Thailand Statistician
April 2023; 21(2): 450-465
http://statassoc.or.th
Contributed paper

Meta-Analysis Unconstraining Method for Two-Class Overbooking Model

Piyachat Leelasilapasart and Murati Somboon*
Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand.
*Corresponding author; e-mail: murati.s@sci.kmutnb.ac.th

Received: 21 June 2022
Revised: 28 October 2022
Accepted: 4 January 2023

Abstract

Accurate demand forecasting is crucial for airline revenue management. However, it is difficult to forecast demand accurately since the historical data does not reflect the actual current demand. In order to obtain a better estimate of current demand, there are a number of unconstraining methods available. In this study, we used the meta-analysis (MA) technique applied to unconstraining data to improve the performance of the two-class overbooking model. The accuracy and expected profit are computed and compared to other methods often used, for example, the expectation-maximization (EM) method and the naïve methods (N1, N2, and N3). Our numerical study found that the MA produces a better MAPE in most situations with high accuracy for demand forecasting and the highest expected profit in all situations, which is greater than other methods, approximately 11.71% to 17.76%.

Keywords: Demand forecasting, static model, airline passenger, airline industry, stochastic model, revenue management.

1. Introduction

With high fixed costs and low marginal costs, the airline industry is a prime candidate for using revenue management (RM) to improve profitability. According to the Air Transport Information Division of Airport of Thailand Public Company Limited (AOT), total air traffic in Thailand increased by 21 percent from 2014 to 2015, aircraft movement increased by 17 percent, and passenger numbers by 21 percent. In value terms, this generated about USD 167 million increased revenue. Accurate demand forecasting is a vital component of RM, and flight booking data from departed flights are used to forecast the demand for future departing flights. A complicating issue is that some canceled booking requests are not included in the departed flights' data. The reservation system accepts booking requests up to a pre-determined booking limit. Hence, demand in that fare class for a given flight may exceed the booking limit; however, historical data shows only the number of reservations. At the booking limit, the demand is called "censored demand" in statistics, while it is referred to as "constrained demand" by airlines.

The statistical method of uncensored data is called “unconstraining”. Thus, unconstraining demand is crucial in forecasting future flights’ actual demand. The popular unconstraining methods, for example, naïve methods (N1, N2, and N3), projection detruncation (PD), booking profile (BP), and expectation maximization (EM), are applied in both single-class and multi-class models. In practice, most commercial RM systems are based upon a two-class rather than a multi-class model. Weatherford (2016) reviewed historical unconstraining methods but did not conduct a meta-analysis, the statistical procedure for combining data from multiple studies. In our study, we applied a meta-analysis to unconstraining data to improve the performance of the two-class overbooking model by estimating the mean demand of two classes and comparing its performance to others in terms of the expected profit through a numerical study.

The concept is that the data on booking demand is divided into groups and functionally identical following the number of updated booking limits. For example, customer’s booking demand data in a year (360 days) is divided into groups in which each group is functionally identical (we see each group as each study in meta-analysis). In this study, the updated booking limit of 13 points was used since (see Somboon and Amaruchkul 2017) informed that these gave the highest profit expectations, so the data is divided into eight groups. Then, the homogeneity test is performed to test that the data in each group are identical so that the data is a fixed effect. Next, we estimate the population’s booking demand by the FE method. Once the estimated booking demand has been obtained, use that demand in the two-class overbooking model and estimate the expected profit.

The paper is organized as follows: two-class models are reviewed in Section 2. Section 3 briefly reviews the two-class overbooking model—Section 4 details the numerical setting and the procedure for simulation. Section 5 presents numerical results, while Section 6 concludes.

2. Literature Review

2.1. Two-class and two-class overbooking model

Littlewood (1972) presented the two-class model, which focuses on the booking control problem. It does not allow overbooking, which means there are no cancellations or no-shows, and all booking requests show up at the time of service. Somboon and Amaruchkul (2016) proposed the static two-class overbooking model, which integrates two strategies, overbooking, and seat inventory, in which low fare (Class-2) customers arrive prior to high fare (Class-1) customers, and each class may have different show up rates. The airline introduces a penalty cost, any of which rejected booking requests. Furthermore, the penalties differ between the two classes; as a result, the airline may overbook Class-2 customers. We refer interested readers to (see Somboon and Amaruchkul 2016) and the literature cited therein for more details on the two-class overbooking model.

2.2. Unconstraining methods

The theoretical study in the statistical unconstraining method gained attention in the mid-1980s with the expectation maximization (EM) algorithm. It is also known as the detruncation and uncensored algorithm (see Dempster et al. 1977, Little and Rubin 1989, McLachlan and Krishnan 1997).

The EM algorithm is an iterative algorithm for estimating the maximum likelihood of problems with missing data. The presence of booking and capacity limits on past demands censors the data in historical booking records. According to Little and Rubin (1989), the EM algorithm formalizes a relatively old ad hoc idea for dealing with missing data.

