1. Introduction

Revenue management (RM) is a fast growing branch in operations research (OR) and has been credited for 3–7% revenue improvement in the airline, hotel, and car rental industries [1]. It was developed in the late 1970s after the deregulation of the US airline industry. There are various customized definitions of RM. For example, Cross [1] defines RM as an application of disciplined tactics that predicts consumer behavior at the micromarket level and optimize product availability and price to maximize revenues. In Smith et al. [2], it is called the application of information systems and pricing strategies to allocate the right capacity to the right customer at the right place and the right time.

In many service industries, capacity of supply is often fixed while demand is volatile. Hence, it is challenging for service companies to achieve a balance between supply and demand. To optimize revenues, RM models need to project demand first based on historical data. It then manages supply and demand through pricing, inventory
control, and overbooking. Obviously, reliable forecasting is essential to the success of the revenue management system (RMS). Erroneous demand forecast may seriously impede the performance of RMS. Lee [3] has shown that a small improvement of 10% in forecasting accuracy contributes to 0.5–3% increase in expected revenues.

Forecasting, however, has not advanced as much as other RM components which ultimately depend on accurate forecasting [4]. A major complicated issue for forecasting in RM is booking data censoring. Because demand recorded is usually affected by managerial decisions, and thus, not genuine, successful forecasting in RM is intricate. It may be systematically biased and lead to further incorrect price and capacity allocation decisions. For example, in the airline and hotel industries, booking limits are set to protect certain customer classes. When the limits are reached, the respective classes are closed and as a result, further demand information for these classes is lost. In statistics this is called censored or constrained data.

With censored data, it is likely to overestimate demand in some situations and underestimate in others. It has been reported that up to 3% of potential revenue may be lost if the forecast used by an RMS has a negative bias [5]. And the impact of underestimating demand by 12.5–25% can hurt revenues by 1–3% on high-demand flights [6]. Moreover, a spiral-down effect on total revenue will occur if historical booking data are left constrained, and true demand is underestimated. This implies that the firm’s expected revenue decreases monotonically over time [7].

To overcome these problems, it is necessary to extrapolate the true demand distribution parameters from censored booking data before putting them into the forecasting models. In the airline and hotel industries, this process is called demand unconstraining. Other terms, such as detruncating, spill analysis, and censored data analysis, have also been used to describe this process. In general, the data estimated by unconstraining methods are referred to unconstrained data.

Wickham [8] claims that unconstrained demand is not easy to measure. Although it is considered to be the “Holy Grail” of RM forecasting, many researchers have found that unconstrained data provides better forecasts and improve revenues. For example, Skwarek [9] has tested the RM systems of two airlines. One of them uses unconstrained data, but the other does not. The results show that even if the booking rate is low, the impact of unconstrained data on revenue can reach 3.5%. In addition, Weatherford and Pölt [10] report that, with actual booking data from a major US airline, the unconstraining process results in 2–12% of the revenue gains.

In view of the revenue benefit, one can see that demand unconstraining deserves attention from both researchers and practitioners. Forecasting based on unconstrained data can overcome the limitation of truncated demand due to the booking limits, better reflect true demand, and improve the accuracy of forecasting. In addition, demand unconstraining is also helpful to the allocation of fleet capacity.

General reviews of the RM literature can be found in [11–17]. As these studies show, despite the importance of demand unconstraining, the research front has received less attention compared to the work on other RM components, and unconstraining methods adopted by RM vendors are largely ineffective [18]. Although some research papers, such as [4, 10] and [9–23], partially review the unconstraining methods, their primary intent is not to provide comprehensive surveys on details of the technique.

Over the last decade, there has been extensive research on data unconstraining. Despite significant development in the area, more research seems needed compared to the advancement in other components of RM. The objective of this paper, thus, is to review
Table 1: Review point to days prior to mapping and booking matrix without demand censoring of a fare class.

<table>
<thead>
<tr>
<th>Review point</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days prior to departure</td>
<td>50</td>
<td>45</td>
<td>40</td>
<td>35</td>
<td>25</td>
<td>17</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>True demand</td>
<td>0</td>
<td>14</td>
<td>38</td>
<td>50</td>
<td>71</td>
<td>91</td>
<td>103</td>
<td>108</td>
<td>104</td>
<td>102</td>
</tr>
<tr>
<td>Booking limit</td>
<td>115</td>
<td>115</td>
<td>115</td>
<td>115</td>
<td>115</td>
<td>115</td>
<td>115</td>
<td>115</td>
<td>115</td>
<td>115</td>
</tr>
<tr>
<td>Observed bookings</td>
<td>0</td>
<td>14</td>
<td>38</td>
<td>50</td>
<td>71</td>
<td>91</td>
<td>103</td>
<td>108</td>
<td>104</td>
<td>102</td>
</tr>
<tr>
<td>Fare class status</td>
<td>Available</td>
<td>Available</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(no censorship)</td>
<td>(cancellation/no shows)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Review point to days prior to mapping and booking matrix with demand censoring of a fare class.

