ODF and scenario level.

Figure 5.8 on p. 126 presents an example of forecasts with the diverse methods. The graph represents the possible speed of adaptation of forecasts at departure time compared to actual bookings during the progression of the booking horizon, from sales opening (snapshot 1) to departure time (snapshot 10), using different forecasting methods, for a certain ODF) of a TDC where advance bookings are below historical data. In details, the abscissa is the progressive number k of the snapshots, each taken at a certain time before departure, while the ordinate represents the absolute value of booked passengers. The continue red line represents the progression of the actual, or advance, bookings (e.g. at the time of snapshot 7, the advance booking was of 10 passengers); this is also reported in [Berto and Gliozzi (2018)].

5.3.3 The Weights of Scenarios

In the adopted methodology, "the potential demand is estimated from elementary data of historical scenarios; it works in a similar way to a smoothing approach, but the weights of the scenarios are here pre-determined, valid across all scenarios and applied to elementary data of each scenarios" as reported in [Berto and Gliozzi (2018)] on p. 274. The scenarios are a historical series with a weight associated to each element. As outlined in [Berto and Gliozzi (2018)], two are the factors that determine the weight of a past scenario:

(i) The difference in seasonality, which relates to the degree of similarity of paths and current levels of bookings of a train/date with respect to the past scenario(s) of its history. Two most relevant kind of seasonalities have been identified, related to month of the year and day of the week; the first weight is the combination of the two, which are multiplied. In most cases if the day of the week is the same the resulting weight will be 1, and 0 in the other cases. As for the hour of the departure instead, it is implicitly taken into consideration: by default, the YMS tends to search the train of the past which is most similar to the current one for departure hour, among the set of trains in the history with same route and ordered set of stops. The weight of seasonality for the j^{th} scenario is denoted as:

$$w_{js} = w_{jm} w_{jf} \tag{5.6}$$

where w_{jm} is a fixed weight, which measures the similarity between the departure month of the forecasted train and of the j^{th} scenario; w_{jf} is a fixed weight that measures the distance between day-of-week of departure of the forecasted train and of the j^{th} scenario of the (5.6) on p. 127.

(ii) The distance, measured in days, between the date and time when forecasts are made and the date and time of departure of the train in the past scenario. This second weight is computed with an hyperbolic function:

$$w_{jg} = w(g,m,e,k) = m + \frac{1}{\left[\frac{g}{k} + \left(\frac{1}{1-m}\right)^{\frac{1}{e}}\right]^{e}} \tag{5.7}$$

where m, k, and e are system parameters and g is the difference in days between the forecast date and the departure date on the j^{th} scenario. In particular m should be such that $w_{j0} = 1; w_{j\inf} = m$ and $0 < m < 1$.

The two weights are then composed in the raw weight w_j of the j^{th} scenario, through a system parameter $0 \le h \le 1$ such that :

$$w_j = hw_{js} + (1-h)w_{jg} \tag{5.8}$$

Finally there is the need to normalize the weights, since the number of available scenarios in the history may not be always the same. The sum of all normalized weights should be 1; the normalized weight $\bar{w}_j$ will therefore be:

$$\bar{w}_j = \frac{w_j}{\sum_{j=1}^{m} w_j} \tag{5.9}$$

Using these normalized weights, the mean forecast of any given category c will simply be:

$$\hat{X}_c = \sum_{j=1}^{m} \bar{w}_j X_{cj} \tag{5.10}$$

This method provides results that are completely comparable with those obtained with the exponential smoothing models and show a substantial concordance of the two methods, even when the observations are very variable, but this is much simpler than smoothing.

Figure 5.9 on p. 129 shows an example of weights over a set of eleven scenarios of the same day of the week.

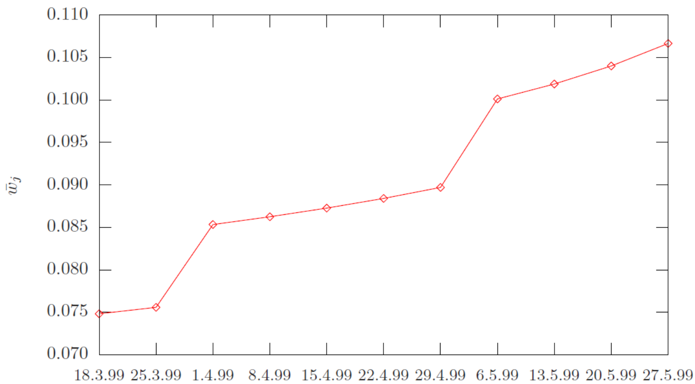

Figure 5.9: Example of weights over a set of eleven scenarios — same day of the week

5.4 Optimal Demand Selection

In our implementation, the optimization method is used to select, for each segment-category (K_{sc}), the 'protection' quantity which will lead to the combination of passengers which maximizes revenue or other parameters, at the end of the booking horizon. The 'protections' are described in [Williamson (1992)] in the context of small airline networks. Since a train can be considered as a small network, the availability computation was built along Williamson's guidelines, but here it is also defined a variable nesting order with possible ties.

Here, the optimization approach is a two stage stochastic model, which gets in input the forecasts for each scenario, without any strong assumption about demand distribution. Based on elementary data (60 observations or booking profiles per train on a certain day of the week, corresponding to 1 year and 2 months of historical data). All potential demand forecasts for each scenario are considered (60 profiles *times* the number of segments and categories), as input for the decision-making step. They are associated with their respective probability, which is pre-assigned based on time proximity and seasonality (day of the week and month of the year). Thus, co-variation among segments and categories is implicitly taken into consideration. Demand forecasts at departure are then rounded to the nearest integer; in this way, any extreme point will be integer even if the model is formulated as a Continuous Two Stage Stochastic Linear Program.

In details, the Stochastic, Scenario based Linear Program maximizes the total weighted revenue of all scenarios for each TDC, taking into account demand variability and determining a single partition $\vec{K}$ of capacity, valid over all scenarios, to allocate the unconstrained demand.

We will introduce the optimization model presenting initially its deterministic version, then the adopted scenario-based one which takes into consideration the stochasticity of the demand and implements some corrections to (i) ensure that capacity constraints (i.e. inventory) are duly taken into consideration and (ii) round the resulting number of passengers to the nearest integer. The basic *deterministic* optimization model, to maximize revenue under a demand which is given, is fairly simple in its logic and it can be represented as:

$$\max(z) = \sum_{i} \sum_{r} \check{X}_{ri} v_{ri} \tag{5.11}$$

$$\sum_{r} \delta_{r\ell} \left(\sum_{i} \check{X}_{ri} \right) \leq C_{\ell} \qquad \forall \ell \tag{5.12}$$

$$\check{X}_{ri} \leq \ddot{\check{X}}_{ri} \qquad \forall r, i \tag{5.13}$$

where r, i, ℓ are respectively the indices used for segments, categories and legs; v_{ri} is the unit value of a reservation for segment r and category i; $\check{X}_{ri}$ is the decision variable representing the number of bookings accepted for each segment-category, which is bounded in (5.13) by the corresponding demand forecast $\ddot{\check{X}}_{ri}$. Moreover $\delta_{r,\ell}$ is the $[0,1]$ value of the incidence matrix between segment r and leg ℓ. (5.11) aims at maximizing the total value of bookings, which are limited not only by the demand forecasts, but also by the capacity C_ℓ on each leg (see (5.12)).

This simple deterministic model has been expanded into a two stage stochastic model. Moreover, a couple of features have been added upon request from the users:

(i) The stochastic model aims at maximizing a linear combination of revenue and Paxkm[5], with coefficients respectively $(1-\alpha)$ and α. By fact the weight of this second part of the objective function is kept very small;

(ii) The model is not required to partition the whole train, but only the necessary part of the capacity;

(iii) It is accepted to approximate the forecast demand $\ddot{\check{X}}_{ris} \forall r, i, s$ rounding to the nearest integer. This allows for building a Continuous model rather than a Mixed Integer one and shorten the overall computation time. Therefore, the performance requirements related to the real-time nature of the problem are met.

Under these new requirements, and keeping in mind that: s will be used as the 'scenario' index in the scenario based stochastic model; $\overline{w}_s$ is the weight (or probability) of scenario s occurring; Λ_r is the length in km of segment r; it results the following formulation:

$$\max(z) = \sum_i \sum_r \sum_s \check{X}_{ris}\Big(\overline{w}_s\big((1-\alpha)v_{ris} + \alpha\Lambda_r\big)\Big) + \sum_\ell U_\ell \varepsilon - \sum_\ell O_\ell \zeta \tag{5.14}$$

[5] Passenger Kilometer, or Paxkm, is a measure of demand where each passenger is multiplied by the length in kilometers of their own itinerary

subject to:

$$\check{X}_{ris} \leq K_{ri} \qquad \forall\, r, i, s \tag{5.15}$$

$$\sum_r \delta_{r\ell} \left(\sum_i K_{ri} \right) + U_\ell = C_\ell + O_\ell \qquad \forall\, \ell \tag{5.16}$$

with bounded real variables:

$$\min\left\{X_{ri}, \mathrm{round}(\ddot{\check{X}}_{ris})\right\} \leq \check{X}_{ris} \leq \mathrm{round}(\ddot{\check{X}}_{ris}) \qquad \forall\, r, i, s \tag{5.17}$$

$$-\infty \leq K_{ri} \leq \infty \qquad \forall\, i, r \tag{5.18}$$

where K_{ri} is the 'protection level', the decision variable of the first stage, which is unique across all scenarios. $\check{X}_{ris}$ is the number of bookings accepted for each segment-category and scenario. The capacity constraint (in the former deterministic model (5.12)) is now transformed into (5.16). In this latter constraint it is foreseen the use of explicit slacks, which are also weighted in the Objective Function:

U_ℓ with a very small positive coefficient ε, which will keep the sum of the relevant $\vec{K}$ as small as possible;

O_ℓ with a ζ coefficient, negative and very large in absolute value, which will keep the model always feasible, allowing the sum of relevant $\vec{K}$ to exceed the capacity if this is the only way to maintain feasibility[6].

The K_{ri} variables, in turn, limit the bookings of the $\check{X}_{ris}$ with inequality (5.15). K_{ri} values are 'unbounded' to be sure they will always be basic, such that the constraint reduced cost $\dot{C}_\ell$ is not null when capacity C_ℓ is matched by the sum of relevant K_{ri}. The $\check{X}_{ris}$ are bounded below by the passengers already booked and above by the integer rounding of forecasted demand per scenario.

[6] This might happen in operations if capacity is reduced below the number of already reserved passengers, e.g. because of a broken car. In such case, it is preferable to have anyway an optimization able to report a valid solution rather than handling infeasibility

At the end of the optimization process, the $\vec{K}$ values are the 'protection', while the dual values are used to compute the value of segment r category i, which is $v_{ri} - \sum_{r\ell}(\delta_{r\ell}\dot{C}_\ell)$; this is the base for the Nesting Order $\vec{o}$ computation, which will be described in the following subsection. To match the analysts parameters, the model solution is then adjusted by altering the nesting order or setting slightly different limits to the model authorization if they exceed the analyst pre-defined range. Currently the model is implemented in C++/concert, and solved by the IBM CPLEX V12.7 library, while initially the library used was the Coin-or suite.

5.4.1 The Combined Use of Route-Topology and Fare Levers

Many YM systems in the airline sector set the inventory control by fare at leg level, while others calculate the levels of seat inventory controls on other basis, such as: virtual nesting, O&D itinerary level, bid prices ([Zeni (2001)]). Rail industry, as described earlier, is characterized by a multi-leg train topology, with the presence of several segments on the same route, and the wide range of fares. By fact, the same train can be 'under-capacitated' in determined legs of its itinerary, which are therefore soldout, even if there are still some availabilities on other segments, not insisting on the contended legs. To exploit this, the Rail YM can use different levers (price selection, segment selection) to increase a mix of Revenue and Load Factor, as described in Table 5.2 on p. 135.

To show the effect of using a combination of segment and fare levers, let us focus for instance on a specific train, leaving from Milan at 16:00 every Friday. This train has been observed during the first semester of 2005 and 2006. In the first instance there was no YMS while in the second the YMS was active and performed a segment and fare selection, changing the O&D and fare mix of the booked passengers (PAX). Here it is displayed one case from the first implementation as it is the most representative; other examples, more recent, can only present results from algorithmic improvements rather than comparing a real situation without the YMS to

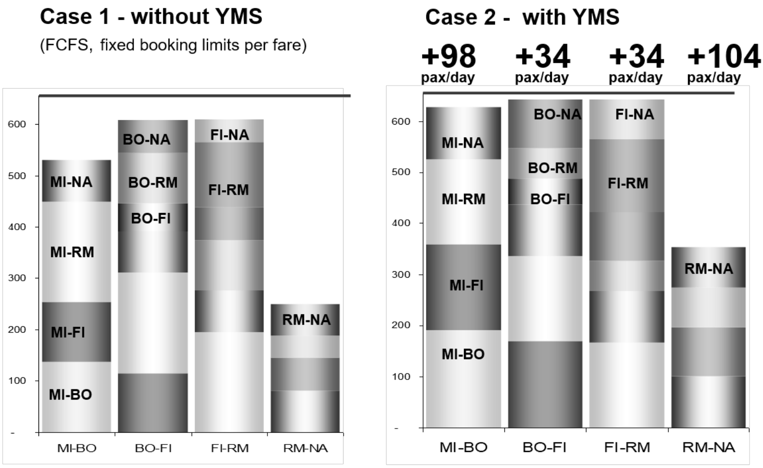

Figure 5.10: Case study: the Milan-Naples train, leaving Milan each Friday at 16:00. The train starts from Milan (MI), and sops in Bologna (BO), Firenze (FI), Rome (RM) before reaching Naples (NA). On the left, the Passengers (PAX) numbers by ODF on average in first semester 2005, before YM System introduction. On the right, the average of the fist semester of 2006, using YM System. There is a clear selection of the demand by the YM system, on the segments MI-RM, and BO-RM, to favor FI-NA, BO-NA and MI-FI, to achieve higher Load Factor and revenue.

Table 5.2: The combined use of levers: the ability of performing a selection by segment, by fare basis, or the combination of the two, allows for an effective rail Yield Management under different conditions

Situation	Lever for demand selection
Over capacity on some legs, under capacity on others: re-balance demand to the given supply	Segment
Over capacity on all legs (because of low demand): discount fares to be introduced and left open to sales	Fare basis
Under Capacity of trains in a context of many distinct fares	Segment x Fare basis
Special events to be recognized and exploited via Marketing and Sales initiatives (i.e.: Scheduling, Upgrading, Last-minute, ...)	Informative and Intelligence tools, Forecasting, Monitoring

another following the implementation.

The results are shown in Figure 5.10 on p. 134. The overall PAX and revenue growth was achieved primarily by increasing the number of passengers on segments traversing the first and last legs, allowing for a different passenger mix per O&Ds and fares. This was achieved not only by limiting the availability on some segments, but also increasing the number of discounted seats available on others, as shown in Figure 5.11 on p. 136. The lower fares were redistributed among segments in order to promote bookings on the ones traversing ‘empty legs’, here defined as legs which had a number of passengers significantly lower than capacity, like Rome-Naples. Stations falling inside the same city node are here aggregated to simplify the example. This is one of the means by which the result displayed in Figure 5.10 on p. 134 was achieved.

The overall effect of the combined use of those levers from the YMS is an increase of revenue, passengers and passengers-kilometer[7]. By fact, the system allows more availability of lower

[7] the ‘passengerkilometer’, that is the sum of the km traveled by each of the train passengers, is the standard metric used in Railways to compare trains, or to compute the ‘Load Factor’ which is the passengerkilometer divided by the ‘seatkilometer’ in the same train.