Applied work in unconstrained airline demand data began in the mid-1990s. Skwarek (1996) examined four unconstrained methods, N2, N3, BP and PD, and discovered that BP and PD were the best, with a revenue increase of 2-3% over the N2 method.

Saleh et al. (1997) focused on the three naïve methods (N1, N2, and N3) and concluded that using N2 can significantly underestimate true demand and that even a simple change like switching to N3 from either N1 or N2 could be beneficial.

Pölt (2000) stated that better unconstraining demand data will improve forecast accuracy and more significant revenues. They were the first to investigate the EM algorithm in the airline context and reviewed four approaches (N1, N2, N3, and BP). The results show that N1 is better than N2 in realistic cases, and N3 gives better results than N1. However, the study concluded that EM was the most robust method for unconstraining data.

Zeni (2001) examined all six methods (N1, N2, N3, BP, PD, and EM) and reached the same conclusion as Pölt (2000) that N1 is superior to N2, and EM is the most robust method among the six, as measured by error.

3. Two-Class Overbooking Model

In this section, we are briefly describing the two-class overbooking model proposed by Somboon and Amaruchkul (2016). We refer interested readers to their article for further information.

Note that there are two classes in this study: high fare (Class-1) and low fare (Class-2). Given that $x(t)\in\mathbb{Z}_+$ is the Class-2 booking limit at the update booking limit t days before the departure, where $\mathbb{Z}_+$ is the set of non-negative integers. It implies that the reservations for Class-2 are accepted up to $x(t)$. However, overbooking is allowed if $x(t)$ exceeds capacity $\kappa(t)$, where $\kappa(t)$ is a capacity at the update booking limit point t.

An optimal booking limit at the update booking limit point $x^*(t)$ that maximizes its expected profit is preferred:

$$\begin{aligned}\pi(x(t))= \quad & E\left[\sum_{i=1}^{2} p_iB_i(x(t))-r_i\left(B_i(x(t))-W_i\left(B_i(x(t))\right)\right)\right]-E\left[h\left(W_2\left(B_2(x(t))\right)-\kappa(t)\right)^+\right]\\ & -E\left[g_2\left(D_2(t)-x(t)\right)^+\right]-E\left[g_1\left(D_1(t)-\left(\kappa(t)-B_2(x(t))\right)^+\right)^+\right].\end{aligned}$$

where $(y)^+=\max(0,y)$ for $y\in\mathbb{R}$ when $\mathbb{R}$ is the set of real number, and t is the number of days prior to departure.

Assume that $D_i(t)$ is the demand of Class-*i* reservations at the update booking limit point t for $i=1,2$. Let $W_i(y_i)$ be the number of Class-*i* show-ups and the number of Class-*i* reservations, denoted by $B_i(x(t))=y_i$, is assumed to follow a binomial distribution with parameters y_i and θ_i where θ_i is the show-up probability of Class-*i*.

Assume that p_i is the revenue of class-*i*; where $p_1>p_2>0$, for each class. Let g_i be the penalty cost including the loss-of-goodwill cost and the opportunity cost where $g_1>g_2>0$ if the airline rejects the booking request. In the case of no-show for Class-*i* reservation, the airline will refund $r_i=\gamma_i p_i$, which γ_i is a proportion of the revenue cost; however, a compensation h must be paid to passengers who denied a boarding pass where $h>p_2$.

It can be seen from the equation above that there are two sources of uncertainty, namely demand and the number of show-ups, otherwise we could find a closed form of the optimal booking limit using the theorem below.

Theorem 1. *For* $x = 0,1,...,\kappa - 2$ *and* $x = \kappa, \kappa + 1,...,$ *the expected profit function* $\pi(x)$ *is piecewise and unimodal in each piece. The expected profit* $\pi(x)$ *has a local maximum point* x' *on* $x = 0,1,...,\kappa - 2$ *as*

$$x' = \begin{cases} 0, & 0 \le \tau < P(D_1 > \kappa - 1), \\ \kappa - F_{D_1}^{-1}(1-\tau), & P(D_1 > \kappa - 1) \le \tau \le P(D_1 > 0), \\ \kappa - 2, & P(D_1 > 0) < \tau \le 1. \end{cases}$$

On the other hand, if $0 < \alpha_2 / (h\theta_2) < \bar{F}(\kappa - 1; \kappa, \theta_2)$, *then the expected profit* $\pi(x)$ *has a local maximum point* x'' *for* $x = \kappa, \kappa + 1, \dots$. *It can be written as*

$$x'' = \arg\min\{x \in \{\kappa, \kappa+1, ...\} : \bar{F}(\kappa - 1; x, \theta_2) > \frac{\alpha_2}{h\theta_2}\}.$$

Otherwise, the expected profit function is increasing.

