A taxonomy of demand uncensoring methods

3 authors:

Shadi Sharif Azadeh
Erasmus University Rotterdam/ École Polytechnique Fédérale de Lausanne/ MIT

Patrice Marcotte
Université de Montréal

Gilles Savard
Polytechnique Montréal

Research Article

A taxonomy of demand uncensoring methods in revenue management

Received (in revised form): 17th December 2013

Shadi Sharif Azadeh[a], Patrice Marcotte[b] and Gilles Savard[a]

[a]Polytechnique Montreal, Montreal, Canada; and [b]University of Montreal, Montreal, Canada

Shadi Sharif Azadeh completed her Masters and PhD in Mathematics (Operations Research) at Polytechnique Montreal in 2013. She is a member of GERAD (Group for Research in Decision Analysis) and CIRRELT (Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation). She is presently a postdoctoral fellow at Polytechnique Montreal.

Patrice Marcotte is Professor and Chairman at the Computer Science and Operations Research Department of the University of Montreal. He has published more than 70 papers in the field of optimization, focusing on variational inequalities and bilevel programming, with applications in network equilibrium and revenue management. He is currently Area Editor (continuous optimization) of the journal *Operations Research Letters*, as well as Associate Editor for *Operations Research and Transportation Science*.

Gilles Savard is Professor and Dean of Research and Innovation at Polytechnique Montreal and member of the GERAD (Group for Research in Decision Analysis) and CIRRELT. He has published more than 60 papers in major refereed scientific journals. He is a leading expert in both the theory and practice of bilevel programming and in the development of global algorithms for non-convex optimization.

Correspondence: Shadi Sharif Azadeh, Polytechnique Montreal, 2900 Boulevard Edouard-Montpetit, Quebec, 2900, Montreal, Canada

ABSTRACT Revenue management systems rely on customer data, and are thus affected by the absence of registered demand that arises when a product is no longer available. In the present work, we review the uncensoring (or unconstraining) techniques that have been proposed to deal with this issue, and develop a taxonomy based on their respective features. This study will be helpful in identifying the relative merits of these techniques, as well as avenues for future research.
Journal of Revenue and Pricing Management advance online publication, 28 March 2014; doi:10.1057/rpm.2014.8

Keywords: revenue management; demand forecasting; uncensoring; statistical methods; optimization; customer choice behaviour

INTRODUCTION

The purpose of Revenue Management (RM) is to enhance the profitability of a firm through the optimal management of its inventory. In the service industry (airlines, railways, hotels), this can be achieved by controlling the availability of products, in order to redirect customers to 'products' with high profit margins. Throughout this process, a trade-off must be stricken between the sale of low-cost products when resources are plentiful, and the protection of high-fare products towards the end of the booking horizon. Any such strategy is highly dependent on historical demand forecasts and must cope with the lack of information resulting from censored demand, that is, virtual demand for products that have been withdrawn, because of their 'booking limits' being reached. This demand may either be lost (spill) or recaptured by a more expensive (buy-up) or cheaper (buy-down) available product. In either case, the observed demand of the remaining available alternatives does not reflect the customer's first choice, and may yield unreliable estimates. While one would like to obtain accurate estimates of the entire preference lists, this ideal cannot be achieved in practice, and one must settle for trade-offs. It must also be stressed that no universally 'superior' method being available, the choice will depend on the context. Is it static or dynamic? Are seasonal effects important? Should decisions be taken in real time? What is the quality of the available data? For instance, the arrival processes in the airline and rail industries are different, with rapid changes being more frequent in the former, requiring decisions to be taken in *quasi*-real time. There are other key differences between the two modes of transportation. In the rail industry, the purchase date behaviour does not vary across customer segments. In sharp contrast, leisure travellers purchase their airline tickets well in advance, to take advantage of reduced fares, while business travellers frequently book their flights close to the departure date.

Another relevant example is the competition between low-cost airlines and high-speed trains in medium-range European markets, which hardly exists in North America, and requires both industries to embed modal split within their demand forecasts. In an altogether different context, seasonal effects and data availability in the retail business (fashion seasons) cannot be compared with those of a transportation company, and call for different demand estimation techniques.

According to Weatherford and Belobaba (2002), underestimating demand by 12.5–25 per cent can result in a loss of revenue from 1 per cent to 3 per cent, which is significant. The main goal of this article is to review and propose a taxonomy for the techniques that have been developed to address the issue of missing data.

Following a simple illustration involving censored data, the remaining of this introductory section will put the issue into proper perspective within the RM literature.

Let us consider a service company that sells a high-fare 'product' A and a low-fare product B. An arriving customer may wish to purchase A, B or renege. As long as both products are available, that is, the booking limits have not been reached, sale figures (registered demand) reflect actual demand. If the booking limit set by the RM policy is reached for product B first, the upcoming demand for B is either transferred to A (buy-up) or lost (spill). This is illustrated in Figure 1 for given streams of arrivals. Note that, if streams A and B are independent, then the native demand for A should not change once B is closed. Even in such a simple case, one realizes the difficulties of retrieving the true demand from incomplete historical data, and of striking the right balance between accuracy and practicality in real-life instances. This leads to a variety of approaches, which have been investigated from different viewpoints:

- Wickham (1995) probed statistical forecasting methods for short-term demand in the airline industry. Time series, linear regression and booking pickup models were considered to estimate demand where some historical data is missing.
- Lee (1990) and McGill and van Ryzin (1999) introduced a variety of statistical methods to extract demand features using registered booking data.
- Zeni (2001) and Weatherford (2000) investigated statistical unconstraining techniques at a micro level. They integrated techniques such as imputations or expectation-maximization (EM) within the framework of exponential smoothing, time series, linear regression and pickup models.
- In van Ryzin (2005), the focus shifted from traditional product demand models to the analysis of customer behaviour, based on the theory of discrete choice (random utility).

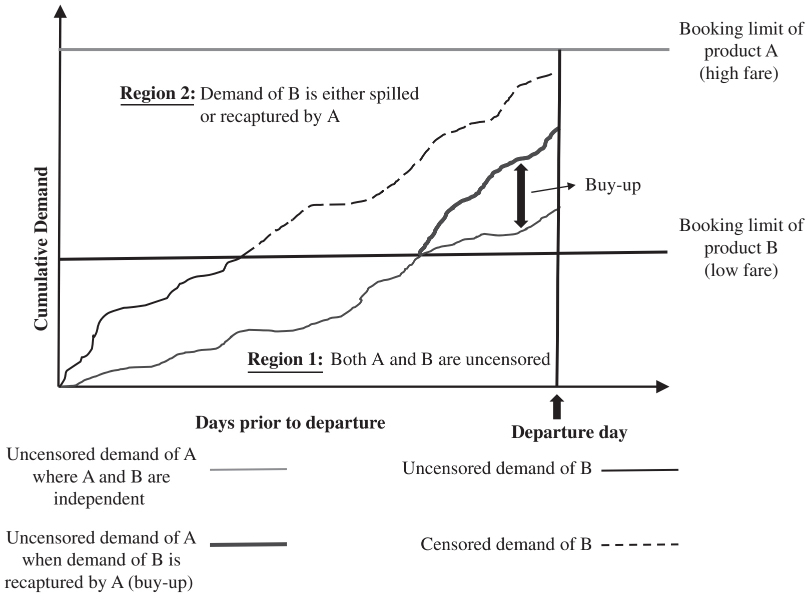

Figure 1: Demand censorship.

- For an airline application, Ratliff *et al* (2008) integrated product dependencies and proposed a hierarchical classification of previous unconstraining models. Three frameworks were considered: (i) single-class models, where product demand is assumed to be independent, (ii) multi-class models, with up-sell and down-sell among different fare classes, (iii) multi-flight methods, which include the most general unconstraining approaches. All models take into account the interactions between the various fare products.

Although the above-mentioned studies cover important subsets of uncensoring methods, there yet exists a need to structure the field, so that adequate techniques be easily matched to areas of application. Hence, our proposal for a flexible and expandable taxonomy should prove to be useful in future settings.