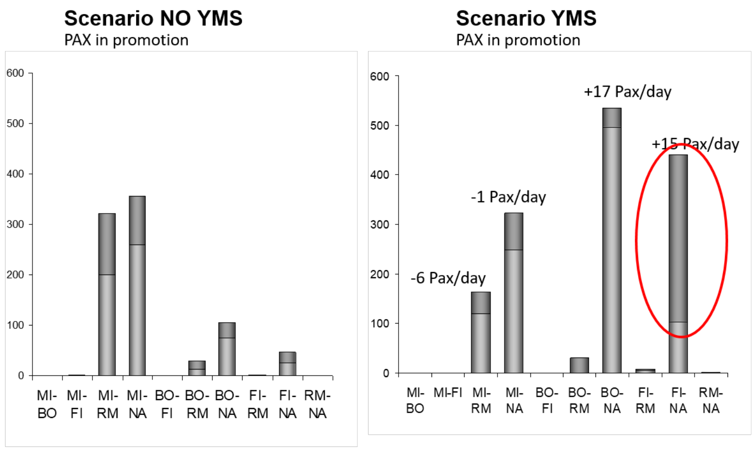

Figure 5.11: Case study: the Milan-Naples train, leaving Milan (MI) each Friday at 16:00 directed to Naples (NA), with intermediate stops: Bologna (BO), Firenze (FI), Rome (RM). On the left, the passengers number by segment, as an average of the discounted fares sold in all the departures of first semester 2005, before YM System introduction. On the right, the same average over the first semester of 2006, using the YMS. It is visible the reduction in the number of discounted seats sold on the MI-RM, and an increase on the number of discounted seats over BO-NA and FI-NA.

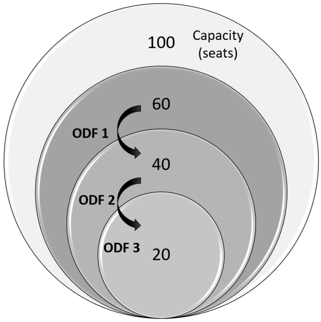

Figure 5.12: Example of Partial Nesting on a hypothetical single-leg train with 100 seats serving therefore just one segment.

fares on segments traversing empty leg(s), while accepting less discounts on the others including legs which are forecasted full.

Optimal Allocation of Seats and Partial Nesting

As described above, starting from the information on passengers already booked and demand forecasts, for each train-date-class the optimization model maximizes an appropriate combination of total revenue and PaxKm, determining the optimal partition of seats by segment-category. However, given the demand stochasticity, applying this exact partition would be risky as it may lead to unsold seats. Again, the uncertainty relates not only to the number of passengers, their partition by fare and the moment they book, but also their order of arrival: it is not known whether low pay customers will book earlier than the high-pay ones.

Therefore, a nesting approach was implemented to hedge the forecast errors minimizing the risk of unsold, which increases with the number of segments and categories and is therefore critical for the rail sector. The nesting technique allows to sell the same space to multiple segment-categories, which are the result of a breakdown in demand. It is established a ranking among combinations of segments and categories depending on their value: the nesting order. This is given by the difference of the price (i.e. the value of the segment) minus the sum of the shadow prices (i.e. a cost measure) per each leg traversed by the segment. Based on the nesting order, the most valuable combinations of segment and category (or 'segment-category') are given greater availability of space.

Specifically, here a 'partial' nesting approach has been adopted. It is aimed to guarantee that the demand with higher value occupies first its own 'protected' space and, only after it is sold out, the space reserved to the demand with the next lower value, then gradually the others. If passengers with high value book first, the higher ranked categories cannot subtract seats to the lower ones, if they still haven't exceeded the protection; so the possibility to have empty seats even on a potentially full train is minimized, muffling the forecasting errors. This is illustrated in Figure Figure 5.12 on p. 137 for a hypothetical single-leg train with 100 seats serving one segment, that coincides with one leg. Based on the optimization, the model assigns a protection of 60 seats to the first ranked segment-category, 40 to the second and 20 to the third. With the partial nesting approach, if 70 passengers are willing to book the highest price, the first 60 will occupy the space assigned to the highest ranked segment-category; only then the other passengers booking in first segment-category will use the space assigned to the 2^{nd} segment-category if still available, and so on. On the other way around, if 30 low pay customers are willing to book, only 20 will be accepted as 20 is the number of seats assigned to segment-category 3.

Based on [Gliozzi et al. (2014)], every time a booking or a cancellation takes place (so, several hundreds of times per train-date), the availability has to be recomputed in real time by the Inventory component of the Reservation system. Therefore, the complete availability computation has been designed to be simple and fast enough to be easily computed, starting from the following parameters: the 'controls' made by protection and nesting order $(\vec{K},\vec{o})$, the capacity C_ℓ and the passengers already booked $X_r i$.

The availability $\vec{D}$, the calculation of which is implemented on the Reservation system, is the number of seats still for sale at the moment of the booking, for each segment-category. In Table 5.3 on p. 141 is practically explained the algorithm to compute the availability. In that case on segment 2, category 1, the number of reservations greatly exceeded their own protected space and took all the one allocated to other segment-categories. In details, let $\eta = 3$ be the number of nodes; as a consequence $\tau = 2$ is the number of legs and $\rho = 3$ the number of segments; there are $\iota = 2$ categories and the train capacity is $C_\ell = 100\,\forall\,\ell$. Let us define $Z_{ri} = K_{ri} - X_{ri}$; $\Theta_\ell = C_\ell - \sum_{ri} K$ is the amount of seats which are not allocated to any protection. Θ_ℓ could also be negative if the protections exceed the capacity. The segments and categories are ranked based on the nesting order, from the best one to the worst. The steps of the computation process are the following ones:

(i) compute $\Theta_\ell, \forall \ell$;
(ii) compute $Z_{ri} \forall r, i$;
(iii) build $\Gamma_{\ell ri}$ by cumulating Z_{ri}, independently per each leg, starting from the Θ_ℓ and up from worst to best nesting order. In case of same nesting order, they are considered together and cumulated;
(iv) compute the availability by leg: again, starting from the last ranked, the leg availability is the minimum of $\Gamma_{\ell_r i}$ with a nesting order that is better of, or equal to, o_{ri};
(v) compute the final Williamson availability D^w_{ri}, for each ri as the minimum $D_{ri\ell} \forall \ell$.

It should be noted that, if a segment doesn't traverse leg ℓ, in the table there is the symbol "–" which represents "Null". The operator min is utilized so that $\min\{\text{Null}, x\} = x$ for any $x \in \mathbb{R}$

Here the choice, described in [Gliozzi et al. (2014)], requires to have the set of controls for the corresponding segments c,r, and i to be a quadruplet. K_{cri}, named 'Protection', is the integer rounding of the number of seats being reserved 'ideally' for a particular category i in a passenger class c and O&D r occurring on a determined train date. Secondly, O_{cri} can be referred to as 'Nesting Order', an integer number representing a hierarchy of values occurring in a TDC c, $O_{r\neq 0i} < O_{r=0,i}$. This means that $O_{r\neq 0i}$ is more valuable than $O_{r=0,i}$. All these can be computed and updated within the inventory system, using the YM system models. This number is computed and updated in the main inventory system by using the YM system model or can be derived from prevailing analyst parameters.

As described in [Gliozzi et al. (2014)], I_{cri} represents the 'Inhibition', a class or category that might have been closed to sales because of timetable rules or analyst choices ([Mahajan and van Ryzin (2001)]) as well as operational variations and constraints. This parameter is coded to be an integer that can be altered at schedule changes or through the YM system models, deriving it from analyst parameters. A_{cri} being the 'Authorization' denotes the maximum number of seats that will be accommodated in the segment r and category i that haven't yet been presented for booking. D_{cri} is the 'Availability' function representing the number of seats available for sale in the present moment, net of bookings already in place. Both A_{cri} and D_{cri} can be computed either based on demand or being stored in the inventory system. Here, in order to improve the performance, it was managed to store and later compute them every single time.

As a reservation will be done based on the train-date-class category, Availability and Authorization can be negative numbers that can result to bias, according to [Brown (1967)]. In that case, it would represent the number of cancellations that should be done to allow for a reopening. In general, 'Bias' are defined as the integers of the number of seats to be added to the value computed by the models following the analyst choice, based on Federgruen and Zipkin (1984). Here the Bias are represented by B_{cri}. Finally, based on [Mahajan and van Ryzin (2001)], it is important that the YMS is able to override some constraints from time to time, to find feasible solutions even if they do not respect some of the limits.

5.5 Constrained Demand Forecast

The simulation module provides an estimation of the ‘constrained’ forecast, that is the number of passengers at departure time and at predetermined moments of the booking horizon, called ‘snapshots’ or ‘photographs’ k; their value is obtained by the controls generated in the optimization phase. As described earlier, the constrained demand forecast depends on the followings:

(i) unconstrained forecast at each snapshot up to $t = 0$;
(ii) ‘advance’ bookings;
(iii) cancellations between couples of adjacent snapshots;
(iv) values of the controls $(\vec{K}, \vec{o})$;
(v) order of arrival of the forecasted demand between each couple of consequent snapshots.

The simulation process generates several ‘queues’ of pseudo bookings (or pseudo-PNR, where PNR is the reservation code) for each scenario. Each queue is consistent to the forecast of potential demand by segment-category at each snapshot. Within each interval between adjacent snapshots, different queues of the same scenario differ for the arrival order of pseudo-PNRs. Each queue is processed by a ‘Reservation System Simulator’ which computes the availability

r, i	o_{ri}	K_{ri}	X_{ri}	$K_{ri} - X_{ri} = Z_{ri}$	$\Gamma_\ell = \Theta_\ell + \sum_{\{jk \mid o_{jk} \geq o_{ri}\}} Z_{jk}$		$D_{ri\ell}$		D^w_{ri}
					$\ell = 1$	$\ell = 2$	$\ell = 1$	$\ell = 2$	
1, 1	1	30	0	30	30	–	30	–	30
2, 1	1	15	70	-55	30	30	30	30	30
3, 1	2	10	0	10	–	85	–	30	30
1, 2	3	15	0	15	55	–	30	–	30
3, 2	4	25	0	25	–	75	–	30	30
2, 2	4	18	0	18	40	75	30	30	30
$\Theta_\ell = C_\ell - \sum_r \left(\delta_{r\ell} \sum_i K_{ri} \right)$					22	32			—

Table 5.3: Practical example of availability computation, from [Gliozzi et al. (2014)].

and either accepts a booking request, if availability is non negative, or discards it.

The final averages and standard errors of number and revenue of accepted pseudo-PNRs, for all the queues-scenarios, represents the estimation of the constrained demand stemming from the set of controls proposed by the models. A report is generated and the data on average and standard errors are saved for each segment-category.

5.6 Role of the YM Team

The YMS operational work flow is massively automated, nevertheless it should work under the supervision of a team of analysts for a set of reasons, comprehending:

(i) to ensure the coherence of the price availability with the overall Company's strategy;
(ii) to properly manage exceptions and outliers, i.e. trains or dates differing significantly from their history of booking paths;
(iii) to protect against the risk of dilution;
(iv) to ensure the models are well functioning and producing sensible outcomes;
(v) to use the model results to identify possible changes in demand patterns. This could be used to perform corrective actions in the medium and long-term, e.g. improve schedule planning or pricing.

On daily basis, the users interact directly with the YMS through the 'User Interface', to: check or challenge the system decisions, change parameters and business rules, re-optimize and send a new set of controls to the Reservation System in real time. From the user perspective, running the models to re-optimize a TDC takes less than one minute and sending controls to the Reservation system is a matter of one click.

The analyst role should focus on setting YMS parameters and business rules, checking the YMS daily operations as well as managing exceptions. In particular, in order to limit the risk associated with dilution and in accordance with the Company's and brand pricing strategies, analysts have the opportunity to limit the authorizations resulting from the optimization model (i.e. parameters for maximum authorization per category) based on the expected demand. The analysts also monitor the presence of YM phenomena including spillage, spoilage and stifling.

The analysts use the information from the YMS tools and in particular the Informative Module, updated daily with detailed and pre-aggregated data tailored to the analyst needs, and others available both from the Company and outside (e.g.: statistics and researches, web data); they also receive directions from the Company, e.g. on how to limit the discounted fares per train brand, route and cluster of historical load factor using YM parameters. In Figure 5.13 on p. 144 is shown a visual example of the guidelines given to the analysts to set default parameters, based on historical Load Factor of the trains on certain routes.

The availability per fare category is therefore limited through business rules and parameters from the analysts, such that any hypothetical booking requests falling into segments and fare classes with lower value will find available -at most- the number of seats set manually by the analysts. One of the parameters is used to control the consistency of discounted seats available per train through a measure its Load Factor. Given the multi-leg topology of the train, such aggregation is a raw representation of reality, where there can be one or few 'under-capacitated' legs and many available seats on the other legs. However, it is still an useful aggregated measure for coordinators and managers.

At Trenitalia, the team of analysts manage up to 260 average trains per day, with up to 4 service levels and 15 (nested) categories, with a booking window of up to 180 days. It was estimated that each analyst controls an average of 1.5 million price/inventory decisions per week, with a high productivity rate (trains per analyst). Finally, a subgroup of "super users" also monitors the functionality of the system and provides input for the periodic maintenance of the models and processes.

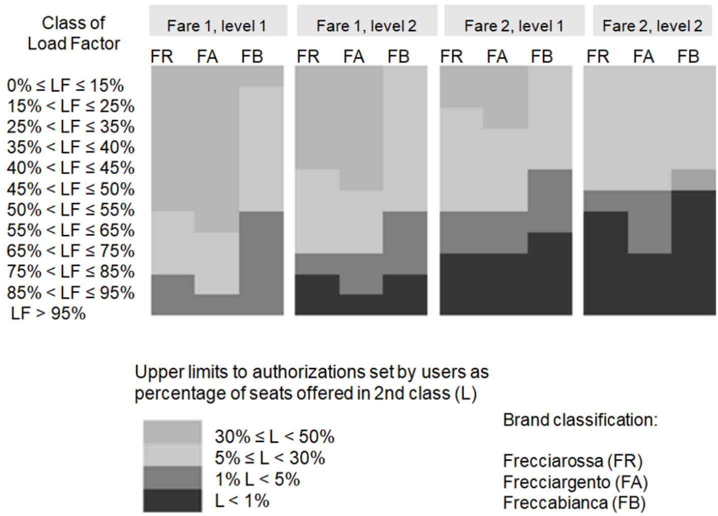

Figure 5.13: Visual representation of analysis guidelines on default parameters. The fares and levels are represented so that there is a decrease on the price from the left hand side to the right, so that for instance Fare 1, level 1 is higher than Fare 2, level 2. FR, FB and FA stand respectedly for Frecciarossa, Frecciabianca and Frecciargento and are three brands of HS trains.

6 Assessing YMS Performance Through Live Testing

We have seen already how in RM pre-departure and post-departure performances assessment can be used to monitor RM decisions over time and provide corrective actions. Such measures can provide a basis for comparisons and identify weaknesses in the revenue management processes that need to be corrected. Also the components of Performance Measurement have been detailed already, with references to other industries. In particular, it was considered expected (pre departure) and historical (post departure) performance, the first being focused on Corporate and Model performance and the second relying mainly on Standard performance and Revenue opportunity ([Vinod (2006)]).

In the following paragraphs those concepts will be fit in the rail context at Trenitalia, with the aim to present the rigorous methodology of evaluation and testing that has been developed during the past years. Some indicators consist of objective data, such as Revenue, Passengers or PassengerKm of each train, compared to an appropriate average of the history. This type of data are directly retrievable from post-departure data, and shouldn't be considered properly monitoring indicators; they will be presented in the followings. Other indicators, instead, belong to the Monitoring system; they are based on the estimation of the "real" potential demand of the departed train. The following part of this paragraph will focus on those indicators and on the underlying “X Algorithm” which provides such estimation.

6.1 Definition of a Set of KPIs

A working group from Trenitalia, Almaviva and IBM defined a set of KPIs for the YMS in a dedicated project. As presented earlier, they are not properly Monitoring indicators as they do not rely on the Revenue Opportunity estimation, but simply on YMS pre and post-departure data which are provided in the YMS Informative Module as well as in the Company’s data-warehouses.

In details, a set of 39 indicators have been identified and detailed in a Technical Report (unpublished). They are differentiated for user (e.g. Manager, Analyst), phase (Planning, Pre-departure or Operations, Post-departure), purpose and special cases. The work also focused on: how to compute them, their expected values and actionable meanings, users and purposes. The identified KPIs mostly relate to:

(i) Revenue;
(ii) Passengers and PaxKm;
(iii) Potential demand and contention, that takes place in case a train is undercapacitated in at least one leg;
(iv) Availability and related unbalance among classes;
(v) Load Factor;
(vi) RASK;
(vii) RPK, or Revenue per Passenger Km.