Proof: See in Somboon and Amaruchkul (2016).

4. Unconstraining Methods

One of the tactics for success in RM is accurately forecasting demand. The reservation system accepts booking requests up to a predetermined booking limit. Hence, demand in that fare class for a given flight may exceed the booking limit, but historical data shows only the number of reservations. At the booking limit, the demand is called censored demand in the field of statistics or constrained demand for a passenger airline. The method of uncensored data is called unconstraining. Weatherford and Pölt (2002) reviewed unconstraining methods; the simplest of which are as follows:

4.1. Expectation maximization (EM)

The EM algorithm designates a relatively general idea for dealing with missing data and follows four basic steps: (1) Replace missing values with estimated values (2) Estimate parameters (3) Re-estimate the missing values assuming the new parameter estimates are correct (4) Re-estimate the parameters, and so forth, iterating until convergence as a result, each EM method iteration consists of an expectation step (E Step) and maximize step (M step). The benefit of the algorithm is that it can be demonstrated to converge reliably. However, the drawback of EM is that its convergence rate can be excruciatingly slow when a large amount of data is missing.

4.2. Naïve methods

Naïve Methods provide the three different averaging methods which are:

1) Naïve 1 (N1) – replace all constrained observations with the average of all observations.

2) Naïve 2 (N2) – replace all constrained observations with the average of all unconstrained observations.

3) Naïve 3 (N3) – replace constrained observations less than the average of all observations with the average of all unconstrained observations.

4.3. Meta-analysis method

Meta-analysis is a statistical method to combine the results of independent empirical research studies, which can be divided into two traditions: tests of the statistical significance of combined results and methods for combining estimates across studies. As mentioned, the data on booking demand is divided into groups and functionally identical following the number of updated booking limits in this study. Then, the homogeneity test is performed to test that the data in each group are identical. In statistical analysis, a fixed-effects (FE) model is one in which the model parameters are fixed quantities. In contrast to random-effects models in which all or some of the parameters of the model contain random variables.

Fixed effects can be used if two conditions are met. First, all studies or populations are functionally identical. Second, the variability of sample statistics is exclusively due to sampling errors. Hedges and Olkin (2014) presented the most commonly used FE procedure. Let k be the number of studies and y_i be sample mean of study $i; i=1,2,...,k.$ Assume that v_i is sample variance of study $i; i=1,2,...,k.$ The sample sampling variance (V_{e_i}) is computed, and the inverse variance are the weights (w_i) used to calculate the average, which is a maximum likelihood estimator for θ as

$$\hat{\theta}=\sum_{i=1}^{k} w_i y_i \Big/ \sum_{i=1}^{k} w_i ,$$

where $w_i = 1/v_i$ and $V(\hat{\theta}) = 1/\sum_{i=1}^{k} w_i.$ However, we use the function "rma" in the metafor package in R for computation in this study. "rma" is the function to fit the meta-analytic for a fixed-effect model with or without moderators; $\text{rma}(x_i, z_i, \text{method="FE"}),$ where x_i is vector of the observed effect sizes or outcomes and z_i is vector of the corresponding sampling variances. The procedure of calculation is presented in Section 5.

5. Numerical Procedure

As stated in the introduction, accurate demand forecasting is critical for determining optimal booking and overbooking. It is cumbersome to find precise forecasting demand because historical data, which are censored data, do not reflect actual demand. The booking limit determines how many seats on a flight can be sold. An airline accepts a booking in a fare class until the booking limit is reached. The airline then discontinues sales of seats in that fare class. It also stops collecting useful data. Because the booking limit is censored or "constrained," demand for travel in that fare class may exceed the booking limit, but the data does not reflect this.

Because of this, actual demand cannot be obtained directly, only estimated demand. In this setting, the N1, N2, N3, and EM methods are the most commonly used to correct constrained data in quantity-based RM (see Little and Rubin 1989). The MA method is the technique we applied to calculate estimated demand.

In this setting, we used the same parameter presented by Somboon and Amaruchkul (2017). We refer interested readers to their article for further information. Let λ_i are the seat demand of customers Class-*i*; $i=1,2.$ In this experiment, we assume that λ_1 = 40, 50, 60, 70, and λ_2 = 90, 100, 110, 120, respectively. Assume that the plane has 162 seating capacity (κ) and the fare for Class-*i*; $i=1,2$ is denoted by $p_1 = 3{,}043$ and $p_2 = 945,$ respectively. Let r_i be the refund for passenger in each class where $r_i = 0.5p_i$ and $r_i = 0.8p_i$. The compensation cost $h = 2{,}000$ baht must be paid to all passenger when the number of passengers exceeds capacity. The show-up probability θ_i on the travel day for customer in each Class-i; $i = 1,2$ are 0.7, 0.8, 0.9, and 0.95, respectively.