Our classification makes use of a tuple notation, which allows for a concise review of existing models, identifying the key elements that distinguish them from one another. In this framework, a model is represented as a tuple $[\mu|\delta|\alpha]$, where μ is the set of attributes of the supplier, δ is the set of demand features and α identifies uncensoring approaches. Whenever an element is not considered in a specific model, it does not appear in the tuple.

The remainder of the article is structured as follows. In subsections 2.1 and 2.2, we respectively introduce the supply- and demand-side assumptions upon which our demand modelling can be set within the framework of RM. In Section 3, we provide a thorough review of the uncensoring literature, whose methods are partitioned into four classes: basic, statistical, choice based and optimization based. In Section 4, the main contributions of the field are presented and a taxonomy is introduced, allowing a classification according to key parameters. Avenues for further research are outlined in the concluding Section 5.

FEATURES OF DEMAND MODELS IN RM SYSTEMS

In this section, we introduce the features according to which the demand models will be classified.

1. Supply side
   - *Customer type (for data gathering)*

- *Domain of application*
- *Competition*

2. Demand
   - *Dependencies among products*
   - *Diversion* (spill or recapture)
   - *Seasonality*
   - *Segmentation (internal or external)*

They are detailed, together with their respective domains of application, in the following subsections.

Supply-side features

Forecasting techniques can address demand either at the macro or micro level. Macro-level analyses consider total demand, whereas micro-level forecasting is typically conducted on a booking date and fare-class basis (see Lee, 1990; Zeni, 2001). In the context of micro-level forecasting (which is the main focus of this review), supply-side assumptions influence the choice of uncensoring method to a great extent. We consider the following classification:

1. Customer type (μ_1)
   - *Myopic*
   - *Strategic*
2. Domain of application (μ_2)
   - *Airline*
   - *Rental–Retail*
   - *Railway*
   - *Hotel*
3. Competition (μ_3)
   - *Competition*

The parameter $\mu_1 \in \{myop, strat\}$ refers to inter-temporal substitutions that involve (or not) delaying one's purchase (see Shen and Su, 2007). In the standard models, *myopic* customers make their final decision at the time of arrival, whereas more recent models allow *strategic* customers to reconsider their choice in the future (see Su, 2007; Liu and van Ryzin, 2008; Cachon and Swinney, 2009; Yin *et al*, 2009; Levin *et al*, 2010; Yang *et al*, 2010; Cachon and Feldman, 2011; Cachon and Swinney, 2011; Swinney, 2011; Bansal, 2012). In our context, it is desirable to classify customer type as a supply-side attribute, especially when the system involves auctions or loyalty programs that are aimed at specific segments of the population.

The second attribute $\mu_2 \in \{\text{air, rent−ret, rail, hotel}\}$ refers to application domains. Although the initial research focused on airlines, RM has subsequently made its way into the realms of rental and retail (Ja *et al*, 2001; Stefanescu *et al*, 2004; Zhu, 2006; Conlon and Mortimer, 2008; Ratliff *et al*, 2008; Stefanescu, 2009; Talluri, 2009; Armstrong and Meissner, 2010; Haensel and Koole, 2010; Vulcano *et al*, 2010; Haensel *et al*, 2011; Huh *et al*, 2011; Crevier *et al*, 2012). In the hotel industry, Queenan *et al* (2009) have assessed unconstraining techniques using actual data, whereas Haensel and Koole (2010) have done so for the hotel industry. Other proposals can be found in the recent literature (see Bodea, 2008; Ferguson and Queenan, 2009; Meissner *et al*, 2012).

The last parameter μ_3 refers to *competition*. Surprisingly, this feature of RM, which actually motivated the very field, has only recently been paid close attention (see Belobaba, 1987; Perakis and Sood, 2006; Jiang, 2007; Gallego and Hu, 2014; Kwon *et al*, 2009; Jiang and Pang, 2011; Martínez and Talluri, 2011). Including competition within a RM system can significantly impact the demand model, which could embed competition between companies that offer similar products. For instance, Netessine and Shumsky (2005) have considered quantity-based games of booking controls under horizontal and vertical competition, and Liu and Zhang (2011) have addressed the issue of dynamic pricing competition between two firms offering vertically differentiated products to strategic consumers.

Demand characteristics

The impact of demand representation over the choice of an unconstraining technique is important. We characterize the demand process through the following four attributes δ_i:

1. Product dependency (δ_1)
   - *Dependent*
   - *Independent*
2. Diversion (δ_2)
   - *Spill*
   - *Recapture*
3. Seasonality (δ_3)
   - *Seasonal effects*
4. Segmentation (δ_4)
   - *Internal* (latent characteristics, such as income)
   - *External* (time dependent arrivals)

We now discuss the parameters in some detail. In reality, products are correlated. However, since this significantly increases the complexity of the demand models, this feature is frequently ignored (that is, the product independence assumption (δ_1)) when large problem instances are considered (Haensel and Koole, 2010; Chapter 2 of Zeni, 2001; Section 6 in Queenan *et al*, 2009; Meissner and Strauss, 2012a; Meissner and Strauss, 2012b). However, it is clearly an over simplification, and recent studies have explicitly considered correlations, either linear or non-linear, either inter-temporal or not (McGill, 1995). Actually, competition among different products within a company is addressed through the product dependency assumption.

In most situations, customers who are denied their preferred choice have recourses within service companies' products. If products are nested, that is, a discontinued low fare cannot be reactivated further in time, then customers can either purchase at a higher fare (*buy-up*) or renege (*spill*) (Swan, 1979; Swan, 1999). Its demand could be seized by other products from the same company or it could be captured by a competitor (δ_2). Some researchers have considered mass balance equations that link spill and recapture (Andersson, 1998; Ja *et al*, 2001; Ratliff *et al*, 2008). Vulcano *et al* (2010) and Vulcano *et al* (2012) have considered a method for estimating substitute and turned away demand for the case of incomplete data. Similar research has been conducted by Talluri and van Ryzin (2004) and Haensel and Koole (2010).

In service companies such as rentals, airlines, railways and hotels, reservations are made days or weeks in advance. These periods of time are divided into *booking intervals* during which customers register for a specific day. Reservations usually face a considerable degree of seasonality, which may be inadequately captured if only a small portion of data is used to estimate the parameters of the model. In order to capture *seasonal effects* (δ_3), the time scope of the historical data needs to be specified, that is, one must determine the number of booking intervals included in the historical data. Too much information makes the forecasting model inflexible, whereas too little does not allow to capture seasonality in a meaningful fashion. To address the issue, a time-series approach (ARIMA) has been adopted by Lee (1990) in Chapter 5, Sa (1987) and Queenan *et al* (2009).

Market segmentation (δ_4) can affect both the choice of uncensoring and optimization approaches in RM systems. Ideally, one would tailor the fare of a product to the willingness to pay of each individual (see Meissner and Strauss, 2009; Talluri, 2010; Gurbuz *et al*, 2011). In this framework, *internal segmentation* refers to customer features (income, purpose, age and so on), while *external segmentation* refers to time-based customer behaviour. For instance, customers who book late are more likely to be business travellers who opt for high-fare products, while weekenders are more likely looking for economy fares.

The tree-like Figure 2 summarizes the elements of demand models and their related components.

ESTIMATION OF UNCONSTRAINED DEMAND

At the heart of the RM is a twofold process of parameter estimation and optimization. These

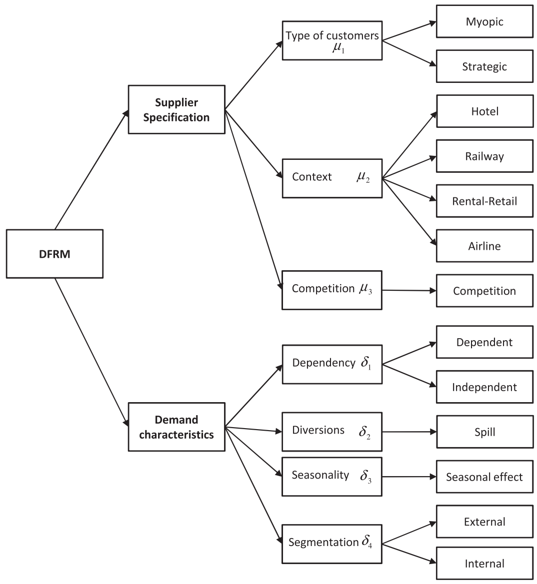

Figure 2: Elements of demand forecasting revenue management.

can be conducted sequentially (estimate then optimize) or in parallel (estimate and optimize). Depending on the strategy adopted, and also on demand specification, different unconstraining methods can be applied, either parametric or non-parametric. Note that classical forecasting methods, such as time series and linear regression, are unable to properly capture customer behaviour and product availability at a given time.