For each indicator, the following aspects were analyzed:

(i) Target and description;
(ii) Formulas and calculation;
(iii) Aggregation level, e.g. per train and departure date, train frequency, period, route;
(iv) Timing of the computations, actors, levels of users involved;
(v) Data sources where the information resides;

Nr.	Indicator	Planning	Pre-departure	Post-departure
1	Analyst loading index			
2	Information on the standard history			
3	Analysis of the schedule			
4	Budget information			
5	Revenue per available seat Km (RASK)			
6	Potential Demand			
7	Potential Contention			
8	Degree of potential contention			
9	Average revenue per passenger-km			
10	Degree of potential contention			
11	Critical train for load factor			
12	Critical train for revenue			
13	Revenue opportunity			
14	Average revenue per Fare Category			
15	Seat-km potentially contended with respect to the total offered			
16	Load factor in advance			
17	Average expected load factor			
18	Book status comparison between photos			
19	Variability of the categories for average revenue at sales opening			
20	Variability of the categories for pax booked			
21	Imbalance between classes			
22	Physical availability per route			
23	Physical availability by train			
24	Level of authorization granted			
25	Post-departure load factor			
26	High spill			
27	High spill degree			
28	Spoilage			
29	Cancellation index			
30	Stiffing			

Figure 6.1: Main KPIs identified per phase

(vi) Type: i.e., management, commercial or institutional KPI.

Table Figure 6.1 on p. 147 lists main indicators identified. Finally, as side activities of the KPI project, a series of activities have been identified on clustering trains per revenue, linking with budgeting and planning, managing analysts activities, etc.

6.1.1 Monitoring Module and Revenue Opportunity

We have seen in previous chapters how the use of monitoring tools is common in RM discipline to analyze performances after the departure, assess whether the YMS and the analysts performed well or not and define proper corrective actions. We have also seen how the use

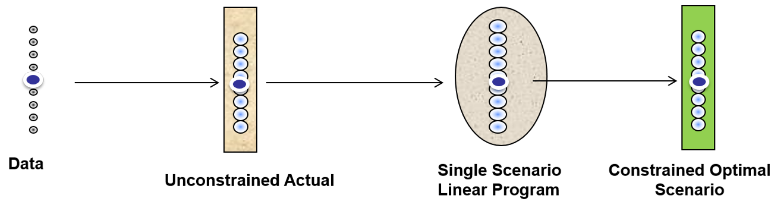

Figure 6.2: YMS Monitoring process

of the Revenue Opportunity Model is diffused for post-departure performance measurement, to determine the possible incremental revenue, computed as the reachable percentage of the maximum revenue opportunity. The Revenue Opportunity can be described as the amount of revenue which could have been obtained through a ‘perfect’ RM. Main reference for this Chapter is [Marchetti (2004)].

Here below a more detailed description of the Monitoring module is outlined. The Monitoring Module implemented in the YMS runs a *post mortem* for each single departure, to spot any possible ‘systematic’ issue which limits the ability to fully maximize the revenue. The core components of the Monitoring module are the unconstrained forecast, the use of the parameters, the optimization and the simulation models, the post-processor, each controlling different aspects of the whole system. As stated in [Marchetti (2004)], an algorithm (later named “X Algorithm”) has been developed to reconstruct potential demand and related Revenue Opportunity, then compute the values of a set of indicators for each departed train. The set of procedures, data and analyses on a departed train compose the Monitoring system, a fundamental instrument to control the Yield Management system. This is presented in Figure 6.2 on p. 148.

Based on [Marchetti (2004)], in further details, the X Algorithm estimates the potential demand and computes the related Revenue Opportunity at train, departure date and class (TDC) level. The main distinctive element for the monitoring analysis is that here, differently from the YMS models, the truncated demand is known and deterministic, being the one which took place and could book given the availability at the moment the reservation was attempted. Therefore, in the absence of closures, the X Algorithm provides the exact potential demand in the analyzed case, measured after its departure. Contrarily, if there have been closures, then the X Algorithm takes into consideration the periods in which the bookings were closed for lack of capacity and uses the same approach of the unconstrained demand forecast, but here without the proportional correction. Notably, it is necessary to estimate the potential demand only in those intervals and segment-categories in which closures have occurred. This is obtained through the use of gradient or finite potential demand *Uk* in those intervals, similarly to what has been done in the YMS models for all future snapshots, both in the unconstrained forecast that for the multiplicative correction.

According to [Marchetti (2004)], the 'exact potential demand' cannot be computed as the potential demand with the same proportional correction provided in the forecasting models. By fact, a train without 'closures' in its own booking path, but having many closures in its scenario history, has a potential demand that exactly coincides with the real train demand; contrarily, for that train the demand forecasted by the YMS models would certainly be greater, at least initially. In addition, here in presence of closures only the additive method is used: the multiplicative correction was not considered necessary as the aim here is not measuring the exact amount of the unconstrained demand, but assess its existence and the fact that those closures took place.

As stated in [Marchetti (2004)], the choice of the history should be done carefully and based on different criteria than in the YMS models. In partucular, it should be as similar as possible to the own booking path and and as "standardized" as possible. For this purpose, a specific issue was related to the weights of scenarios in the history, as in the YMS models they are dependent on the time in which the forecast is made, compared to historical data. The Monitoring models use a different way to determine the weight of the scenarios. It was chosen to calculate the weights based on the 'similarity' per each scenario on O&Ds, categories and snapshots. Such similarity is based on multidimensional euclidean distance of demand and its euclidean similitude index. The weights are then normalized to obtain the probabilities associated with each scenario of the train history. Therefore, the overall similarity of scenarios and their demand across all segments will be taken into consideration.

Once the potential demand for a departed train has been estimated, it is possible to optimize this unique scenario to determine the best combination of the booking requests to be accepted, and thus the total potential revenue. The differences between the optimal passengers and the real ones, with their associated revenues, allow to determine the performances and the actions that could have improved results. The X Algorithm demonstrated to be able to identify, with a reasonable margin of error, the main phenomena that allow us to judge the behavior of the system as a whole towards a departed train. High spill, spoilage and stifling are computed at a fare category and market levels. Such analysis is developed for each segment, class, category and for each departure date. It shall be noted that it is very disaggregated and detailed one, able to provide a fruitful and precise input. However, related flow are related to the complexity of reading results from the analyst, together with the need of all relevant data to work and the possible issues in case of missing data.

Diverse results can be present on different O&Ds of the same train, class and departure date, due to the multi-leg topology. Therefore, for instance, a certain group of markets can be well managed, while in others spillage, spoilage or stifling can be present. This makes the monitoring results not easy to understand and requires deep analysis from skilled analysts.

It should be noted that here the final results, denoted as 'post departure', include a certain period after departure, long enough to encompass what happens after the departure of the train from its initial stop. By fact, it should be considered that a trip may last up to several hours across many stations, therefore there is a high presence of bookings, go-show, no-show, recapture.

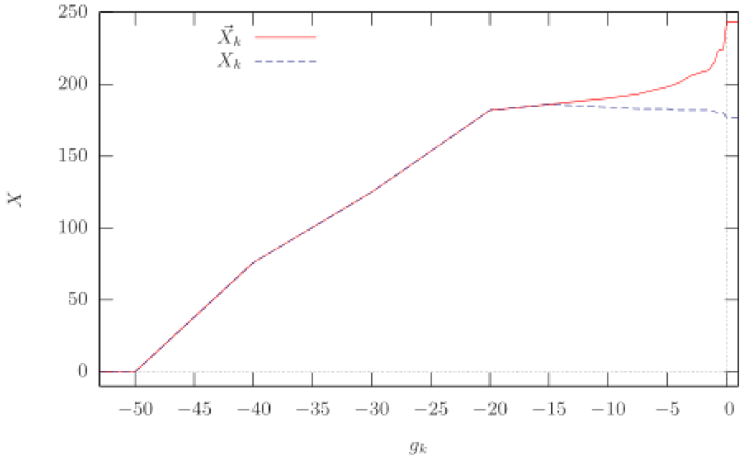

Figure 6.3: An example of application of X Algorithm for an Origin Destination Frequency

Figure 6.3 on p. 151 presents the passengers booked in an ODF at consecutive observations during the booking horizon, from the sales opening to the train departure. In the graph, the blue dotted line represents the booked demand, with the flat part representing the closure (with some cancellations and no-shows), while the red continue line is the estimation of potential demand. This example relates to a day of exceptional peak, side-by-side to a major holidays. Normally, that train in that day of the week is off-peak, therefore its own scenario history is quite different from the actual one. In this case the analyst set for such train the history of Sundays instead of its own frequency, in the attempt to boost forecasts, but this was not enough to allow the YMS to catch the real demand. Most probably, the un-constraining here is not enough to catch up with the real demand of that date, but, as described above, for our purposes this estimation was enough.

The X Algorithm provides atomic suggestions per O&D and Category on a certain departed train. They are in particular the differential optimal passengers which should have been accepted with respect to what has been done in reality. In the case illustrated in Figure 6.1 on p. 153, where Fares are ordered from the lowest (Fare 1) to the highest (Fare 6), the main points were the following:

(i) 31 more passengers should have been accepted on the segment Rome-Florence, Standard fare;
(ii) 43 more passengers on Florence-Milan, Standard fare;
(iii) 9 more passengers on a discounted fare on Naples-Rome, that corresponds to an overcapacitated leg.

The first two cases are related to Spill or Spoilage, and the third to Stifling. Figure 6.4 on p. 154 display such suggestions per ODF with respect to resulting revenues. It allows visualization of the differential revenues that could have been achieved following the suggestions on passengers, which were provided previously.

Figure 6.2 on p. 155 reports main Monitoring indicators, comprehending: the contention (i.e. if and how much the potential demand surpasses the accepted one) and related potential Load Factor, analysis of typical YM phenomena of Spill, Stifling and Spoilage, critical issues related to PaxKm and revenue, unsold capacity and the cost of associating some categories (e.g. discounted tickets for special and 'protected' categories of travelers, owners of annual tickets or cards and others which shall have guaranteed same availability of Standard category for policy reasons, rules and social acceptability constraints. In this example, the computed Revenue Opportunity resulted of 3.12%.

Due to the presence of several O&Ds, on the same train there is commonly the coexistence of diverse phenomena, that take place in different O&Ds. So, for instance, a train which is not saturated in the initial and final legs of its route and sold-out in the central legs can perform (i) stifling in the case it sells few discounted fares due

	Difference Optimal - Departed Passengers					
	Fare 1	Fare 2	Fare 3	Fare 4	Fare 5	Fare 6
Naples-Rome		0.02		9		
Naples-Florence	0.12	9.52		1.5	0.1	0.24
Naples-Bologna		9.02			0.06	0.18
Naples-Milan						
Rome-Florence	0.18	31.6			0.22	
Rome-Bologna						
Rome-Milan						
Florence-Bologna	0.06				0.1	
Florence-Milan	0.74	43.28		0.22	0.6	
Bologna-Milan				2		

Table 6.1: An example of elaboration of X Algorithm for a train/date. Differential passengers per O&D and fare category

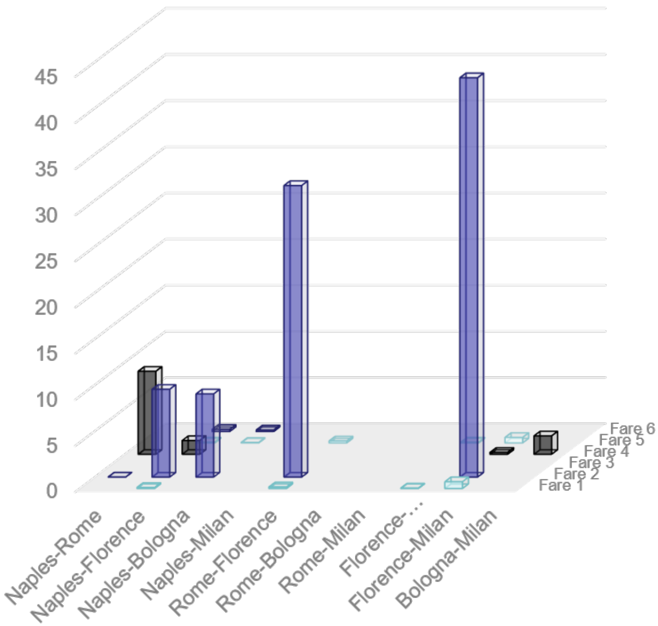

Figure 6.4: Computed difference of the Optimal Revenue with respect to the actual Revenue of the departed train

to closures in certain periods of the booking horizon, or (ii) spoilage if such denied bookings took place in every fares and segments insisting on those legs such that the train left with empty seats. It is useful to note that the nature of spoilage is here different from other industries, due to the presence of several O&Ds insisting on the same legs and the absence of overbooking policies. In the other legs of the central part of its route, for instance, the same train may report high spill if the sale of too many low fare bookings was allowed on short O&D traversing those legs.

For clarity purpose, it is noted that here we refer to 'Revenue Opportunity' as the *difference* between the optimal revenue and the revenue from departed passengers; elsewhere, in some cases of literature o practice like [Vinod (2006)], it is named revenue gain. At the same time the potential revenue is elsewhere defined as Revenue Opportunity. More specifically, in our analysis, 'Revenue Opportunity' (RO) is defined as the possible revenue gain beyond the actual gain, while in other approaches in literature the baseline was identified as the minimum revenue. Figure 6.5 on p. 156 represents the main

	Index	Numerator	Denominator
Contention Index	1		
Degree of contention	199.85%	939	470
Average Potential Load Factor	173.03%	701.001	405.14
Potential demand (passengers)	1327		
High Spill	1		
High Spill Degree	9.14%	36.767	402.288
Spoilage	0.00%		
Spoilage Degree	0%	0	402.288
Stifling	0		
Stifling Degree	0%	0	402.288
Critical Train for PaxKm	148.30%	402.288	271.268
Critical Train for Revenue	156.96%	42.097	26.819
Unsold Capacity	0		
Cost of Associating Categories	0.34%	43.556	43.409

Table 6.2: An example of elaboration of the 'X Algorithm' for a train/date. Computation of the values for the Monitoring indicators

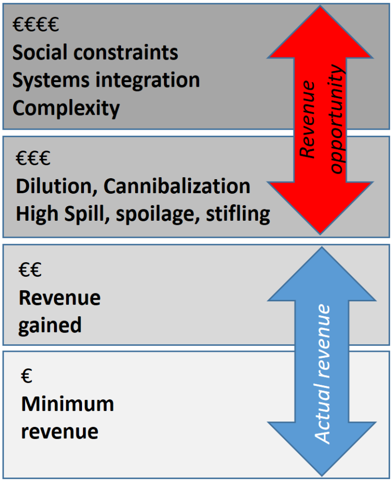

Figure 6.5: Components of the Revenue Opportunity in our case

costs (or opportunity) of a company to extract the maximum value from customers, selling all possible inventory to the best passenger mix and fully reach the revenue opportunity, currently lost. To do so, a company's efforts should focus not only to solve typical YM issues, but also on other points such as systems and organizational integration. Other topics, which limit the full achievement of the revenue opportunity, are real limitations and reside in social aspects which are overarching for a State-owned Rail Operator, as well as limits due to the Company's agreements and regulations.

6.2 Assessing the Performance of a New YM Model

The objective of this study focuses on conducting an experiment and determine a correlation and causal relationship between the experiment results and treatments, to assess advantages and disadvantages of the prototype towards the incumbent algorithm and decide on the future implementation of a new one. To evaluate diverse models, it has been adopted the approach from [Talluri and Van Ryzin (2004)], stating that an ideal YM algorithm is the one that maximizes (a certain measure of) revenue. In the current case, such measure is represented by an optimal combination of Revenue and Load Factor measured per each combination of train, class and departure date (TDC).

This does not apply only at the beginning In the steady state and maturity of a Yield Management system, the opportunity of periodically calibrating or improving the models should be carefully evaluated, as well as the adoption of a new one. In our YMS, from its inception up today, several modifications have been designed, tested and implemented, namely: multiplicative correction to the forecasting model, estimation of the value of new segments and categories, development of a monitoring module, others. A major algorithmic change, impacting on all modules – forecasting, optimization and simulation, has been recently developed, tested and implemented; it is not disclosed in the current document for Trenitalia's choice. This modification was tested live on a set of high-speed trains Frecciarossa and Frecciargento starting from February, 2018. After 11 weeks of experimentation, the results were good and supported by a solid methodology, therefore the test was stopped earlier and the change was implemented. The experiment, from its planning to the results analysis, will be described in the followings.