We partition the demand into two classes by employing the demand function proposed by Suriya (2009) for the Thailand's airline industry. Let q_i be the demand function for Class-i; $i = 1,2$.

Class-1 demand function:

$$p_1 = 3{,}588 - 188.605q_1. \tag{1}$$

Class-2 demand function:

$$p_2 = 1{,}763 - 188.605q_2. \tag{2}$$

After substituting $p_1 = 3{,}043$ in (1) and $p_2 = 945$ in (2), we obtain the demand for Class-1 and Class-2 as, $q_1 = 4.6$ and $q_2 = 6.9$ million customers, respectively. The proportion of Class-1 (υ_1) and Class-2 (υ_2) passengers is

$$\upsilon_1 = 0.4 \text{ and } \upsilon_2 = 0.6. \tag{3}$$

We considered the proportions in (3) to divide the number of reservations into two classes. Next, the MA, N1, N2, N3, and EM methods are used to estimate the censored demand data. Then, the homogeneity test is performed to test that the data in each group are identical. Assume that y_i and v_i are the sample mean and variance of study $i; i = 1,2,...,k,$ respectively, and $\hat{\theta}_i$ is the estimated average. The test statistic is

$$Q = \sum_{i=1}^{k} \frac{(y_i - \hat{\theta}_i)}{v_i}.$$

If the null hypothesis is true, Q will follow a central $\chi^2_{(k-1)}$ distribution.

The booking limit is calculated using the two-class overbooking model, then estimate the demand for each class. The profit can be computed as

$$\hat{\pi} = \sum_{i=1}^{2}\left[p_i b_i - r_i(b_i - W(b_i))\right] - h\left(W_2(b_2) - \kappa\right)^+. \tag{4}$$

The numerical study as mentioned is used to find the expected profit. The accuracy index (MAPE) of the forecast demand for the two-classes follows these steps:

1) Generate the demand for both Class-1 and Class-2 over 360 days, which is assumed to be a Poisson distribution with means $\hat{\lambda}_1$ and $\hat{\lambda}_2$, respectively.

2) Calculate the optimal initial booking limit $x^*(360)$ using Theorem 1.

3) Compute the number of Class-1 and Class-2 bookings using $x^*(360)$.

4) Recalculate the optimal booking limit $x^*(t)$ at update booking limit point t and estimate the demand for each class $(\hat{\lambda}_1, \hat{\lambda}_2)$ by using the proposed MA, EM, N1, N2 and N3 over the remaining days before departure.

a) In the MA method, partition the booking demand into eight groups by the number of updated booking and estimate the demand for each class $(\hat{\lambda}_1, \hat{\lambda}_2)$ by the FE model

b) Uncensored data by EM, N1, N2, and N3 method before estimating the demand for each class $(\hat{\lambda}_1, \hat{\lambda}_2)$ by mean.

5) Used the forecast demand for the two classes to calculate a new optimal booking limit $x^*(t)$

6) At the departure time, generate the number of show-up passengers following binomial distribution with a mean equal to the number of class-i reservations. The show-up probability is $\theta_i;\ i=1,2,$ and the profit is calculated using (4).

7) Repeat 1,000 iterations for all steps and average the profit.

6. Numerical Result

In this experiment, we simulate a booking demand as detailed in Section 5 and estimate censored booking demand data using the MA, N1, N2, N3, and EM unconstraining methods.

For the MA method, we have to test homogeneity to check that the data in each group are identical so that the fixed effect can be used. However, we uncensored data by EM, N1, N2, and N3 methods before estimating the demand for each class by mean.

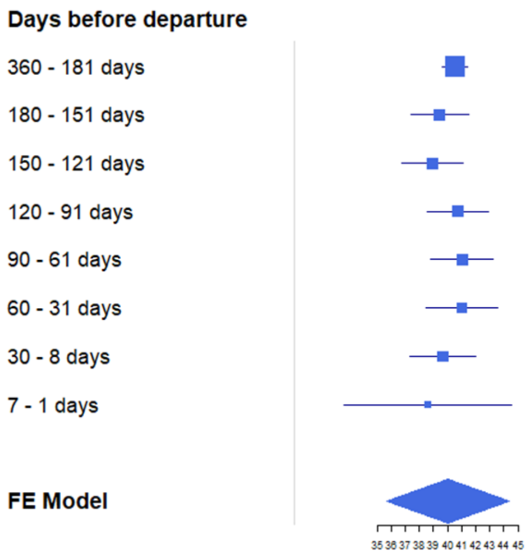

Figure 1. Testing homogeneity test for the simulated booking demand for MA method

Figure 1 and the test statistics value for homogeneity test is equal to 0.15 which less than $\chi^2_{(0.05,7)}$ so that the fixed effect can be used in this data. Then, the two-class overbooking model estimates the demand for each class and finds the optimal booking limit. The MAPE and expected profit in the same situation can be summarized below.