Uncensoring methods can be classified into four main categories, each one associated with a symbol α_i: basic (α_1), statistical (α_2), choice based (α_3) and optimization based (α_4). The first three categories fit the 'estimate then optimize' framework, whereas optimization methods are of an 'estimate and optimize' nature.

Basic methods

Basic methods are non-parametric and may either (i) limit themselves to observed booking data, (ii) ignore censorship altogether, (iii) discard censored data or (iv) use 'imputations' to make up for missing data. They are now discussed in more detail.

Direct observation

One of the simplest methods to tackle the problem of censored demand is not to tell customers that their required product is unavailable. Rejected requests are then appended to registered bookings, resulting in an unbiased estimation of the true demand. In practice, the existence of several booking outlets (online or not) makes this 'ideal' approach unsuitable, notwithstanding the additional burden of processing this data, and the impediment on the perceived quality of service. Moreover, this strategy could not cope with dynamic variations of customer behaviour (Orkin, 1998; Queenan *et al*, 2009).

Ignoring censorship

Assuming that data is uncensored is tantamount to setting demand estimates to their booking limits, whenever these are reached, and will obviously lead to underestimation (Saleh, 1997; Little and Rubin, 2002; Cooper *et al*, 2006).

Discarding censored data

This strategy limits the size of the sample and may yield either over or underestimation, depending whether products with low or high demand levels are censored. This method usually performs adequately when the arrival process is totally random and the number of sell-outs is small. If these conditions are not fulfilled, a negative bias can occur (Section 2.8 in Saleh, 1997; Zeni, 2001).

Imputations

The term 'imputation' refers to methods that fill in censored demand. A commonly used method is 'mean imputation', whereby censored data is replaced by the mean of registered booking data, whenever the latter is less than the average (Zeni, 2001; Little and Rubin, 2002; Farias, 2007). In a similar fashion, one obtains an imputation based on the median of the historical unconstrained demand, in place of its mean.

Each approach is illustrated on an example involving three available products whose data is displayed in Table 1. In the first part of the table, general information about these products are provided. Registered demand for 'direct observation' method differs from the other three basic methods. Hence, we have represented registered and uncensored demand of this method, separately from 'ignoring censorship', 'discarding censorship' and 'mean imputation'.

We close this subsection with a list of the acronyms corresponding to each uncensoring approach:

$$\alpha_1 = \begin{cases} \textit{dir\_obs} & \text{Directly Observed booking data} \\ \textit{ign\_cen} & \text{Ignore Censorship} \\ \textit{dis\_cen} & \text{Discard Censored data} \\ \textit{imp} & \text{Imputations.} \end{cases}$$

Statistical methods

In RM, statistical methods are broadly expressed in three categories (Lee, 1990; Weatherford and Kimes, 2003): historical, advanced and combined booking models.

Historical booking models

Historical booking models resort to traditional parametric forecasting such as time series, exponential smoothing or linear regression (see Sa, 1987; Pölt, 2000; Weatherford, 2000; Littlewood, 2005; Kachitvichyanukul *et al*, 2012).

Time series describe the random nature of the data, and are based on final booking numbers. Despite their relatively simple mathematical structure, they are rich enough to embody a wide range of data features. For one, the ARIMA model comprises autoregressive and moving average components (Box *et al*, 2011). It can be mathematically expressed as follows:

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (1)$$

where Y_t represents the demand at time t, μ is the mean of a stationary process, θ_i's are random coefficients and ε_t's are uncorrelated random terms with 0 mean and common variance σ_ε^2. The first p terms from Y_{t-1} to Y_{t-p} in the above equation represent the autoregressive component and the second q terms from ε_{t-1} to ε_{t-q} depict the moving average elements. Time series explicitly exploit the correlations between successive data points to improve forecasts.

On the basis of data observed up to time $t-1$, *Simple exponential smoothing* adjusts the next value $\hat{Y}_t$ through the formula

$$\hat{Y}_{t+1} = \hat{Y}_t + \alpha\left(Y_t - \hat{Y}_t\right) \quad (2)$$

where the parameter α lies between 0 (no adjustment) and 1 ('strong' adjustment). This method, which relies on a weighted average of the most recent observations (Hyndman *et al*, 2008), is not recommended for the analysis of time series characterized by a large number of null values and a high variability among the non-zero data.

Linear regression assumes a linear trend of registered bookings in successive time periods, the key issue being to properly select the number and nature of the descriptive variables entering the model. The parameters of the regression are usually estimated via least squares.

Table 1: Uncensoring demand: Basic methods

Product	*A*	*B*	*C*	*D*
General information				
Availability	1	1	0	1
Booking limit	8	6	0	23
Actual demand	15	0	9	22
Uncensored demand by direct observation				
Registered demand (Direct observation)	12	0	0	22
Direct observation	12	0	0	22
Uncensored demand by:				
Registered demand (Other three methods)	8	0	0	22
Ignoring censorship	8	0	0	22
Discarding censorship	—	0	—	22
Mean imputation	10	0	0	22

For a case involving two descriptive variables over two successive booking intervals, we have that

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} \quad (3)$$

where Y_t is the current booking and Y_{t-1}, Y_{t-2} represent the total bookings for the two preceding time intervals. The drawback of this regression model is the underlying linearity assumption, which may not always hold.

Advanced bookings

Advanced booking models (including *pickup*, *advanced pickup*, *booking profile*) are based on registered bookings over time, and can be of the *additive* or *multiplicative* type. Both types have been considered in the transportation literature (see L'Heureux, 1986; Skwarek, 1996a; Skwarek, 1996b; Zickus, 1998; Gorin, 2000; Weatherford and Polt, 2002; Mishra and Viswanathan, 2003; Section 5.5.2 in Lee, 1990; Wickham, 1995; Zakhary *et al*, 2008). Classical or advanced pickups are additive models that differ in their treatment of historical data (Lee, 1990; Wickham, 1995). In classical methods, an overall average on the products that are no longer available represent uncensored demand. However, advanced booking methods base their estimates on incremental dynamic averages to better depict small variations in demand. In general, they assume no proportionality relationship between current registered bookings (at the time when the product is no longer available) and final bookings. Instead, they assume that the absolute growth (pickup) in bookings between the current time interval and the last interval of other open similar products is a good indicator of the booking history, had the closed products been still open. This yields

$$Y_0^i = Y_t^i + \frac{1}{J}\sum_{j=1}^{J}\left(Y_0^j - Y_t^j\right), \quad (4)$$

where Y_0^i is an estimate of the final uncensored booking of product i, J represents the number of remaining booking intervals and Y_t^i is the current booking (at time t) of the closed product. In this formula, the average term corresponds to the mean number of pickups following the closure of available products.

L'Heureux (1986) has suggested that the inclusion of data drawn from all reservation intervals (that is, taking incremental pickups into account) provides valuable information about demand behaviour.

Multiplicative pickup models operate in a similar fashion, but base their forecasts on the 'pickup ratio', defined as

$$pick\_ratio(t,0) = \frac{1}{J}\sum_{j=1}^{J}\left(Y_0^j - Y_t^j\right) \times \frac{1}{Y_t^i} \quad (5)$$

$$Y_0^i = Y_t^i \times pick\_ratio(t,0). \quad (6)$$

It is important to point out that these methods only rely on historical data and neglect socioeconomic or behavioural features of the population. They are of course highly dependent on the quality of the data collection process.