Rigorous testing is required to compare the implemented YMS, or 'incumbent', to the new one(s), or 'prototype'. This can be performed through a comparison of post-departure performance measures (e.g.: revenue, passengers, O&Ds fares mix, other factors), but can also comprehend other metrics (e.g.: measures of forecasting errors at certain moments before the departure) as well as post-departure analysis of the revenue opportunity (which is here performed by the Monitoring module described earlier). Some of those are the KPIs already presented, while others are calculated in the Monitoring tool. Furthermore, other criteria are also important to be implemented in the evaluation phase as provided by [Talluri and

Van Ryzin (2004)] that includes, among parameters, the controllability, robustness and adaptability of a YMS; in further details:

(i) Controllability: the system should allow for analyst control, as their knowledge is broader;
(ii) Robustness: against diverse possible market circumstances, it should work well in any case;
(iii) Adaptability: the YMS should ensure a prompt reaction in case of quick changes in external circumstances.

To assess the prototype performances towards the incumbent, it was initially looked at the Design of Experiments approaches and tools, in particular to minimize the number of necessary tests, in other words to optimize the experiment by minimizing the associated costs and effort[1]. Broadly speaking, and as described in the dedicated section in 3.4, Design of experiments (DoE) was considered a useful tool to:

(i) identify all possible factors to be considered as independent variables;
(ii) estimate the contribution of individual factors to the main outcomes;
(iii) determine which factors could determine any faulty developments;
(iv) eliminate or block some factors and consider others.

As for the test methodology, the following ones have been evaluated:

(i) analysis of historic data;
(ii) simulation;
(iii) sandbox testing;
(iv) live testing.

The methodologies listed above are described in the followings, with a focus on their positive and negative aspects, evaluated for the

[1] `https://www.isixsigma.com`, retrieved on September 29, 2018

methodological choice. The idea is to choose the approach which best fits the nature of the problem, the research question(s) and the data and resource available for the analysis. At the end, due to industry specificities and operational constraints, it was managed to adopt a different approach, which follows [Talluri and Van Ryzin (2004)], also reported in [Temath (2010)]; its new methodology, called 'sandbox testing', fits better to our case. Let us present how the choice was made.

According to [Talluri and Van Ryzin (2004)], an experimentation based on historical data is incomplete. Specifically, it does not provide full information on the market, as it doesn't cover but a small share of relevant customer and competitor data. Therefore, it isn't suitable for generalizations.

A second approach evaluated is the use of simulation, to reflect customer and competitive behaviors through an adequate modelling and including various scenarios. According to [Talluri and Van Ryzin (2004)], simulations are relatively easy to be designed and analyzed and allow a full control on external factors. However, this falls short of solving the main problem of subjectivity or arbitrariness in selecting a suitable model that can be implemented in simulations. Traditionally, RM models and algorithms are tested using simulations which are often based on the RMS models themselves. Therefore, such simulations are not able to provide further information on the same model validity. On the contrary, they may be wrongly in favour of the implemented RMS which is based on their same models. Furthermore:

(i) customer behavior is too complex to be recreated accurately by simulations which are based on models;
(ii) competitive reactions are hard to model and recreate, unless it is known which models and objectives the competition follows;
(iii) unexpected events cannot be modeled, and they have a broad impact on RM outcomes.

Another option for test methodology was sandbox testing. As described in [Talluri and Van Ryzin (2004)] it has the following benefits:

(i) covers an extended analysis period, without models or assumptions;
(ii) allows the use of actual trains and markets in the analysis;
(iii) provides an overall comparison of the system as a whole: analysts, information, automated models, real world issues, and is therefore presented as "holistic";
(iv) yields insights on reactivity of the models and tests RM components and design, allowing for continuous improvements;
(v) produces sound and reliable outcomes.

Finally, a live test is suitable and applicable for this case because it is sufficient to account for customer behavior and allows for the application of experimental design methodologies as well as the analysis of results. It has similar benefits of sandbox testing methodology, but in addition it is able to recreate all the iterative process of customers reaction to RMS decisions at every re-optimization during the booking window. A live testing approach could be developed, incorporated and run, so that it can consider the effect of the YMS at various moments of the broader booking horizon, when optimization is performed.

The test was centered on a complex modification impacting all the YMS models: forecast, optimization and simulation. Before the go-live, a lot of effort was dedicated to the choice of proper criteria to select the "best" new algorithm to be tested. Then, the selected one has been implemented in the models, so that the incumbent and prototype algorithms can run alternatively on the diverse sets of trains.

Many tests have been performed, mostly focused on:

(i) improvement of the forecast accuracy in the simulations run;
(ii) overall optimization results, based on main post-departure metrics.

Furthermore, the reciprocal influence of different allocations of the capacity on the combinations of O&Ds and fare categories has been studied. A special effort has been dedicated to test the possible effects not only on the total amount of revenue and passengers, but also on their distribution per O&D and fare

For the live test, a careful choice of the test set has been made so that the trains managed from the incumbent and the prototype were as similar as possible for demand amount and mix, routes and stops and other relevant factors; this will be described later.

In the preceding part, it has been explained how the choice of the best approach for the experiment should reflect the nature of the problem and fit to the context. With this regards, the organizational challenges, design and setting up the experiment, implementation difficulties and analysis of results will be also detailed in the followings.

6.2.1 Experiment Methodology and Overview

The following part presents and justifies the methodological framework and outlines the experiment results. A live test was performed using two competing algorithms, the incumbent and a prototype, controlling comparable sets of trains during adjacent days.

Key phases have been, among the others, to:

(i) choose and challenge the methodological approach;
(ii) define the test perimeter;
(iii) run the test, limiting the influence of external factors as much as possible;
(iv) analyze the test results and derive the outcomes.

This hasn't been a linear process, as the initial results led to improvements and refinements during the test run.

As for the first point, as anticipated it was resorted to perform a sandbox test as a more fitting model for this study. Commercial simulators, mostly used in aviation, didn't fit well to our context where there is much more offer and interchangeability between trains and also the competition from alternative transport modes which are largely available and, for shorter O&Ds, mostly used. In particular the reference is the Passenger Origin—Destination Simulator (PODS) research at the Massachusetts Institute of Technology (MIT) which investigates the interactions between different RM optimization, forecasting and estimation models, incorporating an estimation of the "willingness to pay" and the revenue impacts of using sell-up expectations in combination with various forecasting and optimization tools[2]. A comprehensive simulation model would still be applicable in this part; however, according to [Talluri and Van Ryzin (2004)] it could neither demonstrate the strategy for validating the simulation model, nor convince researchers and management beyond reasonable doubts that the results were trustable, as it is based on the same model to be tested, therefore biased towards it. Another option was to use the home-made simulator already used in YMS context for other purposes, but it didn't appear fit to the current case as well. It shall be noted, however, that the YMS simulator was adapted and used for other purposes within the same experiment.

The second step has been to define the experiment scope and sampling test and control sets. This has been challenging as an adequate numerosity of of observations had to be preserved while the test perimeter consisted in few comparable trains and there were constraints on the duration of the experiment to avoid a major timetable change. The sample was built so that not only the demand volume, but also the mix (e.g. business and leisure) was balanced as much as possible. Given the complex nature of the problem and its limited perimeter it was not convenient the use one of the theoretical approaches developed within the Design of Experiments literature, adopting for instance Linear Programming or Mixed Integer-Linear Programming approaches for planning and test results.

The behavior and results were observed for a relatively long period of time: in details, the test run from February to April 2018, for 11 weeks (or 77 days). A longer experiment of 18 weeks was planned, but the structured methodology allowed to trust the initial good results and anticipate the implementation of the prototype.

[2] `http://podsresearch.com`, retrieved on September 29th, 2018

Two homogeneous sets of trains were defined and it was chosen to alternate the control from incumbent and prototype on daily basis. For post departure analysis, the number of observations was of 718 and 720 combinations of trains and departure dates for prototype and incumbent, respectively. For pre-departure analysis instead, the choice was to collect the observations at 21, 15, 7 and 3 days before the departure date and for each one to perform an assessment of the forecast accuracy. Figure 6.3 on p. 163 and Figure 6.4 on p. 164 report the test key facts. This initial testing led to the decision of implementing a determined new algorithm.

The prototype and the incumbent algorithms ran both unsupervised, as the prototype has been implemented in the YMS which is characterized by a high level of automation. In addition, the team of analysts was trained to avoid any influence to the set of trains for test and control. Therefore, they didn't set different parameters to special dates or other limits that could influence the performance. This was done to avoid as much as possible the 'observer' bias and prevent analysts from interfering with the test.

	Test key dates		Nr. Of observations during the test period (trains/dates)			Number of weeks	Number of days
	Start	End	Prototype	Incumbent	Total planned test period	18	126
2018	01-feb-18	18-apr-18	718	720	Up to 18/04/2018	11	77
Equivalent dates 2017	02-feb-17	19-apr-17	715	716	% Completion	61%	
Prototype	Even trains observed in even dates, odd trains in odd dates						
Incumbent	Even trains observed in odd dates, odd trains in even dates						

Note: even trains proceed from North to South and from West to East, odd trains the other way around

Table 6.3: Key test facts - Post departure

Definition of the testing perimeter

The main challenge was initially the definition of two sets of trains which should be as homogeneous as possible. Finally, the choice of the live test method had some constraints. In particular, live tests cannot be run twice or three times, so the comparison leaves some error margins. This has been mitigated in particular through a proper definition of the test perimeter and timing. Therefore, the process for the selection of the test set (i.e., trains and period) for the experiment was very careful and aimed at:

(i) comparability of performances of the control group to the test set under many points of view, implying for instance the selection of a similar number of dates of circulation per day of the week;
(ii) absence of other parallel experiments running on the same trains (another test, not reported here, was running in the trains of the morning);
(iii) minimal presence of other factors of disturbance on performance results, such as: modification of analysts parameters, timetable, seasonality, special events or holidays.

The objective was to choose train sets with similar performance and serving a relatively small set of markets, to be representative of

Departure date		Days to Departure	Group	Number Of Observations	Null Values
Start	End				
07-feb-18	18-apr-18	3	Prototype	593	0
			Incumbent	595	0
07-feb-18	18-apr-18	7	Prototype	623	0
			Incumbent	614	0
13-feb-18	18-apr-18	15	Prototype	557	0
14-feb-18			Incumbent	534	0
19-feb-18	18-apr-18	21	Prototype	492	0
20-feb-18			Incumbent	489	0

Table 6.4: Key test facts - Pre departure

the variety of the high-speed market. For this purpose, the following were considered as key factors of homogeneity:

(i) the performance of the metrics of results to be tested (e.g. revenue, passengers) in historical scenarios;
(ii) the 'proximity' of trains in terms of: time of departure, routes and markets (O&D) served, movements schedule (e.g.: days of circulation).

The test sample was composed of 20 high speed trains, of which 10 even and 10 odd. All trains belonged to the HS brands Frecciarossa or Frecciargento. Therefore, it should be noted that the choice may have a bias towards the highly performing markets.

As anticipated, the trial period planned was the following: Thursday 1^{st} February 2018 to Wednesday 6^{th}June 2018, for a total of 18 weeks (126 Days). After 61% completion it was stopped. The final analysis takes into consideration this early stop, which balances the effects on the two sets. Also the overall period of the test was chosen carefully, avoiding major changes in the planned schedule which take place in December and in June. Particular periods of holiday banks (such as Christmas Holidays or August break) or not representative seasonality were avoided as well.

Differently from "classical" examples of live tests within the airline sector made on alternate weeks, here the prototype and the incumbent algorithm had to manage the two selected train sets on alternative days. In particular:

(i) the prototype had to manage the even trains in even dates, and the odd trains in odd dates;
(ii) the incumbent algorithm had to manage the even trains in odd dates, and odd trains in even dates.

The train number identifies the commercial service and, among the others, reflects the direction: odd trains proceed from North to South and from West to East, and even trains the other way round. Therefore, this choice reflected the tentative to counterbalance the peaks taking place in different days of the weeks and hours of the day according to the direction.

Due to the periodicity of train circulation, the numbers of observations were not perfectly balanced. This was considered an ineliminable part of the "real world" problem, as other choices would have had more impacts on other factors of imbalance. Furthermore, considering periods of at least 2 or 4 weeks such unbalance is either eliminated or negligible.

Finally, an attempt was tried to support the analysis with an year-on-year (YoY) comparison, largely used in business contexts. The management, who shall evaluate the YMS performances, is used to this measure and usually requires it. Such comparison is very useful to provide a measure of change towards a baseline (the performance on the same period and trains on equivalent dates of last year). This can help to spot a set of phenomena; for instance, if the revenue and fares mix improve for the incumbent system with respect to the past year, it may be related on an YoY growth of demand as well as performance improvements of YMS models. In our case, and based on similarity in day of the week, the corresponding period for YoY comparisons was from 2^{nd} February 2017 to 7^{th} June 2017. For our purposes, this comparison was deemed not very useful as any variations incorporated sensible changes in macroeconomic situation, competitive landscape and perimeter, or other factors. For this reason, the YoY analysis results are here omitted.

Operational constraints and difficulties

This part is dedicated to the description of main challenges, difficulties and constraints faced during the experimentation, which are listed here below and detailed in the followings. They comprehend, among the others, the need to:

(i) follow and challenge the methodological approach;
(ii) define an adequate sample, for which observations and demand are balanced and other factor of unbalance negligible,

in presence of a relatively short period and a limited number of trains;

(iii) adjust and refine the perimeter during the test;

(iv) spot and manage exceptional demand situations, or outliers, and limit their influence on the overall results;

(v) take into consideration how the last dates of the test were more influenced by the prototype decisions, with respect to the initial ones;

(vi) limit external aspects, such as the analyst parameters to the YMS booking limits;

(vii) perform specific analysis and perform any corrective actions, where necessary;

(viii) consider cross-elasticities;

(ix) evaluate the role played by cannibalization;

(x) minimize unbalanced impacts of punctual decisions and parameters from the analysts;

(xi) decide on whether to stop the test earlier than planned;

(xii) definition of a baseline to monitor performances over time, through a YoY comparison or other references.

As explained earlier, the choice of sandbox testing, similarly to live test, was supported by the fact that the final results are built upon a series of subsequent choices made by the YMS during the booking horizon, which is actually up to 180 days. When the test started, the first departure dates were close to the first observation dates, so the prototype could impact few on results, which were much more related to the decisions from the incumbent algorithm all along the booking horizon.

As for the definition of the test perimeter, it has been discussed separately due to its importance for the test. As a reminder, the necessity to balance the test sets through alternating days of tran circulatons under the control of the two algorithms was related to the presence of short periods and relatively few observations.

Another challenge was represented by the need to refine the test perimeter. By fact, the initial choice of the test perimeter was further refined during the proceeding of the experiment. In particular, the early analysis displayed that the variability for upper classes, due to data sparsity, was very high. This added unpredictability to related demand levels and impacted negatively on forecast accuracy; therefore, it was decided to focus our analysis on 2^{nd}/Std class. Additional factors considered included the seasonality and the presence of stringent users parameters; as a consequence, the test duration has been divided in two periods and we finally resorted to focus on the second one, starting from 19^{th} March.

Thirdly, as the involved booking horizon became longer, in the subsequent departing dates, more and more share of results were attributable to the prototype (which had more time to influence results). The results displayed this clearly.

In addition, it was considered and managed the presence of outliers, which decreased forecast accuracy but had ambiguous effects on revenues. An initial broad analysis excluded some dates with exceptional demand (not fairly distributed in the two sets); they have been removed from the test set perimeter to avoid disturbances on results analysis. In particular, they are:

(i) the week of the 26^{th} of February, 2018 (exceptional heavy snow with service interruption);
(ii) the days of the weekend of 4^{th} of March, 2018, when national elections took place in Italy;
(iii) the days of the Easter weekend in April 2018 (Easter was April, 1^{st}), that falls on different days every year.

A further analysis on those exceptional observations led to the definition of an 'outliers' interval that is specific for our test, i.e.: the observations where the demand is out of the +- 1.5 SD. They represent in this case the 14% of the observations. The core of the analysis was done on the observations without the outliers. The elimination of a set of observations implied a limited unbalance on the number of observations per of days of the week, that has been carefully taken into consideration in the results analysis, as diverse days of the week are characterized by a different amount and composition of demand.

Another point was to limit external aspects as much as possible, in particular the analyst parameters to the YMS booking limits on special dates.

In addition, it has been necessary to perform specific analysis and define corrective actions. For instance, a focuesd analysis was done on ‘closures’, i.e. the cases when a certain segment and fare at a certain time is closed to sales for a certain train. This analysis aimed to reveal the proportion of closures decided by the system model in comperison to the others (i.e., parameters from analysts and business rules). It revealed the opportunity to limit the action of the analysts on parameters to leave more space to the YMS decisions; so upper booking limits defaults were modified in the second part of the test.