6.1. Comparison of the accuracy of the estimation of the average of the booking demand for both classes of customers

Table 1 shows that when $\lambda_1 =$ 40 and $\theta_1 =$ 0.95, the MAPE of the MA method gives the least MAPE, while the EM, N1, N2, and N3 methods gave similar MAPE values and increased when λ_2 has increased.

It can be seen from Table 2 that when $\lambda_1 = 50$, and $\theta_1 = 0.95$, the MA method for estimating λ_2 gives the least MAPE value followed by the EM method. N1, N2, and N3 methods gave similar MAPE values and MAPE is increased when λ_2 has increased.

Table 1 MAPE for estimating λ_2 when $\lambda_1 = 40$, $r_1 = 2{,}434.4$, $r_2 = 472.5$, and $\theta_1 = 0.95$

λ_2	θ_2	MA	EM	N1	N2	N3
90	0.7	0.07	13.32	13.42	13.41	13.34
90	0.8	0.02	13.32	13.42	13.41	13.34
90	0.9	0.39	13.32	13.42	13.41	13.34
90	0.95	0.98	13.32	13.42	13.41	13.34
100	0.7	0.72	16.02	16.60	16.63	16.09
100	0.8	0.03	16.02	16.60	16.63	16.09
100	0.9	0.24	16.02	16.60	16.63	16.09
100	0.95	0.99	16.02	16.60	16.63	16.09
110	0.7	0.18	18.24	19.94	20.25	18.82
110	0.8	0.48	18.24	19.94	20.25	18.82
110	0.9	0.69	18.24	19.94	20.25	18.82
110	0.95	0.30	18.24	19.94	20.25	18.82
120	0.7	0.46	20.20	22.51	24.52	22.17
120	0.8	0.23	20.20	22.51	24.5 2	22.17
120	0.9	0.37	20.20	22.51	24.52	22.17
120	0.95	0.53	20.20	22.51	24.52	22.17

Table 2 MAPE for estimating λ_2 when $\lambda_1 = 50$, $r_1 = 2{,}434.4$, $r_2 = 472.5$, and $\theta_1 = 0.95$

λ_2	θ_2	MA	EM	N1	N2	N3
90	0.7	0.73	6.68	7.32	7.34	6.76
90	0.8	0.32	6.68	7.32	7.34	6.76
90	0.9	0.83	6.68	7.32	7.34	6.76
90	0.95	0.26	6.68	7.32	7.34	6.76
100	0.7	0.60	10.07	11.94	12.30	10.75
100	0.8	0.71	10.07	11.94	12.30	10.75
100	0.9	0.53	10.07	11.94	12.30	10.75
100	0.95	0.39	10.07	11.94	12.30	10.75
110	0.7	0.24	13.00	15.45	17.69	15.06
110	0.8	0.16	13.00	15.45	17.69	15.06
110	0.9	0.15	13.00	15.45	17.69	15.06
110	0.95	0.79	13.00	15.45	17.69	15.06
120	0.7	0.88	15.53	17.10	23.08	20.03
120	0.8	1.01	15.53	17.10	23.08	20.03
120	0.9	0.24	15.53	17.10	23.08	20.03
120	0.95	0.12	15.53	17.10	23.08	20.03

Table 3 MAPE for estimating λ_2 when $\lambda_1 = 60$, $r_1 = 2{,}434.4$, $r_2 = 472.5$, and $\theta_1 = 0.95$

λ_2	θ_2	MA	EM	N1	N2	N3
90	0.7	0.53	0.11[a]	2.14	2.53	0.79
90	0.8	0.40	0.11[b]	2.14	2.53	0.79
90	0.9	0.02	0.11	2.14	2.53	0.79
90	0.95	1.25	0.11	2.14	2.53	0.79
100	0.7	0.23	4.25	7.02	9.43	6.56
100	0.8	0.27	4.25	7.02	9.43	6.56
100	0.9	0.20	4.25	7.02	9.43	6.56
100	0.95	0.19	4.25	7.02	9.43	6.56
110	0.7	0.41	7.82	9.54	16.08	12.75
110	0.8	0.14	7.82	9.54	16.08	12.75
110	0.9	0.40	7.82	9.54	16.08	12.75
110	0.95	0.36	7.82	9.54	16.08	12.75
120	0.7	0.05	11.01	10.90	22.15	19.17
120	0.8	0.03	11.01	10.90	22.15	19.17
120	0.9	0.19	11.01	10.90	22.15	19.17
120	0.95	0.43	11.01	10.90	22.15	19.17

Table 3 shows that the MA gives the lowest MAPE when $\lambda_1 = 60$ for all values except for the cases a and b which the EM method has the lowest MAPE. However, when $\lambda_2 = 120$, the N1 method yields less MAPE than the EM method.