Combined models

Combined models use regression or weighted average of historical and advanced booking models to produce forecasts. In order to achieve high accuracy, they may resort to parametric regression, neural networks or distribution-based demand models. The use of *weighted moving average* allows to emphasize the most recent bookings. Given a set of weights summing up to one, we have

$$\hat{Y}_{t+1} = w_1 Y_t + w_2 Y_{t-1} + w_3 Y_{t-2} + \cdots + w_N Y_{t-N+1}. \quad (7)$$

In this context, Wickham (1995) has implemented both simple and weighted averages and found that they were outperformed by pickup methods (see van Ryzin and McGill, 2000; Ja *et al*, 2001; Liu, 2004).

More advanced techniques, such as *supervised learning neural networks*, are akin to complex non-linear regressions, and are able to process large and complex data sets. A neural network comprises an input layer, one or several hidden layers and an output layer. Individual inputs are processed through the network, and their weighted combination is compared with the neuron's threshold value. In the 'training phase' one iteratively adjusts each weight until the difference between expected bookings and actual data falls below a predefined threshold value. Following this phase, the network is used to predict future demand from a data set that should not differ too widely from the training set. Although neural networks have been applied successfully to transportation demand forecasting (Dantas *et al*, 2000; Weatherford *et al*, 2003; Sharif Azadeh *et al*, 2012), supervised learning technique is yet unable to produce accurate forecasts when a large proportion of historical data is censored.

In *Distribution-based demand models*, it is assumed that the statistical distribution underlying the demand process (usually *Normal* or *Gamma*) is known, and that its parameters (mean, variance and so on) are estimated based on historical data. Alongside the Normal or Gamma assumptions, Brummer *et al* (1988) has considered log-normal distributions, while Logistic, Gamma, Weibull, Exponential and Poisson distributions have been advocated (see Kaplan and Meier, 1958; ZF Li and Hoon Oum, 2000; Swan, 2002; Guo *et al*, 2011; Eren and Maglaras, 2009; Huh *et al*, 2011; Popescu *et al*, 2013).

In the following, the statistical methods are partitioned according to the parameter α_2:

$$\alpha_2 = \begin{cases} hbm & \text{'historical booking models': time series (tseries), exponential} \\ & \text{smoothing (exp\_smooth) and linear regression (lin\_reg)} \\ abm & \text{'advanced booking models': Additive, Multiplicative (pickup),} \\ & \text{and Booking Profile (BP)} \\ cm & \text{'combined models': weighted average (weight\_ave),} \\ & \text{parametric regression (par\_reg), Neural Networks (NN),} \\ & \text{and distribution based demand (dist\_dem).} \end{cases}$$

Choice-based models

The integration of a discrete choice framework (McFadden, 2001) within RM systems has provided the flexibility required to take into account strategic customers who base their decisions on the set of available alternatives (*Choice sets*), under the following restrictions (Train, 2009):

- only one choice can be made at any given time period;
- all available choices are included in the choice set;
- the number of alternatives is finite.

In the discrete choice framework, a customer selects the product that maximizes his expected utility, the latter being expressed as the sum of deterministic and stochastic terms that are related to the features of each product. The choice of the random term results in different models: Probit (normal), Logit (Gumbel), Mixed Logit and so on, and their parameters are typically estimated via maximum likelihood. In the much touted Multinomial Logit model, which involves a Gumbel-distributed random term, the probability that a product i with utility u_i be selected is given by the closed form formula

$$P_i(S_t) = \frac{\exp(u_i)}{\sum_{j \in S_t} \exp(u_j) + 1}, \tag{8}$$

where S_t denotes the subset of products available at time t. We will assign the acronym *cb* (choice based) to the parameter α_3 when discrete choice models are considered.

The embedding of discrete choice models within an optimization process has been considered by Talluri and van Ryzin (2004), Vulcano *et al* (2010), Vulcano *et al* (2012), Haensel and Koole (2010), Haensel *et al*

(2011), Conlon and Mortimer (2008) and Zhang and Cooper (2005). In particular, Belobaba and Hopperstad (1999) have studied the impact of customer behaviour on traditional RM systems, while Talluri and van Ryzin (2004) have characterized optimal control policies in a very general discrete choice setting.

Optimization methods

In recent years, techniques that focus on optimization have been introduced in choice-based RM. These can be broadly divided into four main categories: EM, *Projection Detruncation* (PD), *Double Exponential Smoothing* (DES) and *Non-linear Programming* (NLP). The first three methods are parametric, while non-linear programming covers most non-parametric estimation methods.

Expectation-maximization

After its introduction in the late 1990s by Salch (1997), the two-stage EM process has quickly become one of the most popular unconstraining methods. In the first step, *E-step*, unobserved demand of an unavailable product is replaced by its average observed demand, before its reaching the capacity. In the subsequent *M-step*, the parameters of the demand distribution (mean and variance) are estimated via maximum likelihood. The first step is then repeated and the fixed-point process is halted when no significant progress is observed. In this setting, seasonality is usually ignored.

For a given product[1], let $Y_1, \ldots Y_{N_1}, Y_{N_1+1}, \ldots, Y_{N_1+N_2}$ denote a stream of registered bookings consisting of N_2 uncensored and N_1 censored realizations of the random variable Y. The latter obtained after the product has reached its booking limit. Following common practice, the index of booking intervals decreases from n (in this case, N_1+N_2) to 0, which corresponds to departure time in transportation RM. Assuming that demand follows a normal distribution with mean μ and variance σ^2, the procedure goes through the following steps.

Initialization: Estimate μ and σ, based on N_2 uncensored observed data:

$$\mu = \frac{1}{N_2} \sum_{i=N_1+1}^{N_1+N_2} Y_i \tag{9}$$

$$\sigma = \sqrt{\frac{1}{N_2} \sum_{i=N_1+1}^{N_1+N_2} \left(Y_i - \mu^{(0)}\right)^2} \tag{10}$$

E-Step: For a given number C of constrained observations, the first and second moments of the censored data required to form the log-likelihood function are estimated according to the formula: iteratively to replace the missing data to form the complete log-likelihood function where C represents registered constrained observation.

$$\hat{Y}_i^{(+)} = E[Y \mid Y > C, Y \sim N(\mu, \sigma)] \tag{11}$$

$$\left(\hat{Y}_i^2\right)^+ = E\left[Y^2 \mid Y > C, Y \sim N(\mu, \sigma)\right]. \tag{12}$$

For $i = 1, \ldots, N_1$, the data $Y_i, \ldots, Y_{N_1}$ and $Y_1^2, \ldots, Y_{N_1}^2$ are respectively replaced according to the formula (11) and (12).

M-Step: Maximize the log-likelihood function with respect to μ and σ to obtain μ^+ and σ^+.

Stopping criterion: Repeat steps E and M until the difference between successive iterates is less than some predetermined threshold value δ.

Researchers have recognized the EM method as one the most efficient demand uncensoring techniques in RM (see McGill, 1995; Hopperstad, 1996; Hopperstad, 1997; Pölt, 2000; Weatherford, 2000; Section 2.8.9 in Zeni, 2001; Little and Rubin, 2002; Stefanescu *et al*, 2004; Zeni and Lawrence, 2004; Chen and Luo, 2005; Talluri and van Ryzin, 2005; He and Luo, 2006; Hopperstad *et al*, 2007; Stefanescu, 2009; Guo *et al*, 2011; Haensel and Koole, 2010; Karmarkar *et al*, 2010; Haensel *et al*, 2011; Vulcano *et al*, 2012). The drawback is that, if large and correlated data are involved, the maximum likelihood step is difficult to implement (Xu, 1997; Naim and Gildea, 2012).

Projection detruncation

This method is similar to the EM method, but uses the median instead of the mean. Furthermore, a weighting constant may be used to yield aggressive demand estimates. Recently, this method has been applied to RM systems (see Hopperstad, 1995; Skwarek, 1996a; Skwarek, 1996b; Zickus, 1998; Gorin, 2000; Zeni, 2001; Chen and Luo, 2005; Guo *et al*, 2011; Section 2 in Queenan *et al*, 2009; Weatherford and Ratliff, 2010).