<table>
<thead>
<tr>
<th>Review point</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days prior to departure</td>
<td>50</td>
<td>45</td>
<td>40</td>
<td>35</td>
<td>25</td>
<td>17</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>True demand</td>
<td>0</td>
<td>14</td>
<td>38</td>
<td>50</td>
<td>71</td>
<td>91</td>
<td>103</td>
<td>108</td>
<td>104</td>
<td>102</td>
</tr>
<tr>
<td>Booking limit</td>
<td>85</td>
<td>85</td>
<td>85</td>
<td>85</td>
<td>85</td>
<td>85</td>
<td>85</td>
<td>85</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>Observed bookings</td>
<td>0</td>
<td>14</td>
<td>38</td>
<td>50</td>
<td>71</td>
<td>85</td>
<td>85</td>
<td>85</td>
<td>81</td>
<td>79</td>
</tr>
<tr>
<td>Fare class status</td>
<td>Available</td>
<td>Not available</td>
<td>Available</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(no censorship)</td>
<td>(censored)</td>
<td>(cancellation/no-shows)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The rest of this paper is organized as follows. Section 2 provides a thorough discussion of demand-unconstraining problem in RMS from two aspects: the definitions and the relationship between censored data unconstraining and forecasting. Section 3 reviews five alternative methods for data unconstraining when firms face censored sales data, as well as how these methods are applied to different industries. Section 4 lists some future research questions, followed by concluding remarks.

2. Demand Unconstraining

2.1. Definitions

When customers’ booking requests for a certain class are accepted, the recorded booking data show true demand (see Table 1 and Figure 1(a)). However, if the booking limit is reached and demand requests are denied, the historical data only represent censored demand at the view point (see Table 2 and Figure 1(b)). As proposed by Zeni [19], such a problem encountered naturally leads to following definitions.

2.1.1. Censored Observation

An observation is considered censored (or constrained) if the booking limit in a given fare class at a specified review point in the history of the service product is less than or equal to the number of bookings present at that time.


**Figure 1:** Booking pace curve of a fare class.

**Table 3:** Factors impacting on forecasting of RM.

<table>
<thead>
<tr>
<th>No.</th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Seasonality</td>
</tr>
<tr>
<td>2</td>
<td>Day-of-week and time-of-day variations</td>
</tr>
<tr>
<td>3</td>
<td>Special events</td>
</tr>
<tr>
<td>4</td>
<td>Sensitivity to pricing actions</td>
</tr>
<tr>
<td>5</td>
<td>Demand dependencies between fare classes</td>
</tr>
<tr>
<td>6</td>
<td>Group bookings</td>
</tr>
<tr>
<td>7</td>
<td>Cancellations</td>
</tr>
<tr>
<td>8</td>
<td><em>Censorship of historical demand data</em></td>
</tr>
<tr>
<td>9</td>
<td>Defections from delayed flights</td>
</tr>
<tr>
<td>10</td>
<td>No shows</td>
</tr>
<tr>
<td>11</td>
<td>Recapture</td>
</tr>
</tbody>
</table>

2.1.2. *Constrained Fare Class*

A fare class is considered constrained if the observed demand in a given fare class at any review point in the life of the service product is censored (or constrained).

In light of the fact that firms really have a record of the actual number of bookings, a challenge faced by them is to estimate how many true demands would have been accepted without any constraint for their products. This process has commonly been called *demand unconstraining*.

2.2. *Relationship to Forecasting*

In practice, forecasting consumes major resources of development, maintenance, and implementation time of an RMS [11]. Hence, most researchers are mainly interested in the comparison between the existing forecasting methodologies. During the early development stage of revenue management, there was lack of research on the theoretical side of demand forecasting in RM systems due to its complexity [4]. McGill and van Ryzin [12] list the factors contributing to these difficulties (see Table 3). Each of these factors presents a challenge of its
Table 4: Forecasting issues in RMS.

<table>
<thead>
<tr>
<th>No.</th>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>What to forecast</td>
</tr>
<tr>
<td>2</td>
<td>Level of aggregation</td>
</tr>
<tr>
<td>3</td>
<td>Unconstraining methods</td>
</tr>
<tr>
<td>4</td>
<td>Number of periods to include in forecast</td>
</tr>
<tr>
<td>5</td>
<td>Which data to use</td>
</tr>
<tr>
<td>6</td>
<td>Outliers</td>
</tr>
<tr>
<td>7</td>
<td>Reporting forecast accuracy</td>
</tr>
<tr>
<td>8</td>
<td>Measurement and impacts on revenue</td>
</tr>
</tbody>
</table>

In addition, depending on the type of industry, there are many issues, as shown in Table 4, associated with forecasting [24, 25]. Managers must address them before choosing an appropriate forecasting method.

As presented in [20], observed booking data need to be unconstrained before being used in RMS (see Figure 2). Demand unconstraining can fill the gap between what is needed (unconstrained data) and what can be observed (censored data). Its function is to provide true demand information for forecasting models. It usually contains two steps. First, through examining similar historical bookings that have not been censored, one derives unconstrained demand parameters. These parameters then are applied to estimate unconstrained historical demand. This process is viewed as a preforecasting step. Weatherford [21] describes the steps that a complete forecasting system should perform in RMS (see Figure 3). Among them, the choice of unconstraining methods and the unconstraining of censored observations are both essential in the forecasting process.

3. Unconstraining Methods

There are two reasons for obtaining unconstrained data using unconstraining methods. First, the number of forecasting models producing unbiased estimates from censored data is limited. Second, different units within a firm may use various forecasting techniques with no coordination. Although these forecasting models may deal explicitly with censored data, it would be preferable to unconstrain the data collectively and then have all the forecasting models that use the same unconstrained data [19].