Table 4 MAPE for estimating λ_2 when $\lambda_1 = 70$, $r_1 = 2{,}434.4$, $r_2 = 472.5$, and $\theta_1 = 0.95$

λ_2	θ_2	MA	EM	N1	N2	N3
90	0.7	0.03	6.38	3.33	0.63	3.82
90	0.8	0.91	6.38	3.33	0.63[c]	3.82
90	0.9	0.40	6.38	3.33	0.63	3.82
90	0.95	0.01	6.38	3.33	0.63	3.82
100	0.7	1.50	1.36	0.52[d]	7.72	4.03
100	0.8	0.10	1.36	0.52	7.72	4.03
100	0.9	0.85	1.36	0.52[e]	7.72	4.03
100	0.95	0.26	1.36	0.52	7.72	4.03
110	0.7	0.61	2.82	2.80	15.10	11.82
110	0.8	0.77	2.82	2.80	15.10	11.82
110	0.9	0.29	2.82	2.80	15.10	11.82
110	0.95	0.44	2.82	2.80	15.10	11.82
120	0.7	1.03	6.41	5.13	14.93	19.17
120	0.8	0.29	6.41	5.13	14.93	19.17
120	0.9	0.08	6.41	5.13	14.93	19.17
120	0.95	0.94	6.41	5.13	14.93	19.17

From Table 4, it can be concluded that the MA gives the best MAPE for all λ_2 when $\lambda_1 = 70$ except for the cases c, d, and e, where the N2 and N1 methods give the lowest MAPE value. However, at $\lambda_2 = 90$, the N2 method gives the low MAPE, while at $\lambda_1 = 100$, 110 and 120, the N1 method gives the low MAPE value.

The simulation study using the MA, EM, N1, N2, and N3 unconstrianing methods finds that the MA method gives the lowest MAPE, with the MAPE value between 0–1.96, while the others give the MAPE inconsistently. Based on the parameters, the MAPE values are between 0.07–24.54.

In addition, we find that the EM, N1, N2, and N3 methods give the same value of MAPE, although θ_1 and θ_2 are increased. However, in some cases, the EM, N1, N2, or N3 methods probably give a lower MAPE than the MA method; the estimate may not be an accurate approximation. As a result, the MA method is the most accurate estimating unconstraining method.

Table 5 MAPE for estimating λ_1 when $\lambda_2 = 90$, $r_1 = 2{,}434.4$, $r_2 = 472.5$, and $\theta_2 = 0.8$

λ_1	θ_1	MA	EM	N1	N2	N3
40	0.7	0.91	30.02	29.93	29.95	29.98
40	0.8	0.45	30.02	29.93	29.95	29.98
40	0.9	1.02	30.02	29.93	29.95	29.98
40	0.95	0.16	30.02	29.93	29.95	29.98
50	0.7	0.86	12.00	11.40	11.39	11.96
50	0.8	1.06	12.00	11.40	11.39	11.96
50	0.9	0.03	12.00	11.40	11.39	11.96
50	0.95	0.81	12.00	11.40	11.39	11.96
60	0.7	0.06	0.07	1.83	2.58	0.57
60	0.8	0.43	0.07[f]	1.83	2.58	0.57
60	0.9	0.63	0.07[g]	1.83	2.58	0.57
60	0.95	0.84	0.07[h]	1.83	2.58	0.57
70	0.7	1.12	8.78	11.44	13.89	10.90
70	0.8	0.44	8.78	11.44	13.89	10.90
70	0.9	0.30	8.78	11.44	13.89	10.90
70	0.95	0.47	8.78	11.44	13.89	10.90

It can be seen from Tables 5 and 6 that the MA gives the smallest MAPE estimate when $\lambda_2 = 90$ and 100 and $\theta_2 = 0.8$. However, the cases f, g and h in Table 5 show that the EM method produces the lowest MAPE while the N2 method creates the lowest MAPE in the cases of i, j and k as seen in Table 6.

Table 6 MAPE for estimating λ_1 when λ_2 = 100, r_1 = 2,434.4, r_2 = 472.5, and θ_2 = 0.8

λ_1	θ_1	MA	EM	N1	N2	N3
40	0.7	0.34	40.00	39.20	39.08	39.94
40	0.8	0.16	40.00	39.20	39.08	39.94
40	0.9	0.19	40.00	39.20	39.08	39.94
40	0.95	0.05	40.00	39.20	39.08	39.94
50	0.7	0.18	19.93	17.75	16.90	19.32
50	0.8	0.81	19.93	17.75	16.90	19.32
50	0.9	0.62	19.93	17.75	16.90	19.32
50	0.95	0.33	19.93	17.75	16.90	19.32
60	0.7	1.43	6.38	3.32	0.44[i]	3.95
60	0.8	0.49	6.38	3.32	0.44[j]	3.95
60	0.9	1.67	6.38	3.32	0.44[k]	3.95
60	0.95	0.33	6.38	3.32	0.44	3.95
70	0.7	0.69	3.30	5.59	12.07	8.57
70	0.8	0.15	3.30	5.59	12.07	8.57
70	0.9	0.44	3.30	5.59	12.07	8.57
70	0.95	0.67	3.30	5.59	12.07	8.57