Double exponential smoothing

In DES, one predicts the total demand that would have been registered in the absence of booking limits. The first parameter is used for smoothing the *base component* of the demand pattern and the second deals with the *trend component* (Section 3 in Queenan *et al*, 2009). For each instance of censorship, a non-linear optimization model estimates the two smoothing parameters while minimizing the forecasting error. This is achieved in the following manner. Let t be an instant when the booking limit of a given product has not been reached yet, that is, registered demand matches observed demand up to t. On the basis of Queenan *et al* (2009), let Y_t be the actual cumulative demand at time t, B_t the smoothed base component, T_t the smoothed trend component and FT_t the cumulative forecast at time t, taking trend into account. The forecast for the upcoming time period $t+1$ then satisfies

$$\mathrm{FT}_t = B_t + T_t \tag{13}$$

where

$$B_t = \mathrm{FT}_{t+1} + \delta(Y_{t+1} - \mathrm{FT}_{t+1}) \tag{14}$$

$$T_t = T_{t+1} + \beta(B_t - \mathrm{FT}_{t+1}) \tag{15}$$

and the parameters δ and β are optimal solutions of the mathematical programme

$$\min_{\delta,\beta} \sum_t (Y_t - \mathrm{FT}_t)^2. \tag{16}$$

The procedure is initialized on historical data, and the non-convex least-square problem can be solved via metaheuristics such as Tabu Search or Simulated Annealing. This framework has been applied to many demand uncensoring problems, and proved competitive with EM in most cases (Armstrong, 2001; Guo *et al*, 2008).

Non-linear programming

This class includes non-parametric methods such as *least-squares*, *discriminant analysis* or *cluster analysis*. In the literature, Besbes and Zeevi (2006) have discussed a non-parametric algorithm that characterizes the underlying demand behaviour. Farias *et al* (2013) have considered non-parametric methods in the context of choice modelling with limited data. Lee *et al* (2005) have used discriminant and cluster analysis to segment the customer population with respect to its preferences, with an application to the Taiwan Railway Administration.

Our classification of optimization is presented as follows:

$$\alpha_4 = \begin{cases} EM & \text{Expectation} - \text{Maximization} \\ PD & \text{Projection Detruncation} \\ DES & \text{Double Exponential Smoothing} \\ NLP & \text{Nonlinear Programming.} \end{cases}$$

Figure 3 illustrates methods to uncensor demand.

A TAXONOMY

In this section, the various techniques used in the RM literature to deal with missing data are classified with respect to their applicability and capability to deal with issues such as independence, stationarity or seasonality. Throughout, each work is made to fit our tuple notation.

Table 2 focuses on non-choice-based unconstraining methods. These ignore correlations between products, as well as the impact of availabilities on the demand for competing products, or competition from other firms. Within this class, EM and PD assume that demand distributions are known *a priori*, which is not the case in many practical situations.

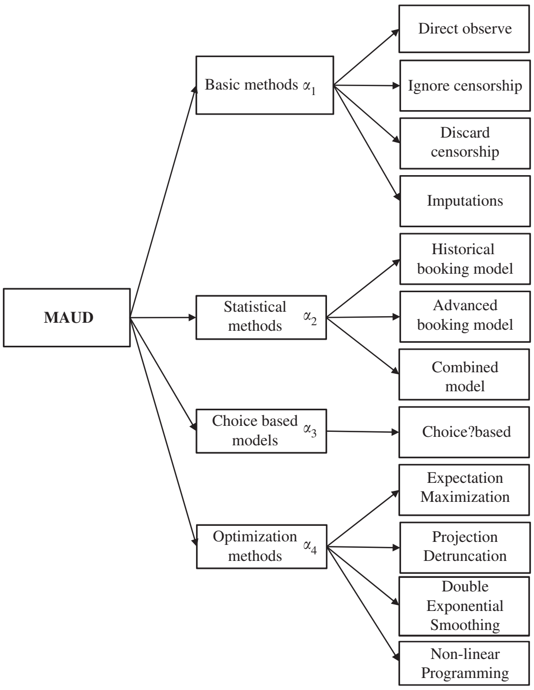

Figure 3: Methods applied to uncensor demand.

Table 3 displays models where customers base their purchasing decisions upon product availabilities (choice set). Apart from a suitable representation of customer behaviour, the flexibility of these models may improve the accuracy of both the estimation and optimization processes. In this respect, two key issues that should be better addressed in the future are:

1. inter-product and inter-temporal correlations ($\delta_1 = dep$)
2. seasonal factors ($\delta_3 = season$).

When firms have the opportunity to set prices dynamically, it is natural to expect price variations to persist. In this environment, customers may react strategically to price fluctuations, and ignoring such responses may lead to sub-optimal pricing decisions. Note also that product diversity induces a significant correlation between demands for alternatives within the choice sets. It follows that understanding of

Table 2: $\alpha3$ − Non-choice-based methods

Author	*Model description*
α_1 Basic methods	
Queenan *et al* (2009)	$[\mu_1 = myop, \mu_2 = hotel \vert \delta_1 = ind, \delta_4 = in \vert \alpha_1 \star]$
Zeni (2001)	$[\mu_1 = myop, \mu_2 = air \vert \delta_1 = ind, \delta_4 = in \vert \alpha_1 = imp]$
Orkin (1998)	$[\mu_1 = myop, \mu_2 = hotel \vert \delta_1 = ind \vert \alpha_1 = dir\_obs]$
Saleh (1997)	$[\mu_1 = myop, \mu_2 = air \vert \delta_1 = ind \vert \alpha_1 = ign\_cen, dis\_cen, imp]$
α_2: Statistical methods	
Sa (1987)	$[\mu_1 = myop, \mu_2 = air \vert \delta_1 = ind, \delta_3 = season \vert \alpha_2 = hbm(\text{lin\_reg, tseries})]$
Pölt (2000) – Weatherford (2000)	$[\mu_1 = myop, \mu_2 = air \vert \delta_1 = ind \vert \alpha_2 = abm(\text{BP})]$
Lee (1990)-Wickham (1995)	$[\mu_1 = myop, \mu_2 = air \vert \delta_1 = ind \vert \alpha_2 = abm]$
Zeni (2001)	$[\mu_1 = myop, \mu_2 = air \vert \delta_1 = ind, \delta_4 = in \vert \alpha_2 = abm(\text{BP})]$
van Ryzin and McGill (2000)	$[\mu_1 = myop, \mu_2 = air \vert \delta_1 = ind \vert \alpha_2 = cm(\text{LT})]$
Weatherford *et al* (2003)	$[\mu_1 = myop, \mu_2 = air \vert \delta_1 = ind \vert \alpha_2 = cm(\text{NN})]$
Dantas *et al* (2000)	$[\mu_1 = myop, \mu_2 = rail \vert \alpha_2 = cm(\text{NN})]$
ZF Li and Hoon Oum (2000)	$[\mu_1 = myop, \mu_2 = air \vert \delta_1 = ind, \delta_2 = spill, \delta_4 = in \vert \alpha_2 = cm(dist\_dem)]$
Liu *et al* (2002)	$[\mu_1 = myop, \mu_2 = hotel \vert \delta_1 = ind, \delta_3 = season \vert \alpha_2 = cm(param\_reg)]$
Belobaba and Farkas (1999)	$[\mu_1 = myop, \mu_2 = air \vert \delta_1 = ind, \delta_2 = spill, \delta_3 = season, \delta_4 = in \vert \alpha_2 = cm(dist\_dem)]$
Mishra and Viswanathan (2003)	$[\mu_1 = myop, \mu_2 = air, \mu_3 = compl \vert \delta_1 = ind, \delta_3 = season, \delta_4 = in \vert \alpha_2 = abm(\text{BP})]$
Huh *et al* (2011)	$[\mu_1 = myop, \mu_2 = ret \vert \delta_1 = ind \vert \alpha_2 = cm(dist\_dem)]$
α_4: Optimization methods	
Queenan *et al* (2009)	$[\mu_1 = myop, \mu_2 = hotel \vert \delta_1 = ind, \delta_4 = in \vert \alpha_4 = \text{DES}]$
Pölt (2000)-Weatherford (2000)	$[\mu_1 = myop, \mu_2 = air, \mu_3 = compl \vert \delta_1 = ind \vert \alpha_4 = \text{EM}]$
McGill (1995)	$[\mu_1 = myop, \mu_2 = air \vert \delta_1 = dep, \delta_3 = season \vert \alpha_4 = \text{EM}]$

Note: ⋆ address all approaches (Basic methods) presented in α_1.