Generally speaking, a firm facing censored sales data has five options: (1) directly observe and record latent demand, (2) leave data constrained, ignoring the fact of censorship, (3) use unconstrained data only and discard censored ones, (4) replace censored data using imputation methods, or (5) statistically unconstrain the data. These alternatives and related methods are illustrated in Table 5.

3.1. Direct Observation

Direct observations of demand include records of bookings (requests that are met) and rejections (requests that are not met). The method may not be able to uncover true latent demand. Booking data censorship may be caused by availability (denials) or rate (regrets). Bookings declined due to availability are considered latent demand [22]. The boundary between denials and regrets is blurry. Denials occur when customers’ requests cannot be met because of capacity constraints, while regrets happen when the requests can be
accommodated but customers refuse to book. In practice, it is often impossible to categorize a particular call as one or the other [26]. Therefore, it increases the difficulty in distinguishing denials from regrets.

Firms invest in systems and train their managers in order to track turndowns directly and depend on these direct observations to unconstrain their sales data. Queenan et al. [22] point out that there are a number of issues that need to be considered: (1) multiple availability inquiries from the same customer, (2) incorrect categorization of rejections by reservation agent, and (3) the fact that only small portions of customer requests arrive through a channel controlled by the firm [26]. Consequently, direct observations of demand are not an option for most industries because of these drawbacks.

3.2. Ignore the Censored Data

Ignoring censorship and performing estimates as if the censoring never happened, this approach is referred to as method Naïve #1 (N1) in [10]. It often leads to undesirable consequences such as underestimated future demand, a spiral-down effect on total revenue [7], and insufficient number of seats protected for high-fare customers. In the meantime, demand for low fare classes appears to decrease in RMS when it actually increases [19]. Unfortunately, this practice is common for firms using unsophisticated RMS [22]. The postdeparture analysis of how well an RMS performs, as well as its development stages, would certainly be influenced. The performance of forecasting, inventory control policies,
Table 5: Research on unconstraining methods used in RMS.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Model</th>
<th>Reference</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Observation</td>
<td>Directly Observe and Record Latent Demand</td>
<td>Orkin [26] and Queenan et al. [22]</td>
<td>1998</td>
</tr>
<tr>
<td>Ignore the censored data</td>
<td>Naive #1 (N1)</td>
<td></td>
<td>2007</td>
</tr>
<tr>
<td>Discard the censored data</td>
<td>Naive #2 (N2)</td>
<td>Saleh [27]</td>
<td>1997</td>
</tr>
<tr>
<td>Imputation unconstraining</td>
<td>Naive #3 (N3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Proposition</td>
<td>Spill Model</td>
<td>Swan [28–31]</td>
<td>1979–</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1990</td>
</tr>
<tr>
<td></td>
<td>Projection Detruncation (PD)</td>
<td>Hopperstad [33]</td>
<td>1995</td>
</tr>
<tr>
<td></td>
<td>Pickup Detruncation (Pickup)</td>
<td>Skwarek [34]</td>
<td>1996</td>
</tr>
<tr>
<td></td>
<td>Life Table (LT)</td>
<td>van Ryzin and McGill [36]</td>
<td>2000</td>
</tr>
<tr>
<td></td>
<td>Observed Load Factor (OLF)</td>
<td>Li and Oum [37]</td>
<td>2000</td>
</tr>
<tr>
<td></td>
<td>Programming</td>
<td>Gao and Zhu [38] and</td>
<td>2005</td>
</tr>
<tr>
<td></td>
<td>Multi-distribution-Based EM</td>
<td>Guo [40] and Guo et al. [41]</td>
<td>2008</td>
</tr>
<tr>
<td></td>
<td>and PD</td>
<td></td>
<td>2011</td>
</tr>
<tr>
<td>Single-class Airlines</td>
<td>Comparison</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BP, PD, and pickup</td>
<td>Skwarek [34]</td>
<td>1996</td>
</tr>
<tr>
<td></td>
<td>N2, N3, BP, and PD</td>
<td>Skwarek [9] and</td>
<td>1996</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hopperstad [42, 43]</td>
<td>1997</td>
</tr>
<tr>
<td></td>
<td>N1, N2, N3, BP, and EM</td>
<td>Pölt [20], Weatherford [21]</td>
<td>2000</td>
</tr>
<tr>
<td></td>
<td>N1, N2, N3, BP, PD, and EM</td>
<td>Weatherford and Pölt [10],</td>
<td>2002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Zeni [19] and</td>
<td>2001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Zeni and Lawrence [44]</td>
<td>2004</td>
</tr>
<tr>
<td></td>
<td>EM and PD</td>
<td>Chen and Luo [45]</td>
<td>2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2000</td>
</tr>
<tr>
<td>Statistical Model</td>
<td>EM</td>
<td>He and Luo [48]</td>
<td>2006</td>
</tr>
<tr>
<td>Unconstraining</td>
<td>DES</td>
<td>Guo et al. [49]</td>
<td>2008</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liu et al. [50] and Liu [51]</td>
<td>2002</td>
</tr>
<tr>
<td></td>
<td>Parametric Regression (PR)</td>
<td></td>
<td>2004</td>
</tr>
<tr>
<td></td>
<td>Hotels</td>
<td>Double Exponential Smoothing (DES)</td>
<td>Queenan et al. [22]</td>
</tr>
</tbody>
</table>