Table 7 MAPE for estimating λ_1 when λ_2 = 110, r_1 = 2,434.4, r_2 = 472.5, and θ_2 = 0.80

λ_1	θ_1	MA	EM	N1	N2	N3
40	0.7	0.52	49.94	47.21	46.10	49.19
40	0.8	0.99	49.94	47.21	46.10	49.19
40	0.9	0.67	49.94	47.21	46.10	49.19
40	0.95	0.49	49.94	47.21	46.10	49.19
50	0.7	0.15	27.65	23.99	20.54	24.74
50	0.8	0.66	27.65	23.99	20.54	24.74
50	0.9	0.27	27.65	23.99	20.54	24.74
50	0.95	0.65	27.65	23.99	20.54	24.74
60	0.7	1.60	12.84	10.18	2.64	6.67
60	0.8	0.35	12.84	10.18	2.64	6.67
60	0.9	1.11	12.84	10.18	2.64	6.67
60	0.95	0.23	12.84	10.18	2.64	6.67
70	0.7	0.17	1.60	1.45	11.09	7.14
70	0.8	0.44	1.60	1.45	11.09	7.14
70	0.9	0.41	1.60	1.45	11.09	7.14
70	0.95	0.17	1.60	1.45	11.09	7.14

Still, the MA gives the lowest MAPE for λ_1 estimation when λ_2 =110 as shown in Table 7. When λ_1 = 40 and 50, the EM, N1, N2, and N3 produce similar MAPE and high values. Also, when λ_1 is increasing, the MAPE of the EM, N1, N2, and N3 decrease; unfortunately, it still produces a MAPE higher than the MA.

Table 8 MAPE for estimating λ_1 when $\lambda_2 = 120$, $r_1 = 2{,}434.4$, $r_2 = 472.5$, and $\theta_2 = 0.80$

λ_1	θ_1	MA	EM	N1	N2	N3
40	0.7	0.67	59.67	55.02	50.72	55.86
40	0.8	2.31	59.67	55.02	50.72	55.86
40	0.9	1.30	59.67	55.02	50.72	55.86
40	0.95	0.20	59.67	55.02	50.72	55.86
50	0.7	0.26	35.37	32.23	23.15	28.00
50	0.8	0.41	35.37	32.23	23.15	28.00
50	0.9	0.64	35.37	32.23	23.15	28.00
50	0.95	0.46	35.37	32.23	23.15	28.00
60	0.7	0.25	18.37	18.34	3.76	8.33
60	0.8	0.07	18.37	18.34	3.76	8.33
60	0.9	0.17	18.37	18.34	3.76	8.33
60	0.95	1.36	18.37	18.34	3.76	8.33
70	0.7	0.31	7.07	8.55	2.77	7.14
70	0.8	0.40	7.07	8.55	2.77	7.14
70	0.9	0.53	7.07	8.55	2.77	7.14
70	0.95	1.31	7.07	8.55	2.77	7.14

From our numerical study using the MA, EM, N1, N2, and N3 unconstraining methods to estimate the demand, we conclude that the MA method produces the smallest MAPE, with the MAPE value being between 0–2.68, while the others give the inconsistent MAPE. The MAPE values were between 0.03–59.67, and we find that the EM, N1, N2, and N3 methods give the same MAPE, although θ_1 and θ_2 are increased. In addition, we find that the MAPE of λ_1 estimated from the EM, N1, N2, and N3 methods decrease when they increase until $\lambda_1 = 60$. However, the MAPE will increase at $\lambda_1 = 70$. The MA method produces a MAPE that is relatively stable. Although the EM, N1, N2, or N3 methods may give a lower MAPE, the estimate may not be an accurate approximation. We conclude that the MA method is the most accurate unconstraining method for demand estimation.

6.2. Comparison of the expected profit for all unconstraining methods

In this section, we present some results of the expected profit of all unconstraining methods. Table 9 shows that the expected profit at $\lambda_1 = 40$, $\lambda_2 = 90$, $r_1 = 2{,}434.4$ and $r_2 = 472.5$. The MA method produces the highest profit expectation at all θ_1 and θ_2. Even, the expected profit at $\lambda_1 = 50$, $\lambda_2 = 90$, $r_1 = 2{,}434.4$ and $r_2 = 472.5$. The MA method produces the highest profit expectation for all θ_1 and θ_2 as shown in Table 10.