Table 3: $\alpha 3$ – Choice-based methods

Author	*Model description*
α_2: Statistical methods	
Skwarek (1996a)-Hopperstad (1996)-Gorin (2000)-Zickus (1998)	$[\mu_1 = myop,\ \mu_2 = air,\ \mu_3 = compl \vert \delta_1 = ind,\ \delta_2 = spill,\ \delta_4 = in \vert \alpha_2 = abm(\mathrm{BP})]$
Ja *et al* (2001)	$[\mu_1 = myop,\ \mu_2 = air,\ \mu_3 = compl \vert \delta_1 = ind,\ \delta_2 = spill,\ \delta_4 = in \vert \alpha_2 = cm(lin\_reg)]$
Guo *et al* (2011)-Swan (2002)	$[\mu_1 = myop,\ \mu_2 = air \vert \delta_1 = ind,\ \delta_2 = spill \vert \alpha_2 = cm(dist\_dem)]$
Zhang and Cooper (2009)	$[\mu_1 = myop,\ \mu_2 = air \vert \delta_1 = ind \vert \alpha_2 = cm(dist\_dem)]$
Kunnumkal and Topaloglu (2010)	$[\mu_1 = myop,\ \mu_2 = air \vert \delta_1 = ind,\ \delta_4 = in \vert \alpha_2 = hbm]$
Cachon and Swinney (2009)	$[\mu_1 = strat,\ \mu_2 = ret \vert \delta_1 = ind,\ \delta_3 = season,\ \delta_4 = in \vert \alpha_2 = cm(dist\_dem\_Gamma)]$
α_4: Optimization methods	
Skwarek (1996b)-Hopperstad (1997)-Zickus (1998)-Gorin (2000)	$[\mu_1 = myop,\ \mu_2 = air,\ \mu_3 = compl \vert \delta_1 = ind,\ \delta_2 = spill,\ \delta_4 = in \vert \alpha_4 = \mathrm{PD}]$
Karmarkar *et al* (2010)-Vulcano *et al* (2012)	$[\mu_1 = myop,\ \mu_2 = air \vert \delta_1 = dep,\ \delta_3 = season \vert \alpha_4 = \mathrm{EM}]$
Stefanescu *et al* (2004)	$[\mu_1 = myop,\ \mu_2 = air \vert \delta_1 = dep \vert \alpha_4 = \mathrm{EM}]$
Talluri and van Ryzin (2004)	$[\mu_1 = myop,\ \mu_2 = air \vert \delta_1 = ind,\ \delta_4 = in \vert \alpha_4 = \mathrm{EM}]$
Dempster *et al* (1977)	$[\delta_1 = ind \vert \alpha_4 = \mathrm{EM}]$
Conlon and Mortimer (2008)	$[\mu_1 = myop,\ \mu_2 = ret \vert \delta_1 = ind,\ \delta_3 = season,\ \delta_4 = in \vert \alpha_4 = \mathrm{EM}]$
Bansal (2012)	$[\mu_1 = strat,\ \mu_2 = ret \vert \delta_1 = ind,\ \delta_3 = season,\ \delta_4 = in \vert \alpha_4 = \mathrm{NLP(DP)}]$

customer response to market mechanisms is an issue that should be addressed properly. In the same vein, the analysis of optimal purchase timing (inter-temporal substitution) is being monitored more closely in both the industrial and academical worlds.

Other relevant issues include capacity rationing (creation of artificial scarcity to influence purchase timing), valuation uncertainty and consumer learning effects. These relate the dynamics of consumer demand to the seller's dynamic pricing strategies, a dependency that is not captured by conventional models based on exogenous arrival processes.

With respect to seasonality, statistical methods hold some promise, under simple assumptions. However, the nature of seasonality can be complex as well as highly context-dependent. It may involve dimensions such as day of week, moth of year, holidays, with variations across the customer population. Embedding all this information within an estimation-optimization framework, and setting the right widths of booking intervals, are challenges that have not been properly addressed yet.

CONCLUSION

Currently, demand forecasting may well be the most critical area in RM, and demand unconstraining clearly lies at the heart of the matter.

The selection of a flawed demand model can have an adverse effect from the point of view of RM, through bad pricing and inventory management decisions. In the airline industry, seat availabilities for high-fare classes that are based on uncensored historical data may underestimate the required protection levels and yield suboptimal revenues or, worse, dynamically induce a downward spiral in both high-fare protections levels and revenue, as has been recognized by Kuhlmann (2004), Cooper *et al* (2006) and Eijken (2009).

Alternatively, ignoring the fact that data is censored may also result in protection levels that are too high, and thus reduce revenue by denying access to low-fare classes towards the end of the booking horizon.

Finally, we stress that the choice of an uncensoring method is context dependent. Indeed, while an airline must adjust dynamically its booking policy in *quasi*-real time, within an environment where partial information about the competition is available through the use of shared databases, this is not the case in the rail or retail industry. In the latter case, prominent issues are seasonality, just in time, the management and the management of the supply chain, all elements intervening in a different way and on a different time scale. We hope that the present article has shed some light on the latest developments in this area, through a novel taxonomy, and will trigger further research in this key area.

NOTE

1 Products are assumed to be independent.

REFERENCES

Andersson, S. (1998) Passenger choice analysis for seat capacity control: A pilot project in Scandinavian Airlines. *International Transactions in Operational Research* 5(6): 471–486.

Armstrong, A. and Meissner, J. (2010) Railway Revenue Management: Overview and Models. UK: Lancaster University Management School. Technical Report .

Armstrong, J. (2001) *Principles of Forecasting: A Handbook for Researchers and Practitioners*. April 2001 edn. New York, USA: Springer.

Bansal, M. (2012) Modeling customer behavior for revenue management. PhD thesis, Columbia University.

Belobaba, P. (1987) Air Travel Demand and Airline Seat Inventory Management. Cambridge, MA: Flight Transportation Laboratory, Massachusetts Institute of Technology. Technical Report.

Belobaba, P. and Farkas, A. (1999) Yield management impacts on airline spill estimation. *Transportation Science* 33(2): 217–232.

Belobaba, P. and Hopperstad, C. (1999) Boeing/MIT simulation study: PODS results update. In: AGIFORS Reservations and Yield Management Study Group Symposium Proceedings. Renton, USA.

Besbes, O. and Zeevi, A. (2006) Blind nonparametric revenue management: Asymptotic optimality of a joint learning and pricing method. Working paper.

Bodea, T. (2008) Choice-based revenue management: A hotel perspective. PhD thesis. Georgia Institute of Technology.

Box, G., Jenkins, G. and Reinsel, G. (2011) *Time Series Analysis: Forecasting and Control*. 4th edn. Vol. 734. Hoboken, New Jersey: Wiley.

Brummer, M. *et al.* (1988) Determination of potential load capacity based on observed load factor data, a study for Northwest Airlines. Northfield, MN: St. Olaf College Undergraduate Practicum Group. Technical Report.

Cachon, G. and Feldman, P. (2011) Dynamic versus static pricing in the presence of strategic consumers. MSOM. In press, accepted manuscript.

Cachon, G. and Swinney, R. (2009) Purchasing, pricing, and quick response in the presence of strategic consumers. *Management Science* 55(3): 497–511.

Cachon, G. and Swinney, R. (2011) The value of fast fashion: Quick response, enhanced design, and strategic consumer behavior. *Management Science* 57(4): 778–795.

Chen, Y. and Luo, L. (2005) Approach of uncensored demand in airline revenue. *Journal of Chengdu University of Information Technology* 20(6): 747–750.

Conlon, C. and Mortimer, J. (2008) Demand Estimation under Incomplete Product Availability. National Bureau of Economic Research. Technical Report.

Cooper, W., Homem-de Mello, T. and Kleywegt, A. (2006) Models of the spiral-down effect in revenue management. *Operations Research* 54(5): 968–987.

Crevier, B., Cordeau, J. and Savard, G. (2012) Integrated operations planning and revenue management for rail freight transportation. *Transportation Research Part B: Methodological* 46(1): 100–119.