Table 5: Continued.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Model</th>
<th>Reference</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Censored Demand EM</td>
<td>McGill [52]</td>
<td>1995</td>
<td></td>
</tr>
<tr>
<td>Procedure</td>
<td>Spill Model</td>
<td>Farkas [53] and Belobaba and Farkas [54]</td>
<td>1996</td>
</tr>
<tr>
<td></td>
<td>Q forecasting</td>
<td>Boyd and Kallesen [56]</td>
<td>2001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Boyd et al. [57] and Hopperstad et al. [58]</td>
<td>2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hopperstad [59]</td>
<td>2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Karmarkar et al. [60]</td>
<td>2011</td>
</tr>
<tr>
<td></td>
<td>Regression-Based Estimation</td>
<td>Ja et al. [61]</td>
<td>2001</td>
</tr>
<tr>
<td></td>
<td>Correlated Demand Forecasting</td>
<td>Stefanescu et al. [62]</td>
<td>2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stefanescu [63]</td>
<td>2009</td>
</tr>
<tr>
<td></td>
<td>EM (Discrete Choice Model)</td>
<td>Talluri and van Ryzin [64] and Vulcano et al. [65]</td>
<td>2004</td>
</tr>
<tr>
<td></td>
<td>Multi-flight Recapture Heuristic</td>
<td>Ratliff et al. [23]</td>
<td>2008</td>
</tr>
<tr>
<td></td>
<td>Log Risk-ratio Estimation Heuristic</td>
<td>Talluri [66]</td>
<td>2009</td>
</tr>
<tr>
<td></td>
<td>EM (Customer Choice Sets)</td>
<td>Haensel and Koole [67] and Haensel et al. [68]</td>
<td>2011</td>
</tr>
<tr>
<td></td>
<td>EM (Substitution Effects and Indirect Competitor Estimation)</td>
<td>Vulcano et al. [69]</td>
<td>2012</td>
</tr>
<tr>
<td></td>
<td>Two-step Decomposition</td>
<td>Newman et al. [18]</td>
<td>2012</td>
</tr>
</tbody>
</table>

and the impact on revenue is unlikely to be evaluated reliably due to the potentially significant difference between the estimated and actual demand parameters.

3.3. Discard the Censored Data

As noted in [19], the method of using unconstrained data only and discarding the censored data can be viewed as a complete-data method of dealing incomplete data. Known as method Naïve #2 (N2) in the reference, it is simple and easy to implement. The method performs reasonably well when the data are censored completely at random, and there is only a small amount of missing data. In other words, if the variables in the study are not related to the mechanism causing censorship, the analysis is done for the remaining data as if the censoring never had happened. But if there are some correlations between particular variables, discarding them is potentially harmful. While it is likely that this will lead to a
negatively biased forecasting because of the limited remaining sample size, there may also be a positive bias, as explained in [19].

3.4. Imputation Methods

The definition given in [19] describes “Imputation is a generic term for filling in missing data with plausible values by transforming incomplete data into complete data set.” After replacing the incomplete data, the imputed values are treated as unconstrained demand. There are various approaches for imputing the censored data, such as the mean (Naïve #3, N3), median, and percentile imputation methods. These methods are discussed in more detail in [70]. Zeni [19] further expands the discussion in the context of RM. Saleh [27] compares three imputation methods; that is, N1, N2, and N3 and concludes that N2 can vitally underestimate the true demand while N3 performs the best.

3.5. Statistical Model Methods

In recent years, statistical methods focusing on solving censored data problem have become a hotspot in research. “These models avoid the ad hoc nature of imputation methods and are built on a foundation of statistics theory. This is done at the cost of additional complexity and model assumptions that must be validated” [19]. As addressed in [22], statistical unconstraining methods cover an array of optimization and heuristic techniques that rely only on observed bookings and whether each booking class is open or closed. More discussions on the statistical methods are presented in the following section where unconstraining methods are categorized into three classes.

3.6. Categorization of Unconstraining Methods and Applications

As illustrated in Table 5, research on unconstraining techniques used in RMS, statistical methods in particular, can be classified into three major categories: single-class, multi-class, and multi-flight. The hierarchical classification is similar to the literature review parts in [23, 71]. A number of researchers have addressed deterministic or stochastic approaches to inferring latent demand in RMS of different industries. In the following sections, we provide more details of these methods in conjunction with their applications.

3.6.1. Single-Class Methods

The single-class methods stem from airline revenue management. Most of the early RM models make a potentially problematic assumption; customer demand for each of the fare classes is independent of the control policy implemented by the seller. That is, demand of any fare class does not depend on the selling status of other fares. Obviously, this may not be the case in reality [64].

The single-class algorithms use univariate and disaggregate demand models. Because the RM optimization approaches (e.g., EMSR or EMSR-b) require independent demand inputs, the single-class unconstraining techniques perform best within this framework. The assumption that demands for each flight classes are uncorrelated makes these methods unable to capture demand interactions.
(A) Airlines

(a) Methods Proposition. Swan [28, 29] first studies the problem of spill estimation in airline industry, and develops a theoretical spill formula, which becomes the basis for practitioners to unconstrain demand later. After recognizing the influences of RM on spill estimation, Swan [30, 31] revisits the basic spill model and suggests a few approximation methods.