Lastly, it was considered the presence of cross elasticities among the trains, and in particular between the trains belonging to the test and the control sets. They could materialize with passengers shift from one group to the other in response to different availabilities per price and market. This was reduced by the definition of the test set and by the choice of alternating days and train directions, so that, for instance, a passenger had to travel in a different day to find a cheaper price on the same O&D in the case one of the two running algorithms provided different fare availabilities. Generally speaking, cross elasticities are quite diffused in the rail sector, most probably much more than in aviation, due to the high frequency of train services and the contemporary presence of several means of transport on main O&Ds (i.e. redundancies). More specifically, in our context there are many trains offering similar services per time band, so that customers can choose among different trains and Companies on the same market, with a very high cross-influence between adjacent services an possibility of cannibalization. The latter can be defined as the probability that a customer chooses another train in the case it fits best their objectives (fare or time of departure or arrival), or in other words provide higher values that satisfy their preferences. Therefore, in railways not only the degree of intra-modal competition (between trains) is high within main markets, but also the inter-modal competition, e.g. from cars in shorter O&Ds, or from low-cost buses for longer ones. Consequently, cross-elasticities are higher and modal switch is more common. For instance, it may happen that a delay of one train causes a chain of cancellations and no-show (passengers which have a reservation on one train but don't use their ticket, nor cancel their booking) in the delayed train(s) and late or post-departure reservations and go-shows (passengers without previous reservation getting on board) in the next train(s); this causes many challenges to the YMS.

Another point is the role played by cannibalization. Generally speaking, it is possible that a train performed in a certain way because it cannibalized demand from other trains, or on the contrary it received cannibalization. Cannibalization can take many forms, especially in RM context. It can also be confused with genuine market growth or demand stimulation. As already outlined in [Talluri and Van Ryzin (2004)], there are several forms of cannibalization which can take place in transport RM, relating to:

(i) Demand levels: different availabilities in the two sets can imply diverse demand levels;
(ii) Demand mix: even when total demand levels appear to be similar in the two test sets, the proportion of the diverse segments or fares that compose the demand can be different. This can be due to an exchange of passengers from test and out-of-trial train sets.

For what relates to our test, it may happen for instance in the case the prototype provides higher availabilities on certain combinations of prices and segments, attracting demand from the set of trains managed by the incumbent algorithm, or the other way round.Furthermore, other forms of improper cannibalization are outlined in [Talluri and Van Ryzin (2004)]. They do not relate to direct cannibalization between the two sets and at the same time and are harder to spot and measure. For what can be of interest for our case, tey can consist, namely, in:

(i) Consumer surplus: this happens when the prototype is able to extract, for instance, more consumer surplus, thus altering the overall demand levels;
(ii) Competition: can take place in the case the set of trains managed by the prototype is able to subtract demand from competitors, instead that from the control test set.

Another factor of impact is the parameter set implemented by the analysts, in the case it is not balanced among the test sets. Normally, analysts can influence the performance of the trains under their responsibility, interacting with the system and defining a proper set of parameters. In the course of this experiments, in the tentative to keep the two test sets comparable, they were asked not to modify or override the general set of rules.

As will be discussed later on, during the second part of the experiment and encouraged by the goodness and trustability of early results for the first weeks, from 19^{th} March the analyst's parameters were changed to leave more degree of freedom to the YMS. The same change was done to all trains, both managed by the incumbent algorithm and the prototype, at the same time to avoid any possible unbalance.

Another operational challenge was represented by the definition of a baseline to monitor performances over time, through a YoY comparison or other references. This aimed at reinforce and confirm test results provided on weekly comparison and, most of all, from alternate days comparison on groups of 2 or 4 weeks. The YoY comparison, largely used in business contexts, is computed through the identification of comparable days the year prior to compare the performances. This is simple and accurate enough in the case there are no sensible alterations to the change in control, competitive landscape, departure timings and special events falling differently. However, in our case there was a significant level of variance in those external factors, therefore the YoY comparison resulted not helpful and related relults are not reported here.

Input, process, output of the experiment

An high level DoE approach has been followed to minimize the use of the Company's resources, such as data, effort and others, and deliver a satisfactory experiment. In particular, an analysis was performed on inputs and outputs of the process, to separate the input which are not under control and model the ones which are observable and measurable. In our test however, the reality proved to be not so linear as the DoE approach. In particular, a set of factors have been identified which do not fall strictly under DoE classification of input, but still impact on the test results. They can be envisaged in particular in (i) the customer reaction to the availabilities set by the system and (ii) the several reiterations of YMS process taking place during the booking horizon,

The different elements to be taken into consideration, which are also summarized in Figure 6.6 on p. 174, are outlined below.

Input under direct control from the YMS. They are the YMS levers and in particular:

(i) supply (Trenitalia's Frecciarossa and Frecciargento): the number of seats offered per market (identified by O&D and time of departure) as determined by the YMS controls, based on external factors such as the supply considered as fix in the short term, the advance bookings and the forecasted demand;
(ii) pricing (Trenitalia's FR and FA): availability of the diverse fare levels pre-defined per market and grouped in Categories.

Other factors, not under Trenitalia's YMS control, but impacting on the results, in particular in the short term, are outlined below.

Actions from Trenitalia:

(i) Marketing & Sales actions;
(ii) variations to the supply in the short term (specifically, within the test time frame), due to perturbation factors like: schedule changes, reduced capacity on a certain train for operational issues, strikes and others;

(iii) service improvements, offer of ancillary services and others impacting the demand levels.

Competitor actions:

(i) variation of pricing, communication campaigns;
(ii) offer of new services, variation of capacity;
(iii) improvement or, in general, any modification of the service.

Other factors impacting the demand:

(i) seasonal fluctuations of travel need (within the time frame of the test period);
(ii) special events like holidays, fairs, strikes, disruptions, exceptional weather conditions;
(iii) macroeconomic scenario variations and outlooks;
(iv) cross influences, e.g. from competitor actions, or YMS actions on substitute train services. In particular, we refer to the reciprocal influence of YMS decisions on the two test sets, which was limited by the choice of the alternate days checkerboard implementation.

Other factors, not under YMS control, which can be assumed as fixed in the short term and specifically for the duration of the experiment:

(i) Trains schedule: due to an agreement among European railways, it changes yearly in December and can be adjusted after six months, in June;
(ii) Trenitalia's pricing structure and price points - fare levels defined per markets and grouped in Categories;
(iii) Nuisance elements (present but not relevant for the test results);

The Output here can be defined as the partition of the offer of seats of a TDC per market, time of departure and fare sent to the reservation system.

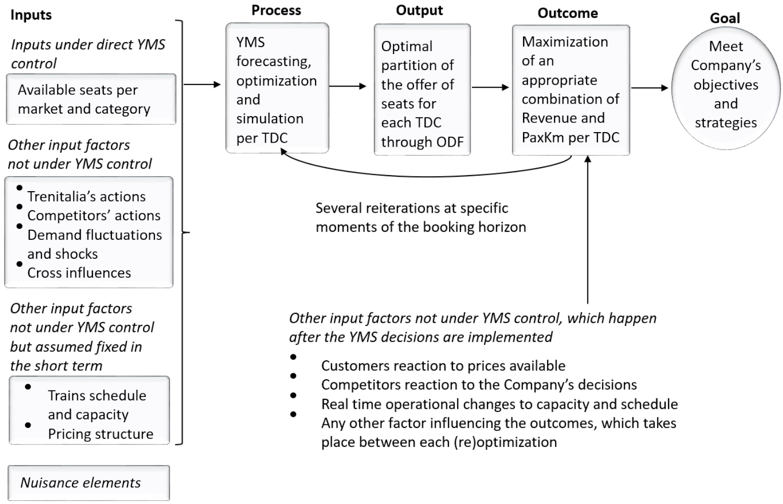

Figure 6.6: A DoE approach fit to our case

The Outcome consists on the maximization of an appropriate combination of Revenue and PaxKm.

The Goal is envisaged in the contribution to the Company's overall objectives and strategies on revenues, passengers, mix and others.

Figure 6.6 on p. 174 presents how our case was modeled with an Experimental Design approach. It shall be noted that part of the process follows different logics and is often iterative. The linear parts which fit into the DoE modelling is represented within the rectangular shapes, while the others are not in shapes and can be found in the lower part of the right hand side of the figure.

Differently from the aviation sector, where a lot of RM studies are focused on round trips and overbooking, here overbooking isn't practiced and there isn't a similar structured policy for round trips, i.e. with a complicated set of restrictions and dedicated pricing depending on which days of the week are included, distance in days between trips and other factors. Furthermore, as the sample was based on similarity in performance and customer mix of the test and control sets, which is enforced by the 'checkerbord' repartition approach, it is possible to assume balanced effects. Therefore this was considered not influent for our test. The same principles apply for groups: for a series of reasons it is reasonable to assume that they are equally distributed in the two test sets, so the presence of Groups is not to be considered among the factors which could influence the results in favor to one or another system.

Unsurprisingly, one of the major difficulties has been to identify and classify the input factors and subdivide the 'other factors' in observable and unobservable. By fact, to separate results deriving from YMS decisions from statistical variations and deviations of demand volume and mix which are due to seasonality, special events, competitor actions and other factors is a major challenge. Even more difficult was to figure out a meaningful, accurate and feasible way to measure them and their influence on results, where possible (i.e. for measurable factors). The overall procedure was made up by several iterations and trials in the attempt to "peel off", block or mitigate disturbance factors as much as possible. After presenting the overall results, a model to isolate the real model improvements (prototype vs. incumbent) from the disturbance factors will be presented.

6.2.2 Key Facts and Results Analysis

This part has the purpose to present the key test outcomes, keeping however in mind that the work is focused on methodological approach rather than results. Therefor, the outcomes are presented with the aim to show how a good experiment design can allow for accurate and reliable results in comparison to others, and this results in cost saving and also revenue gain. By fact, the test was meant to last 18 weeks, but due to the grounded methodology developed the managers trusted the satisfactory results of the first testing period and decided to stop the test earlier. Therefore, after 11 weeks of the 18 planned (or at 61% of test completion), the prototype was extended to the other trains in replacement to the incumbent algorithm. This helped the Company to save costs of further data and effort and the loss in potential revenue gain for the remaining period of trial.

The results described below followed a live comparison of two different RM methods including an incumbent system and a prototype, on alternate dates, as described earlier.It allowed a comparison of revenue performance under both systems.

The final version of the analysis, presented here, comes out from a lengthy process, tuned during the course of the experiment, aimed at getting the 'true' results while removing one by one the main disturbances factor, which could alter the test results in favor of one or the other of the two algorithms tested. Initially, all data were taken into consideration, then it was managed to remove progressively:

(i) exceptional dates (heavy snow, Easter, election days) classified as 'outliers';
(ii) all classes, with exception of 2^{nd} (for Frecciargento) and Standard (for Frecciarossa): for them as they are less populated and more sparse, the effect of the new algorithm could lead to instable results;
(iii) the first weeks of the experiment, as the YM model didn't have the chance to learn from past history and as the analysts' parameters were more stringent up to 19^{th} March.

The analysis presented in the following parts is focused on both:

(i) the final results on post-departure RM Key Performance Indicators (Passengers, Revenue, RASK, PRK);
(ii) the improvement on the accuracy of the forecasts, measured with statistical indicators both in percentage and in absolute values.

Post – departure analysis

The post-departure performances of incumbent algorithm and prototype are compared. The results are measured in variation of the indicators, already described, of Revenue, Passengers, Revenue per Available SeatKm (RASK), Revenue per PaxKm (RPK) and Load Factor (LF). The Key Performance Indicators variation is illustrated in Figure 6.5 on p. 177. All indicators registered an improvement with the prototype, with the exception of Revenue (-0.3%) and RPK which a small negative value. A further analysis focused on the second and standard classes, which are more populated in each combination of market, category and scenario, displayed improved results over the same period. Focusing the comparison on the Second/Standard class, all indicators improve, except for RPK; the reason is that PaxKm (the denominator) increases more than Revenue (the numerator).

The further step has been to divide the analysis in two periods, including the dates before and after 19^{th} March respectively. In that date, the analyst parameters were changed to allow for a higher degree of freedom to the YMS models. Furthermore, in the second part of the analysis time frame the effects of the prototype are more evident, as departure dates falling within this second interval are more

	Delta Prototype vs. incumbent	
	2018 – all classes	2018 – 2° / Std class
Revenue (%)	-0,3%	0,3%
Total Passengers (%)	0,3%	0,7%
RASK (a.v. in € cents)	0,012	0,042
RPK (a.v. in € cents)	-0,020	-0,001
LF (%)	0,22%	0,44%

Table 6.5: Key results - Post departure KPIs variation prototype vs. incumbent

distant from the beginning of the test, when the first re-optimization was done from the prototype and and the new set of controls was sent to the reservations system. In practice, the prototype had more time to produce its effects and the customers to respond to the new set of availabilities.

In the followings a further analysis on the test KPIs is presented, through the dispersion graphs of the prototype and incumbent values, aggregated per departure dates. This is focused on the Second and Standard Classes and covers only the second part of the test run, specifically the dates from 19^{th} March to18^{th} April, 2018. This is done to present more evidently the real impact of the new implementation. Figure 6.7 on p. 179 presents such post-departure analysis. Prototype results are indicated in Y axes, while X axes report values from the incumbent model. In all cases, it results evident how the prototype is able to leverage and exploit the cases in which the demand potential was higher.

Furthermore, a Paired T-Test was performed by coupling two consecutive dates for train service and frequency, for a determined period of time, in order to compare each time, for the same train-frequency, the closest pair of Revenue and Pax of prototype and incumbent. The results on the distribution of delta revenue are displayed in Figure 6.8 on p. 179 and Figure 6.9 on p. 180, which illustrate respectively the distributions of delta revenue for the Paired T-Test for all dates of the period, and without outliers.

Generally speaking, in both analysis the prototype outperforms the incumbent. However, in the first case, which includes the outliers, the performances are visibly better than the second one, where they were removed. We tried to analyze the effect of the outliers but, if their negative influence on the forecast was clear, by fact the results in terms of revenues were more unclear, or they couldn't be generalized. Apparently, for more populated classes the outliers didn't impact badly. A first reason could be related to the fact that they fell more into high demand periods rather than low ones. By fact, they were mostly related to peak periods (holiday, elections), with the particularity that election tickets are free (reimbursed from

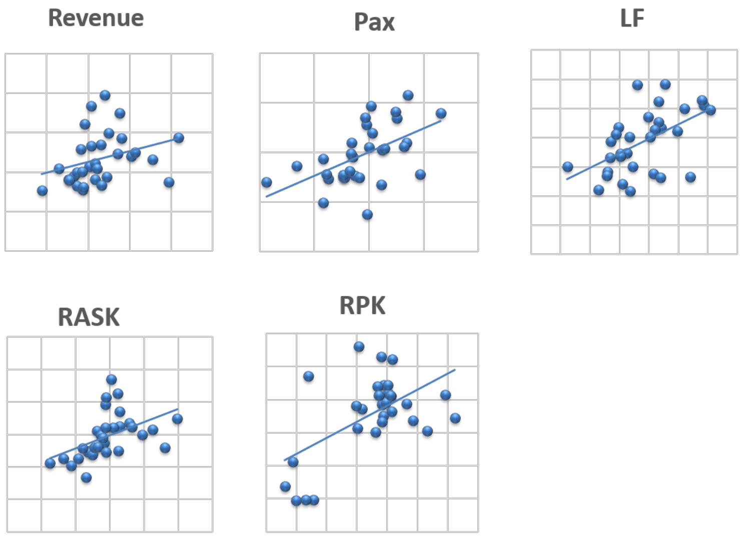

Figure 6.7: Key results - Post departure KPIs - dispersion prototype vs. incumbent

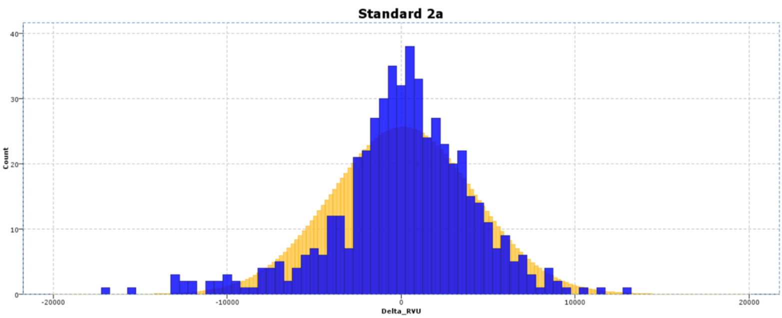

Figure 6.8: Distributions of delta Revenue – Paired T-Test - all dates

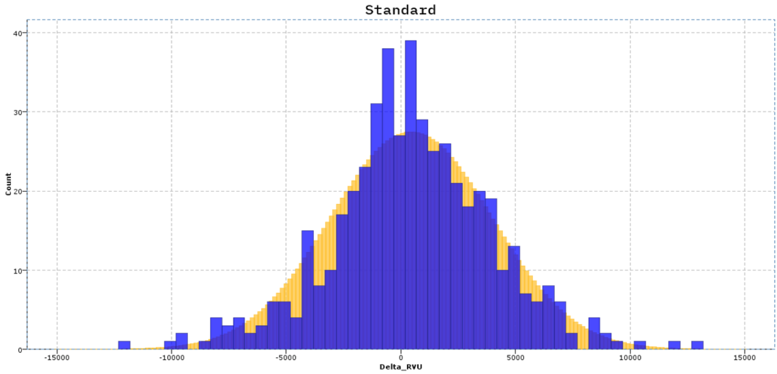

Figure 6.9: Distributions of delta Revenue – Paired T-Test - no outliers (elections weekend)

the State), while only a week of heavy snow determined lower demand than usual. Therefore, it is of common sense that if more peak dates are removed than the off peak ones, the overall effect will be a reduction on average revenues. However, this seems not sufficient. Another possible reason may be related to the use of nesting, for which already [Williamson (1992)] noted that, when there is a lot of uncertainty on low fares, the systems using nesting tend to overprotect, with a bad impact on results.