Dantas, A., Yamamoto, K., Lamar, M. and Yamashita, Y. (2000) Neural network for travel demand forecast using GIS and remote sensing. In: INNS-ENNS International Joint Conference on Neural Networks. Japan: IEEE.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977) Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society Series B: Methodological* 39(1): 1–38.

Eijken, T. (2009) The spiral-down effect in revenue management. Master's thesis. BMI, Vrije Universiteit Amsterdam.

Eren, S. and Maglaras, C. (2009) A Maximum Entropy Joint Demand Estimation and Capacity Control Policy. Columbia Business School, Center for Pricing and Revenue Management. Technical Report.

Farias, V. (2007) Revenue management beyond 'estimate, then optimize'. PhD thesis, Stanford University.

Farias, V., Jagabathula, S. and Shah, D. (2013) A nonparametric approach to modeling choice with limited data. *Management Science* 59(2): 305–322.

Ferguson, M. and Queenan, C. (2009) Starting with good inputs: Unconstraining demand data in revenue management. *INFORMS Transactions on Education* 9(3): 180–187.

Gallego, G. and Hu, M. (2014) Dynamic pricing of perishable assets under competition. *Management Science*.

Gorin, T. (2000) Airline revenue management: Sell-up and forecasting algorithms. PhD thesis, Massachusetts Institute of Technology.

Guo, P., Xiao, B. and Li, J. (2008) Research on the method of demand forecasting without constraint for airline passenger revenue management. *Journal of Transportation Engineering and Information* 2(6): 71–78.

Guo, P., Xiao, B. and Li, J. (2011) The research on unconstraining methods based on multi-distribution of demand in revenue management systems. *Applications of Statistics and Management* 30(6): 1077–1088.

Gurbuz, E. *et al.* (2011) Revenue management operations in hotel chains in Finland. Master's thesis. Imatra, Finland: Saimaa University of Applied Sciences.

Haensel, A. and Koole, G. (2010) Estimating unconstrained demand rate functions using customer choice sets. *Journal of Revenue and Pricing Management* 10(5): 438–454.

Haensel, A., Koole, G. and Erdman, J. (2011) Estimating unconstrained customer choice set demand: A case study on airline reservation data. *Journal of Choice Modelling* 4(3): 75–87.

He, D. and Luo, L. (2006) An improved winters model for airline demand forecast. *Journal of Transportation Systems Engineering and Information Technology* 6(6): 103–107.

Hopperstad, C. (1995) An Alternative Detruncation Method. Boeing Commercial Aircraft Company. Technical Report.

Hopperstad, C. (1996) Passenger O/D simulator–PODS. In: AGIFORS Proceedings, Zurich, Switzerland.

Hopperstad, C. (1997) PODS update: Simulation of O/D revenue management schemes. In: AGIFORS Proceedings Yield Management Study Group. Renton, USA.

Hopperstad, C., Zerbib, G. and Belobaba, P. (2007) Methods for estimating sell-up. In: AGIFORS Yield Management Study Group Meeting. Jeju Island, Korea.

Huh, W., Levi, R., Rusmevichientong, P. and Orlin, J. (2011) Adaptive data-driven inventory control with censored demand based on Kaplan-Meier estimator. *Operations Research* 59(4): 929–941.

Hyndman, R., Koehler, A., Ord, J. and Snyder, R. (2008) *Forecasting with Exponential Smoothing: The State Space Approach*. Berlin, Germany: Springer.

Ja, S., Rao, B. and Chandler, S. (2001) Passenger recapture estimation in airline revenue management. In: AGIFORS 41st Annual Symposium. Sydney, Australia.

Jiang, H. (2007) Network Capacity Management Competition. Judge Business School at University of Cambridge, UK. Technical Report.

Jiang, H. and Pang, Z. (2011) Network capacity management under competition. *Computational Optimization and Applications* 50(2): 287–326.

Kachitvichyanukul, V., Luong, H. and Pitakaso, R. (2012) Hotel room demand forecasting via observed reservation information. Asia Pacific Industrial Engineering and Management Systems Conference 2012.

Kaplan, E. and Meier, P. (1958) Nonparametric estimation from incomplete observations. *Journal of the American statistical association* 53(282): 457–481.

Karmarkar, S., Goutam, D. and Tathagata, B. (2010) Revenue impacts of demand unconstraining and accounting for dependency. *Journal of Revenue and Pricing Management* 10(4): 367–381.

Kuhlmann, R. (2004) Why is revenue management not working? *Journal of Revenue and Pricing Management* 2(4): 378.

Kunnumkal, S. and Topaloglu, H. (2010) A randomized linear program for the network revenue management problem with customer choice behavior. *Journal of Revenue and Pricing Management* 10(5): 455–470.

Kwon, C., Friesz, T., Mookherjee, R., Yao, T. and Feng, B. (2009) Non-cooperative competition among revenue maximizing service providers with demand learning. *European Journal of Operational Research* 197(3): 981–996.

Lee, A. (1990) Airline Reservations Forecasting: Probabilistic and Statistical Models of the Booking Process. Flight Transportation Laboratory, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology. Technical Report.

Lee, C., Lin, T. and Lin, C. (2005) Pattern analysis on the booking curve of an inter-city railway. *Journal of the Eastern Asia Society for Transportation Studies* 6: 303–317.

Levin, Y., McGill, J. and Nediak, M. (2010) Optimal dynamic pricing of perishable items by a monopolist facing strategic consumers. *Production and Operations Management* 19(1): 40–60.

L'Heureux, E. (1986) A new twist in forecasting short-term passenger pickup. In: AGIFORS Annual Meeting. Northwestern University, USA.

Little, R. and Rubin, D. (2002) *Statistical Analysis with Missing Data*. 2nd edn. New York: Wiley.

Littlewood, K. (2005) Special issue papers: Forecasting and control of passenger bookings. *Journal of Revenue and Pricing Management* 4(2): 111–123.

Liu, P. (2004) Hotel demand/cancellation analysis and estimation of unconstrained demand using statistical methods. In: Revenue Management and Pricing: Case Studies and Applications. London Thompson.

Liu, P., Smith, S., Orkin, E. and Carey, G. (2002) Estimating unconstrained hotel demand based on censored booking data. *Journal of Revenue and Pricing Management* 1(2): 121–138.

Liu, Q. and van Ryzin, G. (2008) On the choice-based linear programming model for network revenue management. *Manufacturing & Service Operations Management* 10(2): 288–310.

Liu, Q. and Zhang, D. (2011) Dynamic Pricing Competition with Strategic Customers under Vertical Product Differentiation. Hong Kong University of Science and Technology, Kowloon. Technical Report.

Martínez, V. and Talluri, K. (2011) Dynamic price competition with fixed capacities. *Management Science* 57(6): 1078–1093.

McFadden, D. (2001) Disaggregate behavioral travel demand's RUM side. In: *Travel Behaviour Research*. Oxford: Pergamon Press, UC Berkeley, USA.

McGill, J. (1995) Censored regression analysis of multiclass passenger demand data subject to joint capacity constraints. *Annals of Operations Research* 60(1): 209–240.

McGill, J. and van Ryzin, G. (1999) Revenue management: Research overview and prospects. *Transportation science* 33(2): 233–256.

Meissner, J. and Strauss, A. (2009) Choice-Based Network Revenue Management under Weak Market Segmentation. Lancaster University Management School. Technical Report.

Meissner, J. and Strauss, A. (2012a) Improved bid prices for choice-based network revenue management. *European Journal of Operational Research* 217(2): 417–427.

Meissner, J. and Strauss, A. (2012b) Network revenue management with inventory-sensitive bid prices and customer choice. *European Journal of Operational Research* 216(2): 459–468.

Meissner, J., Strauss, A. and Talluri, K. (2012) An Enhanced Concave Program Relaxation for Choice Network Revenue Management. Lancaster University Management School. Technical Report.

Mishra, S. and Viswanathan, V. (2003) Revenue management with restriction-free pricing. In: AGIFORS Revenue Management and Distribution Study Group Meeting, Honolulu, USA.

Naim, I. and Gildea, D. (2012) Convergence of the em algorithm for gaussian mixtures with unbalanced mixing coefficients. *ArXiv preprint ArXiv 1206: 6427.*

Netessine, S. and Shumsky, R. (2005) Revenue management games: Horizontal and vertical competition. *Management Science* 51(5): 813–831.