Brummer et al. [32] assume log-normal distribution of demand data. The goal of their study is to estimate the mean and standard deviation of unconstrained bookings using the MLE technique. The mathematical derivation of the likelihood function of the censored distribution is provided. Lee [3] models the airline bookings as a censored Poisson process and develops the MLE technique to estimate the unknown parameters in these censored Poisson models. The MLE technique estimates the expected value of the demand rate based on the assumption of infinite capacity. As a result, the censored data are unconstrained.

Wickham [8] proposes a deterministic method called “Booking Profile (BP).” It is one of the earliest applications in RM demand forecasting and still widely used in practice. It assumes that the shape of the true booking profile is independent of the level of demand. Demand is forecasted with either additive or multiplicative methods. It is sensitive to the point where the censorship begins. If the data is censored at any review point all unconstrained demands will equal the observed demands, respectively.

Hopperstad [33] develops a probabilistic “Projection Detruncation (PD)” method at Boeing. Similar to the EM algorithm which will be explained below, it has an E-step and an M-step. Its accuracy is based on the value of the parameter. The PD algorithm is different from, but comparable to, the EM algorithm mainly in the way the expected value of the constrained observations is calculated. It uses the conditional median rather than the conditional mean and enables it to perform similarly to EM.

In contrast to projection methods, the “Pickup Detruncation (Pickup)” proposed by Skwarek [34] assumes that no proportional relationship exists between bookings in hand at the closure interval and final bookings. Instead, the estimate of total unconstrained bookings is obtained by adding the simple average of pickup from the closure interval on unclosed flights to bookings that are already received. In addition, he makes a comparison among Pickup, BP, and PD and expands the BP method.

Salch [35] is the first researcher who looks at the EM algorithm in the airline context and applies it to unconstrain censored data of airline passenger demand. After that, the EM method becomes the most popular statistical technique for unconstraining estimation in quantity-based RM. The name of Expectation-Maximization (EM) is given by Dempster et al. [72] in their pioneer paper. It has been successfully used in circumstances where there are censored observations, missing data, and truncated distributions [73]. The basic idea behind the EM algorithm is to take an incomplete-data problem and associate it with a complete-data problem for which MLE is computationally more tractable [70]. It is a two-step iterative process. The expectation step, replacing censored observations by sample mean, is called E-step. The maximization step, called M-step, is computing the new mean and variance for updated sample. The two steps are repeated alternately until convergence is reached. It is cited as the most accurate unconstraining method though computational intensity is required.

Van Ryzin and McGill [36] first provide the use of life table (LT) method in an RM framework, which is explained for unconstraining estimation in [74]. They use the LT method to estimate the parameters of a linear regression function in a simulation study.
Li and Oum [37] propose a generic observed load factor (OLF) table, which depends on the nominal load factor (NLF) and the value of coefficient of variation (CV). They derive formulas for calculating unconstrained demand when nominal demand is assumed to follow normal, logistic, lognormal, or gamma distributions. With these demand distributions, one can obtain the data of unconstrained demand using the OLF table.

Gao and Zhu [38] and Gao [39] propose a nonlinear programming method to analyze demand characters in the airline industry. Guo [40] and Guo et al. [41] derive formulas of EM and PD methods when nominal demand is assumed to follow certain distributions such as gamma, Weibull, exponential, and Poisson. Using the historical booking data, the simulation experiment demonstrates that the combined distribution based on Extended EM and Extended PD algorithms are more robust and have more significant impact on the expected revenue performance.

(b) Methods Comparison. Skwarek [9] and Hopperstad [42, 43] examine four unconstraining estimation methods: N2, N3, BP, and PD. They find that BP and PD are the best among these four methods, outperforming N2 by 2–3% in revenue. Pölt [20] and Weatherford [21] review five unconstraining methods (N1, N2, N3, BP, and EM). They conclude that EM is the most robust one even if it is measured in different ways. One measures sampling bias and the mean absolute error while the other examines how close the estimation is from the true mean.

Zeni [19], Zeni and Lawrence [44] compare six methods (N1, N2, N3, BP, PD, and EM). They conclude that N1 is better than N2, and EM, especially the extended EM algorithm, is the most robust method in error reduction. Their conclusions are similar to Pölt [20] and Weatherford [21] to some extent. Weatherford and Pölt [10] examine the same unconstraining methods as used by Zeni [19]. Their simulation shows that the EM and PD methods are most robust. For example, as the percentage of censorship increases by 60–80%, their estimates of the unconstrained mean increase by 20–80% over the imputation methods. Based on these findings, Chen and Luo [45] further obtain some useful results through comparing EM and PD methods.

(c) Benefit to Revenue Management. Zickus [46] examines the interactions among forecasting methods (Pickup and Regression), unconstraining estimation methods (BP and PD), and seat optimization algorithms (EMSR-b, VEMSR-b, HBP, DAVN, and Netbid) on a simulated airline network. Through using the Passenger Origin-Destination Simulator (PODS) tool, the simulation results show that a better combination of forecasting and unconstraining estimation leads to higher revenues for all of the tested seat optimization methods.

Similar to Skwarek [34] and Zickus [46], Gorin [47] evaluates the benefits of incorporating sell-up models into forecasting process. He focuses on the impact of the unconstraining models (BP, PD, and adjusted BP) on revenue gains through interacting with forecasting methods (pickup and regression), leg-based optimization algorithm (EMSR-b), as well as O-D control algorithms (GVN, Netbid, DAVN, HBP, and ProBP).