Impact on closures

A specific analysis was run on the closures performed over time. Closures take place when there is no (more) availability to satisfy a certain part of demand. Analyzing a single train for a certain departure date during the course of its booking horizon, we can reasonably assume that the capacity is fixed; this is true unless there are operational issues. Therefore, we can state that closures are the ultimate effect of the choices performed and represent the ''moment of truth" of such decisions. By fact, until there are no closures, any decision would be good – as it won't have any effects; when it comes to refusing bookings instead they have impacts.

Closures can be performed either by the system or the analysts. For our YMS, the possibility to set upper and lower bounds at atomic level has been implemented to ensure the respect of "social constraints" (e.g. finding a minimum number of certain discounted fares on specific O&Ds) and the respect of the commercial strategy of the company (e.g. a certain number of promos are communicated to be on sales and upper and lower bound may help to allocate those quantities according to brand and route managers requirements). Finally, the analysts can manage to set business rules and parameters according to seasonality, e.g. day of the week or time of departure. Generally speaking, it may happen that RM systems decisions are overridden from users. It is the case, for instance, when users trust the system decisions only up to a certain point but need to feel ≪in control≫ of the final decisions.

A massive use of parameters overriding the systems decisions can represent a relevant risk and has an impact both in the short and in the medium term. In the short term the YMS will struggle to find feasible solutions, which can be hardly optimal, while in medium term this behavior will alterate the history from which the YMS works. Contrarily, performing some safe price experiments allows the system to learn from customer reactions to any price change and improves its future performance, building a more useful history, scenario after scenario. Therefore, analyzing where the closures come from, either from the YMS decisions, users limits or other factors (e.g. operational modifications of capacity) provides valuable information. One of the objectives of the prototype was to be so good that the analysts could relax their parameters as much as possible. Extended tests to support such trustability has been done on managing routine and exceptions, where the YMS should react promply and send triggers and alerts to the users.

Figure 6.10 on p. 183 presents the analysis of closures per train-date-class and ODF to assess how much users or YMS impacted on decisions. They are displayed in weighted percentage for passengers per days at departure, considering the departure dates from 1st March to 28th March, 2018. If compared to the incumbent algorithm, the prototype demonstrated to be able to take the control on closures in higher degree with respect to the analysts actions. In addition, the prototype left the balance of closures and availability not too much dissimilar from the previous one; this continuity, facilitated by the consideration of a long and stable scenario history, was paramount to ensure continuity to the previous company choices.

6.2.3 Modeling the Influence of Main External Factors on Performance

This part deepens the analysis of the 'Other Factors' that impact on the test outcome and provides an attempt to model such relation, mostly linked to demand fluctuations and deviations. By fact, the purposes of the experiment comprehend to:

(i) identify any major factor which could interfere with the experimental results;
(ii) model the way those diverse factors contribute to the final results;
(iii) show that the introduction of the new RM method or process caused certain differences in the results.

Furthermore, one objective is to allow that further tests can be optimized to reduce the amount of data required for extracting meaningful information as well as a reduction of data-related effort, time and resources. For this reason, main steps of the experiment have been detailed and motivated, as well as its background, baseline and causal assumptions, based on a methodological backbone.

Also, a simple model on external factors effect has been prepared; this is the subject of the following part. A key focus of such model is on demand variation, that is considered one of main components of the external factors impacting the outcomes. By fact, this also incorporates other factors, such as modifications in the fares or supply from competitors in the market, because their impact on results

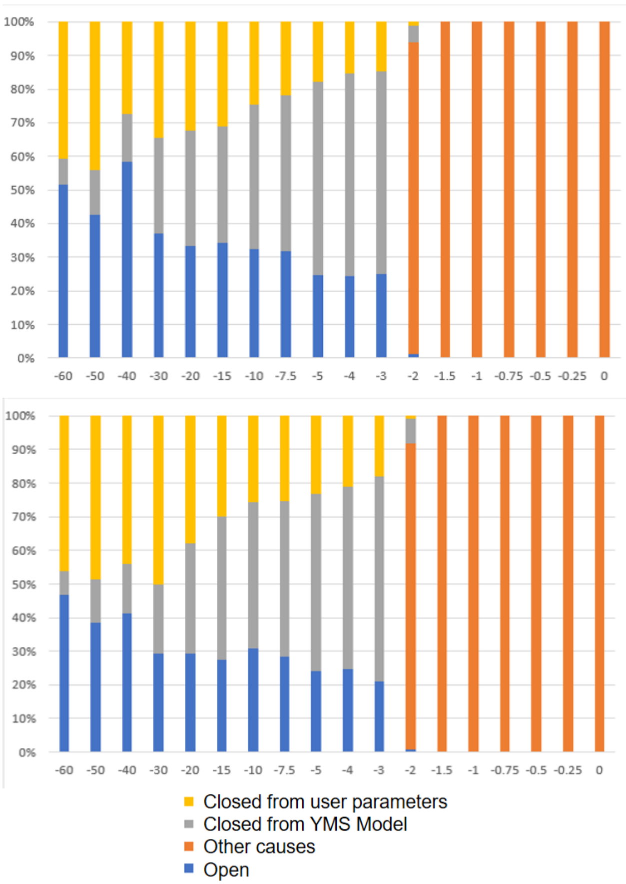

Figure 6.10: Analysis of closures

is mediated by demand response. Other factors, like operational issues and variations or measurement errors, will be considered but not modeled.

For our purposes, the potential market demand for transport can be defined as the part of the population that has a need (or wish, for instance) for any mobility services or modes in a given market on a particular period. Consequently, the total market demand is the part of the potential demand that will most probably materialize, for instance, in bookings or travels. The other part of the potential demand, that doesn't materialize for a series of reasons, can be defined as unobservable. The total market demand is shared by diverse transport modes in the market. A part of this will choose the train; this is a mode specific demand. We can assume (exogenously) that the overall demand is independent across days and markets; it can be estimated and considered relatively stable in the short term and specifically for the experiment duration. Finally we can assume that this demand is composed by a mix of different customer types, characterized by different willingness-to-pay for travel, behaviors and other characteristics.

Several approaches have been used to model the passenger demand for transport, which is difficult to estimate due to a wide variety of factors, as already described. In most cases, a top-down approach is followed, which starts from aggregated forecast and then splits it into modes, e.g. through choice based models which have been used for similar and other macroeconomic applications since the 1970s. For instance, one can refer to the works on passenger demand modeling and forecast for networked transport from [ATAP (2016)], [Ben-Akiva (2008)], [Tsekeris and Tsekeris (2011)], [Zhao and Kockelman (2002)], [Watling and Cantarella (2013)].

For our purposes, we use as references the guidelines for modeling travel demand provided by the Australian Transport Assessment and Planning Council in [ATAP (2016)]. According to [ATAP (2016)], time series analysis is a commonly used approach that fits the mix of deterministic and stochastic processes which are typical of the transportation sector. They allow the study of performance, events and trends over time, as well as economic and movements data. Examples of the use of this approach are Taylor, Bonsall and Young (2000) Chatfield (1984), as well as the following (for freight): Gargett and Perry (1998), Amoako (2002) and Sutfcliffe (2002).

Under this approach, most data stemming from processes under continuous changes are recorded as time-dependent, i.e. traffic data as well as socioeconomic factors and long term trends. To understand and exploit such data, it is necessary to separate the diverse components determining variations, in particular trend, cycles and random effects. In some cases (e.g. stationary processes having stable parameters over time) observations can be replicated to assess data variability, while in other cases (e.g., time series data) this is not possible.

In the followings the focus will be on transport demand modeling. In details, a couple of references from the literature will be presented, then a novel approach from the writer will be outlined.

A first example is provided in [ATAP (2016)]. The equation (6.1) on p. 185 combines a set of factors, assumed to be in additive relation, which build up the stochastic traffic flow:

$$X_t = T_t + S_t + C_t + R \tag{6.1}$$

Being:

X_t dependent variable;

T_t trend component, related to long term changes in demand;

S_t seasonal component, which is the fluctuation of traffic in various times of the year;

C_t cyclic component, related to macroeconomic variation;

R random component, e.g. shocks in demand or traffic.

Four components have been pointed out: trend, seasonal, cyclic and random. This is also displayed in Figure 6.11 on p. 186, also from [ATAP (2016)].

In the followings it is presented another approach retrieved from scientific literature, then the novel one from the writer is presented.

Under normal circumstances, say with an incumbent system and process under steady state, it is possible to assume that a single transport mode is able to capture a part of the total market demand. According to [Talluri and Van Ryzin (2004)] a new treatment can modify or exploit differently the market demand, for instance:

(i) capture underlying market demand;
(ii) extract a different part of consumer surplus;
(iii) stimulate the overall market demand.

In general, from the transport services supply side, a multivariate linear model can be used to estimate which is an adequate amount of trip services to be offered. This will take the form of the equation

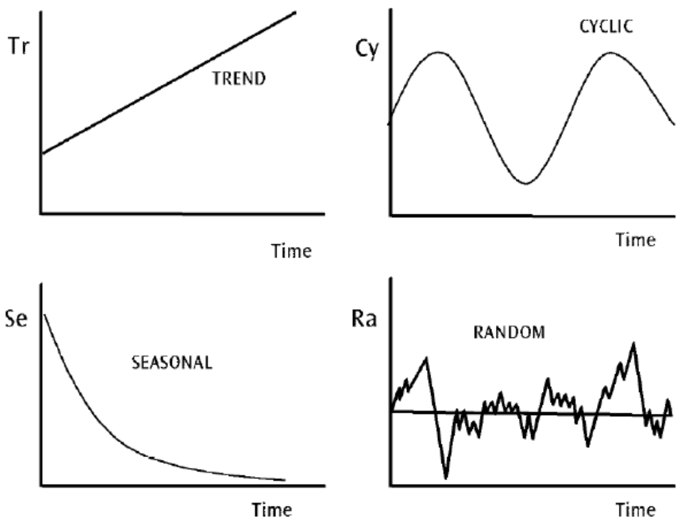

Figure 6.11: Components of a time series, from [ATAP (2016)], page 73

(6.2) on p. 187, which is based on same concepts of the one from [ATAP (2016)]:

$$Y_i = b_0 + \sum_{j=1} m(b_j \times X_{ij}) + \varepsilon_i \tag{6.2}$$

Being:

Y_i dependent variable

X_{im} m independent variables (e.g. demographic), for each observation i, with $1 \leq j \leq m$

b_m model parameters

ε_i error term.

Factors impacting demand and demand-related are the key ones, as they can disturb the test results and cannot be blocked, as well as operational issues or cross-dependencies. Below is provided a basic, novel formula to model main external factors that impact on the test outputs. Here the considered factors are:

(i) a strictly cyclical function representing the influence of seasonality, here denoted by month of the year, day of the week, hour of the day. Therefore, this component is a summation of several and diverse components which should be computed independently, also referred to as 'seasonalities';
(ii) the non-seasonal complement, here the Company's share of the overall market demand (or Market Share) which can be assumed as fixed in the short term;
(iii) the YoY trend, which is a very used measure (also) in RM context, as presented earlier;
(iv) the outliers, in absolute value;
(v) errors in measurements, in absolute value, related to methodology: i.e., the live test approach implies a certain error measure as test runs cannot be repeated;

(vi) cross-effects from test performed on other trains/dates and O&Ds of the control set and cannibalization of, or by, out-of-trial offers;
(vii) random factors, in absolute value, which captures the effects of special events causing demand shocks as well as operational issues impacting the inventory.

Here a change in the own supply is not considered directly. By fact, it was explicitly excluded by the accurate choice of the test perimeter, considering a period without major changes in the timetable. On the other side, any change in competitor supply is considered incorporated indirectly in the own demand component through customer reactions.

In details, under this thesis approach, the demand can be modeled as follows. We can assume:

$$Y_t = X_t \times S_t \times M_t \times T_t - E_t - C_{X_t} + R_t \tag{6.3}$$

Where:

$$S_{t+1(\text{year})} = S_t \tag{6.4}$$

$$T_t = \frac{X_t}{X_{(t-1)}} - 1 \tag{6.5}$$

Being:
Y_t dependent variable
X_t demand at the time t
S_t strictly cyclical function representing seasonalities
M_t non-seasonal component
T_t trend component
O_t outliers
E_t errors in measurements
C_{X_t} cross-effects
R_t random factor.

7 Conclusion and Future research

This thesis describes how the YMS in use at Trenitalia has been designed, tested, implemented and improved by a joint working group from Trenitalia and IBM, where the writer has been involved since its inception. The system design took inspiration and backbone from a previous one, implemented from IBM at Alitalia Cargo, which was leg based and optimized based on continuous variables, namely freight volume and weight. These models have been profoundly modified to fit the rail industry with particular regard to the vast number of O&Ds, actually an average of 36 on High Speed rail services. The resulting system is composed by a forecasting module, an optimizer and a simulator.

During the years the YMS underwent a series of major changes, of which some are reported in this thesis. A first focus is the study and development of such improvements and in particular a multiplicative correction to forecasting models. This implementation is innovative as: (i) the linear mix of additive and proportional methods has an exponential smoothing to regulate and gradually lower to zero the impact of the multiplicative component as departure approaches; (ii) it is performed per O&Ds, and (iii) the combination of forecasting components is a user parameter. Here the challenge was to enable the system to react promptly to any anomalous variations in demand, e.g. demand peaks, without overreacting or producing indeterminate or meaningless results. Therefore, many tests have been made to check performances in contended markets, which concentrated most demand, as well as in less requested ones, where data sparsity or a higher volatility could lead to suboptimal solutions, also at business or social level.

The second part of the thesis is focused on the development of a sound and reusable methodology to assess YMS performances in its steady state, for a more systematic testing of user parameters and algorithm recalibration as well as the evaluation of any major changes. This part encompasses:

(i) the definition of a set of KPIs to assess pre and post departure performances on regular basis;
(ii) the development of a Monitoring module;
(iii) the preparation, run and analysis of a live test comparing a new prototype and the incumbent algorithm, running on alternate days on comparable sets of trains.

In particular, the Monitoring module has the purpose to assess whether the decisions taken for a departed train during its booking horizon were optimal or suboptimal, and suggest any corrective actions to be undertaken. To do so it runs a *post mortem* analysis of the train. Firstly, it estimates the potential demand only where closures took place, otherwise it uses the exact demand that had the chance to book on that train. Then it provides atomic suggestions on the differential passengers that should have been accepted or refused per each fare cluster and market to reach the potential revenue that can be realized from the potential (or exact) demand optimally managed. Finally it calculates the associated Revenue Opportunity and a set of indicators, then quantifies phenomena like spillage, spoilage and stifling.