Orkin, E. (1998) Wishful thinking and rocket science the essential matter of calculating unconstrained demand for revenue management. *Cornell Hotel and Restaurant Administration Quarterly* 39(4): 15–19.

Perakis, G. and Sood, A. (2006) Competitive multi-period pricing for perishable products: A robust optimization approach. *Mathematical Programming* 107(1): 295–335.

Pölt, S. (2000) *From Bookings to Demand: The Process of Unconstraining*. New York: AGIFORS Reservations and Yield Management Study Group.

Popescu, A., Barnes, E., Johnson, E. and Keskinocak, P. (2013) Bid prices when demand is a mix of individual and batch bookings. *Transportation Science* 47(2): 198–213.

Queenan, C., Ferguson, M., Higbie, J. and Kapoor, R. (2009) A comparison of unconstraining methods to improve revenue management systems. *Production and Operations Management* 16(6): 729–746.

Ratliff, R., Rao, B., Narayan, C. and Yellepeddi, K. (2008) A multi-flight recapture heuristic for estimating unconstrained demand from airline bookings. *Journal of Revenue and Pricing Management* 7(2): 153–171.

Sa, J. (1987) Reservations Forecasting in Airline Yield Management. Cambridge, MA: Flight Transportation Laboratory, Massachusetts Institute of Technology. Technical Report.

Salch, J. (1997) Unconstraining passenger demand using the em algorithm. In: Proceedings of the INFORMS Conference. Dallas, USA.

Saleh, R. (1997) Estimating lost demand with imperfect availability indicators. In: Proceedings of the AGIFORS Reservations and Yield Management Study Group. Montreal, Canada.

Sharif Azadeh, S., Labib, R. and Savard, G. (2012) Railway demand forecasting in revenue management using neural networks. *International Journal of Revenue Management* 7(1): 18–36.

Shen, Z. and 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.

Skwarek, D. (1996a) Competitive Impacts of Yield Management System Components: Forecasting and Sell-Up Models. Cambridge, MA: Massachusetts Institute of Technology, Flight Transportation Laboratory. Technical Report.

Skwarek, D. (1996b) Revenue and traffic impacts of alternative detruncation methods. In: Proceedings of the AGIFORS Reservations and Yield Management Study Group. Zurich, Switzerland.

Stefanescu, C. (2009) Multivariate Customer Demand: Modeling and Estimation from Censored Sales. London Business School. Technical Report.

Stefanescu, C., DeMiguel, V., Fridgeirsdottir, K. and Zenios, S. (2004) Revenue management with correlated demand forecasting. In: Proceedings of the American Statistical Association, Business and Economics Statistics Section. Alexandria, VA.

Su, X. (2007) Intertemporal pricing with strategic customer behavior. *Management Science* 53(5): 726–741.

Swan, M. (1999) Spill modeling for airlines. In: 8th World Conference on Transport Research Proceedings, pp. 225–237.

Swan, W. (1979) A Systems Analysis of Scheduled Air Transportation Networks. Massachusetts Institute of Technology, Flight Transportation Laboratory. Technical Report.

Swan, W. (2002) Airline demand distributions: Passenger revenue management and spill. *Transportation Research Part E: Logistics and Transportation Review* 38(3): 253–263.

Swinney, R. (2011) Selling to strategic consumers when product value is uncertain: The value of matching supply and demand. *Management Science* 57(10): 1737–1751.

Talluri, K. (2009) A Finite-Population Revenue Management Model and A Risk-Ratio Procedure for The Joint Estimation of Population Size and Parameters. Universitat Pompeu Fabra. Technical Report.

Talluri, K. (2010) A Randomized Concave Programming Method for Choice Network Revenue Management. Universitat Pompeu Fabra. Technical Report.

Talluri, K. and van Ryzin, G. (2004) Revenue management under a general discrete choice model of consumer behavior. *Management Science* 50(1): 15–33.

Talluri, K. and van Ryzin, G. (2005) *The Theory and Practice of Revenue Management*, Vol. 68, New York: Springer.

Train, K. (2009) *Discrete Choice Methods with Simulation*. 2nd edn. Cambridge, UK: Cambridge University Press.

van Ryzin, G. and McGill, J. (2000) Revenue management without forecasting or optimization: An adaptive algorithm for determining airline seat protection levels. *Management Science* 46(6): 760–775.

van Ryzin, G. (2005) Future of revenue management: Models of demand. *Journal of Revenue and Pricing Management* 4(2): 204–210.

Vulcano, G., van Ryzin, G. and Chaar, W. (2010) Om practice-choice-based revenue management: An empirical study of estimation and optimization. *Manufacturing & Service Operations Management* 12(3): 371–392.

Vulcano, G., van Ryzin, G. and Ratliff, R. (2012) Estimating primary demand for substitutable products from sales transaction data. *Operations Research* 60(2): 313–334.

Weatherford, L. (2000) Unconstraining methods. In: AGIFORS Reservations and Yield Management Study Group. New York.

Weatherford, L. and Belobaba, P. (2002) Revenue impacts of fare input and demand forecast accuracy in airline yield management. *Journal of the Operational Research Society* 53(8): 811–821.

Weatherford, L., Gentry, T. and Wilamowski, B. (2003) Neural network forecasting for airlines: A comparative analysis. *Journal of Revenue and Pricing Management* 1(4): 319–331.

Weatherford, L. and Kimes, S. (2003) A comparison of forecasting methods for hotel revenue management. *International Journal of Forecasting* 19(3): 401–415.

Weatherford, L. and Polt, S. (2002) Better unconstraining of airline demand data in revenue management systems for improved forecast accuracy and greater revenues. *Journal of Revenue and Pricing Management* 1(3): 234–254.

Weatherford, L. and Ratliff, R. (2010) Review of revenue management methods with dependent demands. *Journal of Revenue and Pricing Management* 9(4): 326–340.

Wickham, R. (1995) Evaluation of Forecasting Techniques for Short-Term Demand of Air Transportation. Massachusetts Institute of Technology, Flight Transportation Laboratory. Technical Report.

Xu, L. (1997) Comparative analysis on convergence rates of the em algorithm and its two modifications for gaussian mixtures. *Neural Processing Letters* 6(3): 69–76.

Yang, H., Song, H. and Zhang, S. (2010) Dynamic pricing with strategic and myopic consumers. In: Business Intelligence and Financial Engineering (BIFE), Hong Kong: IEEE.

Yin, R., Aviv, Y., Pazgal, A. and Tang, C. (2009) Optimal markdown pricing: Implications of inventory display formats in the presence of strategic customers. *Management Science* 55(8): 1391–1408.

Zakhary, A., El Gayar, N. and Atiya, A.F. (2008) A comparative study of the pickup method and its variations using a simulated hotel reservation data. *ICGST International Journal on Artificial Intelligence and Machine Learning*, Special Issue on Computational Methods for the Tourism Industry 8: 15–21.

Zeni, R. (2001) Improved forecast accuracy in airline revenue management by unconstraining demand estimates from censored data. PhD thesis, MIT.

Zeni, R. and Lawrence, K. (2004) Unconstraining demand data at us airways. In: Revenue Management and Pricing: Case Studies and Applications. Chapter 11, London, UK: Thomson, Cengage Learning Business Press, pp. 124–136.

ZF Li, M. and Hoon Oum, T. (2000) Airline spill analysis–beyond the normal demand. *European Journal of Operational Research* 125(1): 205–215.

Zhang, D. and Cooper, W. (2005) Revenue management for parallel flights with customer-choice behavior. *Operations Research* 53(3): 415–431.

Zhang, D. and Cooper, W. (2009) Pricing substitutable flights in airline revenue management. *European Journal of Operational Research* 197(3): 848–861.

Zhu, J. (2006) Using turndowns to estimate the latent demand in a car rental unconstrained demand forecast. *Journal of Revenue and Pricing Management* 4(4): 344–353.

Zickus, J. (1998) Forecasting for Airline Network Revenue Management: Revenue and Competitive Impacts. Massachusetts Institute of Technology, Transportation Laboratory. Technical Report.