He and Luo [48] propose an improved time series forecasting method which uses the EM algorithm to unconstrain the historical bookings data. Guo et al. [49] propose a two-step time series forecasting approach. They incorporate the “unconstraining step” with Holt model, and the “forecasting step” with Holt-Winters model. They conclude that the combined time series models outperform others in simulation experiment on single-class airline revenue management problems.
(B) Hotels

Sporadic literature reports have been received for demand unconstraining in the hotel industry. As noted by Orkin [26], even when hotel reservation systems are designed to record data on lost opportunities, the data recorded are insufficient for inferring unconstrained demand. He discusses the complexity of calculating unconstrained demand and outlines how computer software can assist in decision making.

Liu et al. [50] and Liu [51] argue that parametric regression models take into account all relevant information and are computationally more feasible in real-world applications compared with EM algorithm. These models require knowledge of the demand distribution and other specifics of the demand constraints. The MLE method is used to estimate the unknown parameters in the parametric demand distributions (e.g., Weibull, Poisson, and normal regression models).

Queenan et al. [22] consider using the Double Exponential Smoothing (DES) or Holt’s method to estimate unconstrained demand. They evaluate several of the common unconstraining methods (N3, PD, EM, and LT) against their DES approach with constrained data through simulation. The results show that although it is slightly inferior to EM in some situations, it is superior to others in accuracy and implementation.

3.6.2. Multi-Class Methods

As noted in [75], in any origin-destination (O-D) market, passengers will seek the lowest available price. When a passenger’s request for his/her desired flight and fare class is declined, the reservation agent may accommodate the request with a different fare class or different flight. The buy-up behavior occurs when passenger chooses to vertically shift to a higher fare if lower fares are closed. He or she may also “buy-down” a lower fare rather than a high fare when discount fares are available on the same flight. Thus, the airline is making a vertical recapture of the traveler. The passenger could also be referred to horizontally shift to the same airline, but on another flight in the requested fare class. The airline hence makes a horizontal recapture of the passenger. If the traveler refuses the alternatives offered and switches to a competitor, the traveler is lost to the firm.

Restriction-free pricing (RFP) reduces customers’ switching costs between fare types and causes more pronounced downsell. Since then, capacity control becomes the only restrictions conducting fare class sales. In order to avoid multiple counted demands, the multi-class methods are developed. They capture the buy-up and buy-down interactions (vertical recapture) among different fare classes. Although they do not address cross-flight horizontal recapture, compared with the single-class methods, the multi-class methods are more practical and representative in the real world.

(A) Airlines

McGill [52] examines the problem of simultaneously estimating passenger demand models for two or more correlated classes of demand that are subject to a common capacity constraint. He extends the EM method to a multivariate problem. Numerical examples illustrate that good estimates could be obtained with reasonable sample sizes, even when 75% or more of the data have been censored.

Farkas [53], Belobaba and Farkas [54] present a spill model for more accurate unconstraining estimation in cases when multiple booking classes are considered. Farkas
identifies some important characteristics of RM as well as the leg-dependence effects of network that influence the unconstraining process.

Mishra and Viswanathan [55] use flight-level logistic demand curves, bookings, availability, and fare information to estimate dependent unconstrained demands in closed multi-classes. The cumulative expected bookings method proposed by them is used in commercial RM software, primarily for RFP markets with high downsell rates.

Boyd and Kallesen [56] divide demand of fare classes into “yieldable” and “priceable” categories according to passenger behavior, distribution channels, and fare class restrictions. Boyd et al. [57], Hopperstad et al. [58], and Hopperstad [59] develop negative exponential functions to model median upsell from the lowest to higher classes for RFP markets and use bookings, availability, fare, and price elasticity to estimate demand of lowest nested class. They also have “hybrid” versions for markets with partial downsell, which are used in commercial RM software.

Karmarkar et al. [60] extend the EM algorithm to cases in which demand for two dependent fare classes under consideration follows bivariate normal distribution. In an extensive simulation, four different methodologies are proposed for comparison: uncensored versus censored demand, uncorrelated versus correlated demands for two fare classes. The results show that consideration of both censored demand and dependency between fare classes can lead to a significant revenue improvement.

3.6.3. Multi-Flight Methods

Although many researchers have considered a buy-up and buy-down effect in traditional models, horizontal recapture is a complex function of how attractive different products are viewed in a market, especially the O-D environment. Industry practitioners report horizontal recapture rates in the range of 15% to 55% [23]. Neglecting the recapture effect could lead to demand overestimation bias on forecasting and inventory control process due to the “double counting” effect.

Generally speaking, multi-flight methods are probably the most difficult to calibrate because of the complexity of underlying demand models (e.g., multinomial logit, MNL, choice model), but they are able to unconstrain demands from bookings through vertical and horizontal recaptures under almost all combinations of open flights and fare classes. To some extent, they could eliminate “double counting” effects.

(A) Airlines

Ja et al. [61] develop a regression model to balance equations that are similar to the ones in [76]. It treats observed bookings and unconstrained estimates from single-class methods as known values. Reasonable results are presented on a huge test market using one-year historical data from American Airlines.

Stefanescu et al. [62] and Stefanescu [63] develop a multivariate demand model that takes both the time and product dimensions of historical demand into account. The model parameters are estimated using EM algorithm. Its performance with respect to accuracy and running time is reported based on the data from entertainment and airline industries.