In addition, a live test has been prepared, run and analyzed to assess the performance of a new prototype towards the incumbent algorithm. The objective here is to build a reusable methodological backbone for current and future tests on algorithmic changes, which occur periodically during the operational life of an YMS. The umbrella approach of Experimental Design was followed. Under few assumptions and in real market conditions, a proper test perimeter was set to allow direct comparability of prototype and incumbent algorithms over time. For this purpose, it was decided to adopt a live test approach and more specifically sandbox testing methodology, as simulations could not replicate properly demand fluctuations, trends and statistical variations of the real environment. Furthermore, many re-optimizations take place during the booking horizon, followed by the customer reactions to YMS decisions; this non-linear and interactive process is not well taken into consideration from a simulator. The performances of the algorithms were compared across main KPIs, both pre and post departure, and it resulted that the prototype was able to both increase the forecast accuracy and improve performances. Thanks to the endorsement of the structured methodology which assessed good results, it was possible to interrupt successfully the described test after 11 weeks of the 18 planned. Such anticipation allowed the Company to save the costs related to data collection and other efforts to be dedicated to the test for the remaining period, and to earn an increased revenue flow from week 12 instead of 19. Furthermore, it is reasonable to expect that results of the new algorithms will be further improved over time, as:

(i) the controls sent to the reservations system will be more and more similar to the optimal ones determined by the YMS models, as long as the parameters and business rules will be relaxed over time;
(ii) scenario after scenario, the YMS will be able to perform some price experiments, due to less stringent user parameters, and they will increasingly form the history of reference;
(iii) for future trains, the new algorithm will operate from the beginning of the booking horizon, impacting more on final results.

The YMS fruitfully supported Revenue Management decisions, providing satisfactory results on main KPIs, such as revenues and load factor. Those results are at the basis of the Company's choice to maintain and further invest in the same operating YMS in the course of the years. The implementation of the YMS contributed to keep satisfactory economic results, through an optimal balance of customer mix, achieved in consideration to the social role of railways.

It should be noted how this has been a crucial period for the former monopolist Trenitalia, due to:

(i) a period of economic stagnation, pushing companies and customers to pay particular attention to costs and margins;
(ii) the increased price transparency and comparability enabled by digitalization;
(iii) the rise of intra-modal and inter-modal competition and, in particular, the opening of High Speed rail lines to open track competition in Italy.

Currently the YMS faces many challenges:

(i) demand fluctuations and volatility, with a tendency to last-minute materialization;
(ii) increased price transparency and growing importance of price as a key factor for the customer choice;
(iii) integration to the other systems of a Company;
(iv) possibility to add ancillaries, additional services and in general other dimensions to the pricing;
(v) growing demand for door-to-door solutions to be booked in a single click, which may call for wider revenue optimization in the near future;
(vi) changes in business models which may challenge traditional pricing approaches and complicate the mobility landscape, e.g. raise of shared mobility and e-hailing.

Further developments of the YMS are already foreseen or ongoing. Considering the price sensitivity of customers and their willingness to pay can improve forecast accuracy and the optimization decisions. A study is being carried out on cross elasticity versus main potential competitors, such as railways undertakings, airlines, bus services, car sharing. Some others are still under evaluation, e.g. implementation of, and integration to, other automated components of the RM system, integration of data from external sources. In this perspective of actual and future challenges, the specificities of multileg topology and social issues shall be taken into consideration. In particular the high number of O&Ds served should be exploited, as well as the presence of social constraints for a State-owned railway company are of paramount importance. To overcome current and future challenges, research should comprehend a mix of incremental innovation for short to medium term deployments as well as disruptive research to prepare the way for future challenges.

As for the incremental research, future developments may comprehend:

(i) further integration with other systems of the Company to extend the impact of the YMS, from Marketing to Planning and Operations, for reciprocal data exchange and improvement;
(ii) adoption of (improved) tools to perform optimizations and simulations on specific topics, to be managed separately;
(iii) integration of further data in the YMS models, either external or from the Company, to refine forecasting. It should be evaluated carefully as the addition of factors can have opposite effects on forecast accuracy.

As for the disruptive research, it may consider:

(i) forms of wider system optimization, encompassing diverse classes and areas of the network, multimodal transport;
(ii) "sensing" and being being able to influence in real time: mobility demand, competition, offer of diverse travel solutions;
(iii) integration of Machine Learning algorithms to Operations Research modeling to cover specific purposes, e.g. analysis of web data, forecasting customer behavior, preprocessing.
(iv) adoption of new methods (e.g., novel decomposition approaches) and enabling technologies (e.g. diffusion of quantum computing for commercial applications) to address complexity issues and improve time performance.

References

[Alptekinoglu and Semple (2015)] Alptekinoglu and Semple (2015), The Exponomial Choice Model: A New Alternative for Assortment and Price Optimization. Forthcoming in Operations Research.

[Armstrong and Meissner (2010)] Armstrong, A., Meissner, J. (2010). Railway Revenue Management: Overviews and Models, Lancaster University Management School, Working Paper 2010/35.

[AISCAT 2010] AISCAT (2010). Notiziario trimestrale a cura dell'Associazione Italiana Società Concessionarie Autostrade e Trafori. Roma, Italy.

[Alptekino and Semple (2016)] Alptekino glu, A., Semple, J. H., 2016. The exponomial choice model: A new alternative for assortment and price optimization. Operations Research 64 (1), 79–93.

[Anderson et al. (2012)] Anderson, C. K., Xie, X., 2012. A choice-based dynamic programming approach for setting opaque prices. Production and Operations Management 21 (3), 590–605.

[ATAP (2016)] Document prepared by the Australian Transport Assessment and Planning (ATAP) Steering Committee and approved by the Transport and Infrastructure Senior Officials' Committee, T1 Travel Demand Modelling, Australian Transport Assessment and Planning (ATAP), 2016, retrieved at: http://atap.gov.au/ on 10^{th} October, 2018.

[Belobaba (1989)] Belobaba, P. P. (1989). Application of a Probabilistic Decision Model to Airline Seat Inventory Control. Operations Research **37**, 2, 1989.

[Belobaba (2002)] Belobaba, P. P.,(2002),*O&D Control: What Have We Learned?*, MIT International Center for Air Transportation, Presentation to the IATA Revenue Management & Pricing Conference, Toronto, October 2002

[Ben-Akiva and Lerman (1985)] Ben-Akiva, M. and Lerman, S., Discrete Choice Analysis: Theory and Application to Travel Demand: MIT Press (1985).

[Ben-Akiva (2008)] Ben-Akiva, M. (2008), Travel Demand Modeling, 1.201 / 11.545 / ESD.210, Transportation Systems Analysis: Demand & Economics, Massachusetts Institute of Technology (MIT), Available at: https://ocw.mit.edu/courses/civil-and-environmental-engineering/1-201j-transportation-systems-analysis-demand-and-economics-fall-2008/lecture-notes/MIT1_201JF08_lec05.pdf , Fall 2008.

[Berbeglia (2016)] Berbeglia, G., 2016. Discrete choice models based on random walks. Operations Research Letters 44 (2), 234–237.

[Berbeglia and Joret (2017)] Berbeglia, G., Joret, G., 2017. Assortment optimisation under a general discrete choice model: a tight analysis of revenue-ordered assortments, working thesis, University of Melbourne

[Bergantino et al. (2015)] Bergantino, A.S., Capozza, C., Capurso, M., 2015. The impact of open access on intra-and inter-modal rail competition. A national level analysis in Italy. Transp. Policy 39, 77-86.

[Bergantino (2015)] Bergantino, A.S., 2015. 8. Incumbents and new entrants. In: Finger, M., Messulam, P. (Eds.), Rail Economics, Policy and Regulation in Europe. Edward Elgar Publishing, Cheltenham, UK.

[Beria et al. (2018)] Beria P., Raffaele Grimaldi, Daniel Albalatec, Germà Belc, Delusions of success: Costs and demand of high-speed rail in Italy and Spain, Transport Policy, Vol. 68, Pages 63-79, 2018

[Beria et al. (2016)] Paolo Beria, Renato Redondi, Paolo Malighetti, The effect of open access competition on average rail prices. The case of Milan e Ancona, Journal of Rail Transport Planning and Management 6 (2016) 271e283

[Berto and Gliozzi (2019)] Berto, A. and Gliozzi, S. (April, 2019). Comparing Yield Management models performances through live testing at Trenitalia, AIRO - Italian Operations Research Society, International Conference on Decision Science, 2019, Conference Proceedings (short paper), submitted.

[Berto and Gliozzi (2018)] Berto, A. and Gliozzi, S. (August, 2018). Un-constraining the Passenger Demand for Rail Yield Management at Trenitalia. Electronic Notes in Discrete Mathematics, **69**, 269-276.

[Berto and Gliozzi (2018b)] Berto, A. and Gliozzi, S. (submitted in August, 2018). Rail Yield Management at Trenitalia. Accepted for publication by IMA Journal of Management Mathematics

[Berto (2018c)] Berto, A. (2018, July 8-11). Rail Yield Management at Trenitalia. Presented at EURO 2018, 29^{th} European Conference of Operational Research, Valencia (Spain). The project abstract can be found in the EURO 2018 Full book: `http:`

//euro2018valencia.com/wp-content/uploads/2018/07/Conference-Handbook_WITH_ABSTRACTS_FINAL.pdf, retrieved on October 20^{th}, 2018.

[Berto (2018d)] Berto, A. (2018, June 25-27). Rail Yield Management at Trenitalia. Presented at the Joint EURO/ALIO International Conference 2018 on Applied Combinatorial Optimization, Bologna (Italy).

[Berto (2018e)] Berto, A. (2018, April 16-19). Rail Yield Management at Trenitalia. IBM approach. Presented at Transport Research Arena – TRA, Contest Visions 2018, Young Researchers, Wien (Austria). The project abstract can be found in the TRAVisions 2018 Contest book: https://www.travisions.eu/TRAVisions/javax.faces.resource/upload/travision2018_web.pdf.xhtml?ln=bootstrap-layout, retrieved on October 20^{th}, 2018.

[Berto and Gliozzi (2017)] Berto, A. (2017, September 20-22). Rail Revenue Management at Trenitalia. Presented at ECSO 2017, 2^{nd} European Conference on Stochastic Optimisation, Stochastic Optimization in Service Science, Invited Session on Stochastic and Robust Optimization for Railway Operations Management II, Rome (Italy). The project abstract can be found in http://easychair.org/smart-program/ECSO2017/, retrieved on October 20^{th}, 2018.

[Berto and Gliozzi (2015)] Berto, A. and Gliozzi, S. (2015). Rail Yield Management at Trenitalia. Presented at AIRO 2015, 45^{th} Annual Conference of the Italian Operational Research Society, Pisa (Italy). The project abstract can be found in http://www.airo2015.com/libro-degli-atti.pdf, retrieved on October 20^{th}, 2018

[Bharill and Rangaraj (2008)] Bharill R., Rangaraj N., Revenue management in railway operations: A study of the Rajdhani Express, Indian Railways, Transportation Research Part A 42 (2008) 1195–1207

[Bitran and Caldentey (2002)] Bitran G., Caldentey R., An Overview of Pricing Models for Revenue Management December, 2002

[Blanchet et al. (2016)] Blanchet, J. H., Gallego, G., Goyal, V., 2016. A Markov chain approximation to choice modeling. Operations Research 64 (4), 886–905.

[Blau (1964)] Blau, P.M., 1964. Exchange and Power in Social Life. New York. NY: John Wiley & Sons, Inc.

[Breffni et al. (2003)] Breffni M. Noone, Sheryl E. Kimes and Leo M. Renaghan, Integrating customer relationship management and revenue management: A hotel perspective, Journal of Revenue and Pricing Management Volume 2 Number 1, 2003

[Brown (1967)] Brown, R. (1967). Decision Rules for Inventory Management. New York: Holt, Rinehart and Winston

[Brummer et al. (1988)] Brummer, M. et al., "Determination of Potential Load Capacity Based on Observed Load Factor Data," A study for Northwest Airlines: St. Olaf College Undergraduate Practicum Group (1988).

[Campos and de Rus (2009)] Campos J., de Rus G.(2009). Some stylized facts about high-speed rail: A review of HSR experiences around the world. Transport Policy, 16(1); 19–28.

[Cascetta and Coppola (2015)] Cascetta, E., Coppola, P., 2015. New high-speed rail lines and market competition: short-term effects on services and demand in Italy. Transp. Res. Rec. J. Transp. Res. Board 2475, 8-15.

[Cascetta and Coppola (2014)] Cascetta, E., Coppola, P., 2014. Competition on fast track: an analysis of the first competitive market for HSR services. Procedia Soc. Behav. Sci. 111 (2014), 176-185.

[Cascetta et al. (2013)] Cascetta, E., Coppola, P., Velardi, V., 2013. High-speed rail demand: before-and-after evidence from the Italian market. disP - the planning review, ETH, Zurich, 193 (2/2013).

[Cascetta and Coppola (2012)] Cascetta E., Coppola P. (2012). An elastic demand schedule-based multimodal assignment model for the simulation of high speed rail (HSR) systems. Euro Journal on Transportation and Logistics, 1(1); 3-27.

[Chen and Homem-de Mello (2010)] Chen, L., Homem-de Mello, T., 2010. Mathematical programming models for revenue management under customer choice. European Journal of Operational Research 203 (2), 294–305.

[Chiang et al. (2007)] Wen-Chyuan Chiang, Jason C.H. Chen and Xiaojing Xu, An overview of research on revenue management: current issues and future research, Int. J. Revenue Management, Vol. 1, No.1, 2007, pages 97-128

[Ciancimino et al. (1999)] Ciancimino A., Inzerillo G., Lucidi S., Palagi L., A Mathematical Programming Approach for the solution of the railway yield Management Problem, Transportation Science, 33, 2, ProQuest, Page 168, 1999

[Cizaire and Belobaba (2013)] Claire Cizaire and Peter Belobaba, Joint optimization of airline pricing and fare class seat allocation, Journal of Revenue and Pricing Management (2013) 12, 83–93.

[Cleophas (2009)] Cleophas, C. (2009). Simulation-Based Analysis of Forecast Performance Evaluations for Airline Revenue Management. PhD thesis, University of Paderborn.

[Cleophas et al. (2009a)] Cleophas, C., Frank, M., and Kliewer, N. (2009a). Recent developments in demand forecasting for airline revenue management. International Journal of Revenue Management, 3(3):252-269.

[Cleophas et al. (2009b)] Cleophas, C., Frank, M., and Kliewer, N. (2009b). Simulation-based key performance indicators for evaluating the quality of airline demand forecasting. Journal of Revenue and Pricing Management, 8(4):330-342.

[Cooper and Li (2012)] Cooper, W. L., Li, L., 2012. On the use of buy up as a model of customer choice in revenue management. Production and Operations Management 21 (5), 833–850.

[De Boer et al. (2000)] De Boer S. V., Freling R., Piersma N., Stochastic Programming for Multiple-Leg Network Revenue Management, Report EI-9935/A

[Duncanson (1974)] Duncanson, A., "Short-Term Traffic Forecasting, " Proceedings of the 14^{th} Annual AGIFORS Symposium (1974).

[ENAC (2010)] ENAC (2010). Dati di traffico degli scali italiani a cura della Direzione Sviluppo Aeroporti. Roma, Italy

[ExPretio (2009)] ExPretio Technologies Inc., Choice-Based Revenue Optimization applied to the Railway Industry, Technical White Paper, 2009.

[Federgruen and Zipkin (1984)] Federgruen, A., and Zipkin. P. (1984). Allocation Policies and Cost Approximations for Multi-Location Inventory Systems. Naval Res. Logist. Quart. 31, 97-130.

[Fiig and Belobaba (2010)] T. Fiig, P. Belobaba, Optimization of mixed fare structures: Theory and applications, Journal of Revenue and Pricing Management (2010) 9, 152–170.

[Fisher (1971)] Fisher R. A., The Design of Experiments, Macmillan, 9^{th} Edition, 1971.

[Florian and Klein, 1971] Florian, M., and Klein, M. (1971). Deterministic Production Planning with Concave Costs and Capacity Constraints. Mgmt. Sci. 18, 18-20.

[Gallego and Topaloglu (2014)] Gallego, G., Topaloglu, H., 2014. Constrained assortment optimization for the nested logit model. Management Science 60 (10), 2583–2601.

[Gallego et al. (2015)] Gallego G., Li A., Truong V., Wang X., Online Resource Allocation with Customer Choice, Submitted to Operations Research, 2015

[Gama (2017)] Gama A., Own and cross-price elasticities of demand for domestic flights and intercity trains in the U.S., Transportation Research Part D 54 (2017) 360–371

[Giuricin (2018)] Giuricin, A. (2018). Competition is key as high speed rail keeps growing. Railway Gazette International. May 2018

[Givoni (2006)] Givoni M. (2006). Development and Impact of the Modern High-speed Train: A Review. Transport Reviews, 26(5), 593–611.

[Gliozzi et al. (2014)] Gliozzi S., Caponio L., Biancone M., (2014), Notes on the Inventory Properties and their controls in a Multi-leg topology, Project Technical Report, IBM Italy (Unpublished).

[Gliozzi (2006)] Gliozzi S., (2006), Note Metodologiche sul Modello Decisionale e Previsionale Constrained del Sistema YMS ES, Project Technical Report, IBM Italy (Unpublished).

[Gliozzi and Marchetti (2003)] Gliozzi, S. and Marchetti, A. M. (2003). A new Yield Management Approach for Continuous

Multi-Variable Bookings: The Airline Cargo case. In T. Ciriani, G. Fasano, S. Gliozzi, R. Tadei (Eds.) Operations Research in Space and Air, pp.369-391, Kluwer Academic Publisher, ISBN/ISSN 1402012187.

[Gliozzi and Marchetti (2008)] Gliozzi, S., Marchetti, A. M. (2008). Air Cargo Yield Management System For Utilizing Booking Profiles And Unconstrained Demand. US Patent Office (USPTO), US 7,430,518 B2 (Sept. 30, 2008), US 8,117,055 B2 (Feb. 14, 2012), US 8,321,252 B2 (Nov. 27, 2012).

[Goyal, 1974] Goyal, S. K. (1974). Optimal Ordering Policy for a Multi- Item, Single Supplier System. Opns. Res. 25, 293-298.

[Granger and Pesaran (2000)] Granger C. W. J., Pesaran M. H., A Decision Theoretic Approach to Forecast Evaluation, Statistics and Finance, pp. 261-278 (2000)

[Harris and Marucci (1983)] Harris, P. Marucci, G. "A Short-Term Forecasting Model," AGIFOR Symposium Proceedings 23, Memphis, TN (1983).

[Hetrakula and Cirillo (2014)] Hetrakula P., Cirillo C., A latent class choice based model system for railway optimal pricing and seat allocation, Transportation Research Part E: Logistics and Transportation Review Volume 61, January 2014, Pages 68-83

[Hood (2000)] I.S.A. Hood, Merlin: a model to evaluate revenue and loadings for Intercity, Yield management: Strategies for the service industries, 2000, pages 98-119.

[Jagabathula and Vulcano (2017)] Jagabathula S., Vulcano G., 2017. A partial-order-based model to estimate individual preferences using panel data, working thesis, New York University.

[Jagabathula and Rusmevichientong (2016)] Jagabathula S., Rusmevichientong P., A Nonparametric Joint Assortment and Price Choice Model, Management Science, Vol. 63, No. 9, 2016

[JDA (2009)] JDA, 2009 - Inventor: Ronald P. Menich, Dominic Beveridge. Current Assignee: JDA Software Group Inc, JDA SOFTWARE Inc ; Travel price optimization (tpo), US Patent Application: US20120116844A1, 2009-04-14

[Kalyanaram Little (1994)] Kalyanaram, G., Little, J.D.C., 1994. An empirical analysis of latitude of price acceptance in consumer package goods. Journal of Consumer Research 21 (3), 408–418

[Kanafani (1983)] Kanafani, A.K., Transportation Demand Analysis, McGraw-Hill (1983).

[Kouro et al. (2012)] Kouro, S., Rodriguez, J. Wu, B. Bernet, S. and Perez, M. (2012). Powering the future of industry: High-power adjustable speed drive topologies. IEEE Ind. Appl. Mag., 18(4), 26–39.

[Kraft et al. (2000)] Kraft, E R, Srikar, B N, Phillips, R L, Revenue Mangement in Railroad Applications, 2000

[L'heureux (1986)] L'heureux, E., "A New Twist in Forecasting, Short-Term Passenger Pickup," 26th AGIFORS Symposium (October 19-24, 1986).-257-

[Leary (2010)] Leary, M.R. (2010). Affiliation, acceptance, and belonging. New York, NY: Wiley.

[Lederer and Yeoman (2003)] Lederer P., Yeoman I., Future of Revenue Management - The natural extension of marketing, Journal of Revenue and Pricing Management Volume 2 Number 1, 2003

[Lee et al.(1990)] Lee, Anothony, Owen, Airline Reservations Forecasting: Probabilistic and Statistical Models of the Booking Process, MIT Flight Transportation Laboratory Report, R 90-5, 1990.

[Lewbel and Nesheim, 2010] Lewbel, A. and Nesheim, L. (2018), Sparse demand systems: corners and complements, Boston College, First draft March 2013, revised October 2018.

[Legohérel et al. 2010] Legohérel P., Poutier E., Fyall A. (2013). Revenue Management for Hospitality & Tourism, Chapter 3, The Role of the Revenue Manager, pages 37-39, Goodfellow Publishers.

[Link (2004)] Link H., PEP—A Yield-Management Scheme for Rail Passenger Fares in Germany, Railway fare System, part 2, Japan Railway & Transport Review, Vol. 38, pages 50-55, 2004

[Liu and Van Ryzin (2008)] Liu, Q., Van Ryzin, G., 2008. On the choice-based linear programming model for network revenue management. Manufacturing and Service Operations Management 10 (2), 288– 310.

[Mack 2018] Mack C. (2018), From Data to Decisions: Measurement, Uncertainty, Analysis, and Modeling, CHE 379/384, The University of Texas at Austin, http://www.lithoguru.com/scientist/statistics/course.html , retrieved on September 29th, 2018

[Maddala (1983)] Maddala, G. S., Limited-Dependent and Qualitative Variables in Econometrics, Cambridge University Press (1983).

[Mahajan and van Ryzin (2001)] Mahajan, S. and van Ryzin, G. (2001). Inventory competition under dynamic consumer choice. Operations Research, 49(5); 464-657.

[Mahajan and van Ryzin (2001)] Mahajan, S., Van Ryzin, G., 2001. Stocking retail assortments under dynamic consumer substitution. Operations Research 49 (3), 334–351.

[Marchetti (2004)] Marchetti A. M. M., (2004), Note Metodologiche sul Sistema YMSES, Project Technical Report, IBM Italy (Unpublished).

[Masatlioglu and Nakajima (2013)] Masatlioglu, Y., Nakajima, D., 2013. Choice by iterative search. Theoretical Economics 8 (3), 701–728.

[Mason et al. (2003)] Mason R. L., Gunst R. F., Hess J. L., Statistical Design and Analysis of Experiments With Applications to Engineering and Science, John Wiley & Sons, Inc., Second Edition, 2003.

[McFadden and Train (2000)] McFadden D., Train K., Mixed MNL models for discrete response, Journal of Applied Economietrics, 2000

[McGill and Van Ryzin (1999)] McGill, J.I., Van Ryzin, G.J., 1999. Revenue management: Research overview and prospects. Transportation Science 33 (2), 233–256

[Milenković and Bojović (2016)] Milenković, M., Bojović N., (2016), Chapter 5: Railway Demand Forecasting, Handbook of Research on Emerging Innovations in Rail Transportation Engineering, B. Umesh Rai, Chennai Metro Rail Limited India, A volume in the Advances in Civil and Industrial Engineering (ACIE) Book Series, Engineering Science Reference (an imprint of IGI Global), 2016.

[Mitev (1998)] Mitev, N. N. (1998). A comparative analysis of information technology strategy in American Airlines and French Railways, Proceedings of the Thirty First Hawaii International Conference on System Sciences, (cat. 98TB100216),6, pp. 611–621.

[Montgomery (2001)] Montgomery D. C., Design and Analysis of Experiments, John Wiley & Sons, 5th Edition, 2001

[Morlotti et al. (2017)] Chiara Morlotti, Mattia Cattaneo, Paolo Malighetti, Renato Redondi, Multi-dimensional price elasticity for leisure and business destinations in the low-cost air transport market: Evidence from easyJet, Tourism Management 61 (2017) 23-34

[Muckstadt and Roundy (1988)] Muckstadt, J., Roundy, R. (1988). Analysis of Multi- Stage Production Systems. Technical Report No. 806, School of O.R. and I.E., Cornell University, Ithaca, New York.

[Mumbower et al. (2014)] Stacey Mumbower a, Laurie A. Garrow a, Matthew J. Higgins b,c, Estimating flight-level price elasticities using online airline data: A first step toward integrating pricing, demand, and revenue optimization, Transportation Research Part A 66 (2014) 196–212

[Newman et al. (2014)] Newman, J. P., Ferguson, M. E., Garrow, L. A., Jacobs, T. L., 2014. Estimation of choice-based models using sales data from a single firm. Manufacturing and Service Operations Management 16 (2), 184–197.

[Oehlert (2000)] Oehlert G. W., A first course in design and analysis of experiments, W. H. Freeman; 1st edition (January 19, 2000)

[Parla (1988)] Parlar, M. (1988). Game Theoretic Analysis of the Substitutable Product Inventory Problem with Random Demands. Naval Research Logistics. 35(1), 397-409.

[Pearl (2009)] Pearl J., Causality, Cambridge University Press, 2^{nd} Edition, 2009

[Phadke (2013)] Phadke M. S. (2013), Design of Experiments for Software Testing, iSixSigma, https://www.isixsigma.com/tools-templates/design-of-experiments-doe/design-experiments-software-testing/ , October 25^{th}, 2018.

[Roch et al. (2005)] S. Roch, G. Savard and P. Marcotte, “An Approximation algorithm for Stackelberg network pricing”, Networks, Number 46, 2005, pp 57-67

[Roundy (1985)] Roundy, R. (1985). 98%-Effective Integer-Ratio-Lot-Sizing for One-Warehouse-Multi-Retailer-Systems. Mgmt. Sci. 31, 1416-1430.

[Sa (1987)] Sa, J., Reservations Forecasting in Airline Yield Management, Masters Thesis, Flight Transportation Laboratory, MIT, Cambridge, MA (1987).

[Schneider (1986)] Schneider, H., Truncated and Censored Samples from Normal Populations, New York, Marcel Dekker (1986).

[Seltman (2018)] Seltman H. J., Experimental Design and Analysis, July 11, 2018

[Shen and Su (2007)] Shen, Z.-J. M., Su, X., 2007. Customer behavior modeling in revenue management and auctions: A review and new research opportunities. Production and Operations Management 16 (6), 713–728.

[Schroeder et al. (2000)] Schroeder M., Braun I., Eckehard Schnieder, Revenue Management strategies in the railway industry, Forschungs und Technologie Zentrum der Deutschen Bahn AG

[Sibdari et al. (2007)] Sibdari S., Kyle Y. Chellappan L. and S., Multiproduct revenue management: An empirical study of Auto Train at Amtrak, Journal of Revenue and Pricing Management Vol. 7, 2 172–184, 2007

[Simsek and Topaloglu (2017)] Simsek, S., Topaloglu, H., 2017. Technical note: An expectation-maximization algorithm to estimate the parameters of the Markov chain choice model, working thesis, Cornell University.

[Smith et al. (1992)] Smith, B. C., Leimkuhler, J. F., and Darrow, R. M. (1992). Yield management at American Airlines. Interfaces, 22(1):8-31.

[Talluri and Van Ryzin (2004)] Talluri, K., Van Ryzin, G., 2004. Revenue management under a general discrete choice model of consumer behavior. Management Science 50 (1), 15–33.

[Talluri and Van Ryzin (2004b)] Talluri, K., Van Ryzin, G., 2004. The theory and practice of revenue management, Kluwer Academic Publishers.

[Talluri et al. (2010)] Talluri K., Castejon F., Codin B., and Magaz J., Proving the performance of a new revenue management system, Practice Article, Journal of Revenue and Pricing Management Vol. 9, 4, 300–312, 2010

[Taneja (1978)] Taneja, N.K., Airline Traffic Forecasting, Lexington Books, Lexington, MA. (1978).

[Temath (2010)] Temath C., Performance Measurement in Airline Revenue Management – A Simulation-based Assessment of the Network-based Revenue Opportunity Model, Dissertation, 2010

[Topaloglu (2009)] Topaloglu, H., 2009. Using Lagrangian relaxation to compute capacity-dependent bid prices network revenue management. Operations Research 57 (3), 637–649

[Train (2009)] Train, K., 2009. Discrete choice methods with simulation, 2^{nd} Edition. Cambridge University Press.

[Tsekeris and Tsekeris (2011)] Tsekeris T., Tsekeris C. (2011), Demand Forecasting in Transport: Overview and Modeling Advances, Economic Research-Ekonomska Istraživanja, 24:1, 82-94.

[Van Ryzin (2008)] Van Ryzin, G., Vulcano, G., 2008. Computing virtual nesting controls for network revenue management under customer choice behavior. Manufacturing and Service Operations Management 10 (3), 448–467.

[Van Ryzin (2015)] Van Ryzin, G., Vulcano, G., 2015. A market discovery algorithm to estimate a general class of nonparametric choice models. Management Science 61 (2), 281–300.

[Van Ryzin (2017)] Van Ryzin, G., Vulcano, G., 2017. An expectation-maximization method to estimate a rank-based choice model of demand. Operations Research 65 (2), 396–407

[Vinod (2006)] Vinod B., (2006), Measuring Revenue Management Performance, Ascent, (1): 14-19.

[Vulcano et al. (2012)] Vulcano G., Van Ryzin G., Ratliff R., Estimating Primary Demand for Substitutable Products from Sales Transaction Data, Operations Research, Vol. 60, No. 2, March–April 2012, pp. 313–334

[Vulcano et al. (2010)] Vulcano G., Van Ryzin G., Chaar W., OM Practice, Choice-Based Revenue Management: An Empirical Study of Estimation and Optimization, Manufacturing & Service Operations Management Vol. 12, No. 3, Summer 2010, pp. 371–392, 2010

[Wang and Sahin (2014)] Wang R. and Sahin O. (2014), The Impact of Consumer Search Cost on Assortment Planning and Pricing, Management Science, Forthcoming

[Wang et al. (2016)] Wang X. , Wang H., Zhang X., Stochastic seat allocation models for passenger rail transportation under customer choice, Transportation Research Part E: Logistics and Transportation Review, Volume 96, December 2016, Pages 95-112

[Watling and Cantarella (2013)] Watling D. P., Cantarella G. E. (2013), Modelling sources of variation in transportation systems: theoretical foundations of day-to-day dynamic models, Transportmetrica B: Transport Dynamics, 1:1, 3-32.

[Weatherford (1999)] Weatherford, L., Forecast Aggregation and Disaggragation, IATA Revenue Management Conference Proceedings (1999).

[Wickham and Richard (1995)] Wickham, Richard R., Evaluation of Forecasting Techniques for Short-Term Demand of Air Transportation, MIT Flight Transportation Laboratory Report, R 95-7 (May, 1995).

[Williamson (1992)] Williamson, E. L., Airline Network Seat Inventory Control, MIT Flight Transportation Laboratory Report, R 92-3 (June 1992).

[Xiaoqiang et al. (2017)] Xiaoqiang Z., Lang M., Jin Z., Dynamic pricing for passenger groups of high-speed rail transportation, Journal of Rail Transport Planning and Management, Volume 6, Issue 4, January 2017, Pages 346-356

[You (2008)] You P., (2008), An efficient computational approach for railway booking problems, European Journal of Operational Research, Volume 185, Issue 2, 1 March 2008, Pages 811-824

[Zatta (2007)] Zatta, D. (2007). Revenue Management – Come ottimizzare l'uso delle risorse aziendali per massimizzare i profitti, Hoepli, ISBN/ISSN 9788820338633.

[Zeni (2001)] Zeni, R. H. (2001). Improved Forecast Accuracy in Revenue Management by Unconstrained Demand Estimates from Censored Data. Phd Dissertation. Dissertation.com, ISBN/ISSN 15811214152001.

[Zhang and Adelman (2009)] Zhang, D., Adelman, D., 2009. An approximate dynamic programming approach to network revenue management with customer choice. Transportation Science 43 (3), 381–394.

[Zhao and Atkins (2002)] Zhao, X. and Atkins, D. (2002a). Strategic Revenue Management under Price and Seat Inventory Competition. Working Paper, University of British Columbia.

[Zhao and Kockelman (2002)] Zhao Y., Kockelman K. M., The Propagation of Uncertainty through Travel Demand Models: an Exploratory Analysis, Annals of Regional Science 36 (1):145-163, 2002.