Talluri and van Ryzin [64] address the impact of consumer behavior through a discrete choice model, and capture buy-up and buy-down behaviors directly. They develop an EM method to jointly estimate arrival rates and parameters of an MNL choice model based on
panel data under unobservable no-purchases. A simple method for estimating the “no-fly” purchase probability is also provided.

Vulcano et al. [65] provide empirical evidence for the potential of the approach proposed in [64]. Their simulation shows 1–5% average revenue improvements with the choice-based RM model. An MLE method that uses a variation of the EM algorithm is developed to estimate unobservable censored demand.

Ratliff et al. [23] propose a heuristic to focus on demand closure rates based on discrete choice models. The heuristic jointly estimates spill and recapture across multiple flights and fare classes. It uses balance equations that generalize the proposal in [76]. The parameters of discrete choice models are calibrated using EM algorithm. Their method is applied in commercial software for both single-leg and O-D pairs.

Talluri [66] proposes a finite-population model for RM to avoid indeterminacy and nonrobustness in the discrete finite-period model and shows that the proposed model retains good features of finite-period model. In addition, he proposes a heuristic with log risk ratio to jointly estimate the unobservable market size when customer’s purchase probabilities follow the MNL model.

Haensel and Koole [67] use customer choice sets to model customer’s buying behavior. They assume different customer groups representing various buying behaviors and characteristics. The EM method is applied to solve the problem of incomplete data or information. Haensel et al. [68] focus on applications of demand estimation on real airline reservation data.

Vulcano et al. [69] develop a method to estimate primary demand, which is customer’s first choice if all alternatives are available. The approach combines the MNL choice model with nonhomogeneous Poisson arrivals over multiple periods. The EM algorithm is applied to estimate the substitutes and lost demand when the data of sales and product availability are observable.

(B) Hotels

Newman et al. [18] propose several new estimation methods and a benchmark against the estimation procedure recommended in [64] for the MNL model. Their estimation methods are based on marginal log likelihood functions (versus expected log likelihood functions used in [64]). This enables them to eliminate the use of the EM algorithm. The advantages of their methods over EM are demonstrated with real hotel data.

3.6.4. Others

As addressed by Liu et al. [51], Queenan et al. [22], and Stefanescu [63], the problem of censored data analysis exists not only in the hotel and airline industries, but also in many other fields. The data-censoring processes in these areas are similar to the situation in which managers “terminate” demand from a particular customer segment through the use of booking limits.

There has been a great deal of theoretical and applied research on censored data analysis in reliability engineering, biomedical sciences, and econometrics (see [37, 77–88]). Parametric and nonparametric regression models and their variations are the most frequently used modeling techniques. These methods heavily rely on the use of the hazard rate function to determine the probability distribution of lifetime data. In addition, these strands
of research also include literature of inventory management, as well as retail assortment planning with substitutable products (see [89–96]).

4. Issues for Future Research

Research on data unconstraining in RMS has made great advancement in various industries and also unfolds directions for future studies.

4.1. Methods for Choice-Based RM

Customer choice models allow consumers to make their purchasing decisions on alternatives they are given. In recent years they have gained increasing attention in RM. Literature such as Talluri and van Ryzin [64], Zhang and Cooper [97, 98], among others, has focused on the single-leg and parallel flights under the customer choice behavior. Research works by Gallego et al. [99], Liu and van Ryzin [100], van Ryzin and Vulcano [101], Bront et al. [102], Kunnumkal and Topaloglu [103–105], Zhang and Adelman [106], Zhang [107], Talluri [108], Meissner and Strauss [109–111], and Meissner et al. [112] and others focus on both theoretical optimization models and practical implementation in choice-based network RM environment.

Although looking promising theoretically, the choice models have not been widely adopted in practice [18, 113]. One reason is the lack of effective estimation methodology for choice models, which involves solving for choice parameters as well as the arrival rates. The latter represent estimates of unconstrained demand. The existing methods, such as the commonly used EM algorithm, exhibit prohibitively long computational time and often lead to counter-intuitive results (see [65]). Moreover, firms may want to apply a maximum “cap” on the unconstrained estimation results because some of them tend to overestimate the mean when the demand level is low. Therefore, further robust parameter estimating methods need to consider choice models, especially those applied in an RM network environment.

In addition, as shown in Table 5, comparative researches of the single-class unconstraining methods have been conducted by some researchers for years. Although consideration of choice model in demand unconstraining is given in recent years, comparative studies concerning multi-flight methods are minimal. Research along this direction is certainly beneficial.

4.2. Methods for Multi-distribution Demand Data

Ratliff and Research [114] describe that there are three main inputs to the RM optimization models: fare levels, demand forecasts, and demand uncertainty. Among them, demand uncertainty has not received as much (published) attention as the other two. Moreover, as noted by Queenan et al. [22], like most forecasting methods, historical data need to be decomposed into promotion effects, seasonality, and competitive effects before unconstraining methods are applied. Other characters of demand, such as booking processes, market, fare class, observed level of demand, and days prior to departure, also need to be considered.

Generally speaking, most of the existing unconstraining methods, such as EM, PD and others, assume known booking curve or the distribution. But in reality, firms often do not know a priori the shape of the booking curve. In some situation, the traditional assumption of normal distribution for the nominal demand is usually inappropriate. Some
researchers, such as Lee [3], Zeni [19], Li and Oum [38], Guo [41] and Guo et al. [42], Liu et al. [51], Ratliff and Research [114], Belobaba [115] and Swan [116], have considered demand distributions other than normal distribution (e.g., the lognormal, gamma, Weibull, and Poisson distributions). There is little literature focusing on the multi-distribution case, especially under the circumstances of capturing customer choice behavior. “What is needed is a shift from models of product demand to models of customer behavior” [117]. Therefore, in order to find more practical and robust methods to model demand and its uncertainties, especially the multi-flight unconstraining methods based on multi-distribution assumptions, further study along the line is needed.

4.3. Applications of Unconstraining Methods

4.3.1. Robust RM

As shown in [117], building any reasonable model to consider customer behavior requires data at the level of customers. In practice, however, these data are often not available. Concerning the estimation of choice models in RM, a complication is that firms using RM normally have booking data from their own products. As noted in [118], “Collecting product availability from today’s RM systems is a daunting and time-consuming task.” Companies prefer to require data from loyalty programs or those provided by third parties in order to track purchase habits of a random sample of customer. While car rental industry profits from available turn-down data [119], it is a tough challenge for companies in other industries to balance the cost and profit in building customer-level models of demand.

If information on customer behavior is available, then choice models with unconstrained demand data provide flexibility, in addition to gaining forecasting or optimization power [63]. Otherwise, decisions are under limited demand information. The robust RM models hence are useful as they do not heavily depend on accurate demand information [120] and the risk-neutral assumption [121]. Recently, Lan et al. [122] propose robust inventory control methods, which follow the development in [123] by eliminating the need for both assumptions. They analyze robust booking limits for a single-leg problem when upper and lower bounds of demand are known. Birbil et al. [124] introduce robust versions of the classical static and dynamic single-leg seat allocation models, which take into account the inaccurate estimates of the underlying probability distributions. Perakis and Roels [125] derive a limited demand information model using the maxmin and minmax regret criteria under general polyhedral uncertainty sets. They provide a general approach to both single-leg and network RM problems.

Clearly, robust optimization method is another way to solve the problems resulting from forecasting inaccuracy and data insufficiency. It is worth investigating how limited demand information can be estimated through the use of demand unconstraining methods, and what kinds of robust unconstraining methods make the robust optimization policies more effective.

4.3.2. Competitive and Network Markets

The forecast method and the procedure used to update forecast parameters are important factors to determine the choice of the unconstraining method in RMS. There exist coordination problems between unconstraining and forecasting methods. Similarly, attention needs to be paid between the unconstrained estimation and seat optimization methods.
Skwarek [35], Gorin [48], and d’Huart [126] evaluate the benefits of incorporating competitive considerations into existing airline RMS, including the interactions of forecasting, unconstraining methods and seat optimization methods in competitive markets. Oppitz and Pölt [127] and Lough [128] investigate the problems of leg-based unconstraining in an O-D world. Based on various network and demand examples, Farkas [53] also shows that the influences of network and RM effects on demand unconstraining are large enough to change the fleet assignment solution. In addition, the research results of Zickus [47] indicate that the better combination of forecasting and unconstraining methods results in higher revenues on a simulated airline network.

Admittedly the methods reviewed in these studies are still limited. To evaluate latest proposed methods, especially the multi-flight unconstraining methods, it is necessary to conduct simulation tests under various market conditions. The PODS tool or other simulation methods [129, 130] can be used to generate random streams of demand data for competitive [131] and network [132] RM problems.

5. Conclusions

“Revenue management can be defined as the art of maximizing profit generated from a limited capacity of a product over a finite horizon by selling each product to the right customer at the right time for the right price. It encompasses practices such as price-discrimination and turning down customers in anticipation of other, more profitable customers.” [14] Over the past 30 years, it has been an active field of research. Demand unconstraining is one of the key techniques to the successful application of RMS. This paper reviews the latest development of unconstraining techniques in different industries. The definition of demand unconstraining and its relationship to forecasting are presented. Researchers and practitioners have made substantial studies on unconstraining methods in RMS, such as single-class, multi-class, and multi-flight methods.

The problem of censored data analysis not only exists in the RMS of airline and hotel industries, but also in many other fields, such as reliability engineering, biomedical sciences, and econometrics. Despite the growing amount of literature, more research of unconstraining methods in RMS is needed for emerging problems. For instance, what are appropriate techniques that fit choice models, especially those applied in an RM network environment; can new robust approaches reduce the number of iterations and counter-intuitive results in the process of parameter estimation; will robust optimization policies become more effective if limited demand information can be estimated through the use of unconstraining methods; what are more practical and robust multi-flight methods based on multi-distribution assumptions; how to evaluate the performance of currently proposed methods by conducting simulation tests under various market conditions.

Acknowledgments

This work was supported in part by the Major Program of the National Natural Science Foundation of China (71090402), the Program for Changjiang Scholars and Innovative Research Team in University of China (IRT0860), and the Fundamental Research Funds for the Central Universities (SWJTU11ZT32). The authors are grateful to Mark Ferguson for sharing unpublished working papers at various stages of preparation of this manuscript. They also would like to thank the valuable comments and suggestions of two anonymous referees.
References

Advances in Operations Research


[32] M. Brummer et al., Determination of Potential Load Capacity Based on Observed Load Factor Data, A Study for Northwest Airlines, St Olaf College Undergraduate Practicum Group, Northfield, Minn, USA, 1988.