Table 9 The expected profit when $\lambda_1 = 40$, $\lambda_2 = 90$, $r_1 = 2434.4$ and $r_2 = 472.5$

θ_1	θ_2	MA	EM	N1	N2	N3
0.7	0.7	197,101.13	182,801.16	182,586.19	182,574.54	182,784.54
0.7	0.8	202,692.01	186,508.10	186,287.64	186,275.62	186,491.96
0.7	0.9	207,369.99	190,188.07	189,964.51	189,951.73	190,170.27
0.7	0.95	212,111.58	192,017.97	191,792.25	191,779.42	192,000.19
0.8	0.7	208,839.35	195,556.25	195,337.30	195,323.60	195,535.07
0.8	0.8	213,419.96	199,263.19	199,038.75	199,024.68	199,242.48
0.8	0.9	219,937.58	202,943.16	202,715.62	202,700.79	202,920.80
0.8	0.95	223,387.27	204,773.05	204,543.36	204,528.47	204,750.72
0.9	0.7	220,591.42	208,264.89	208,030.42	208,015.50	208,239.63
0.9	0.8	225,598.04	211,971.84	211,731.87	211,716.58	211,947.05
0.9	0.9	231,584.52	215,651.80	215,408.74	215,392.69	215,625.37
0.9	0.95	234,936.83	217,481.69	217,236.48	217,220.37	217,455.28
0.95	0.7	226,965.67	214,506.37	214,255.18	214,240.54	214,482.46
0.95	0.8	230,990.24	218,213.31	217,956.63	217,941.62	218,189.88
0.95	0.9	237,086.26	221,893.28	221,633.50	221,617.72	221,868.20
0.95	0.95	239,136.17	223,723.17	223,461.24	223,445.41	223,698.11

Table 10 The expected profit when $\lambda_1 = 50$, $\lambda_2 = 90$, $r_1 = 2434.4$ and $r_2 = 472.5$

θ_1	θ_2	MA	EM	N1	N2	N3
0.7	0.7	212,644.12	196,754.42	195,482.95	195,462.94	196,681.15
0.7	0.8	218,184.11	200,758.13	199,453.54	199,433.15	200,685.14
0.7	0.9	221,896.62	204,714.81	203,394.79	203,376.39	204,626.94
0.7	0.95	223,218.42	206,722.22	205,387.19	205,366.80	206,634.76
0.8	0.7	226,459.34	210,475.54	209,136.06	209,114.57	210,332.16
0.8	0.8	231,839.42	214,479.32	213,106.64	213,084.78	214,336.15
0.8	0.9	237,018.66	218,436.00	217,047.90	217,028.02	218,277.95
0.8	0.95	238,877.70	220,443.41	219,040.30	219,018.43	220,285.77
0.9	0.7	241,676.55	224,147.91	222,715.77	222,725.58	223,995.92
0.9	0.8	247,417.91	228,151.69	226,686.35	226,695.79	227,999.91
0.9	0.9	250,839.97	232,108.26	230,627.61	230,639.03	231,941.71
0.9	0.95	252,403.00	234,115.67	232,620.01	232,629.44	233,949.53
0.95	0.7	248,523.30	230,721.37	229,201.95	229,187.07	230,575.98
0.95	0.8	251,848.52	234,725.15	233,172.53	233,157.28	234,579.97
0.95	0.9	256,323.46	238,681.83	237,113.79	237,100.52	238,521.77
0.95	0.95	257,143.99	240,689.13	239,106.19	239,090.94	240,529.59

Table 11 The expected profit when $\lambda_1 = 60$, $\lambda_2 = 90$, $r_1 = 2434.4$ and $r_2 = 472.5$

θ_1	θ_2	MA	EM	N1	N2	N3
0.7	0.7	229,308.61	210,273.42	206,523.60	205,625.97	209,698.92
0.7	0.8	231,570.45	214,532.27	210,721.34	209,765.85	213,978.91
0.7	0.9	235,486.01	218,766.84	214,884.62	213,905.63	218,141.16
0.7	0.95	239,284.81	220,892.52	216,999.98	216,031.37	220,267.69
0.8	0.7	244,367.58	224,935.21	220,848.32	219,719.87	224,022.87
0.8	0.8	247,148.94	229,194.02	225,046.06	223,859.74	228,302.85
0.8	0.9	252,665.12	233,428.59	229,209.34	227,999.53	232,465.10
0.8	0.95	254,078.20	235,554.27	231,324.70	230,125.27	234,591.63
0.9	0.7	261,919.61	239,418.33	235,282.01	234,151.91	238,452.19
0.9	0.8	267,200.64	243,698.00	239,479.75	238,291.79	242,732.18
0.9	0.9	269,657.23	247,912.26	243,643.03	242,431.57	246,894.43
0.9	0.95	271,252.90	250,037.39	245,758.39	244,557.32	249,020.95
0.95	0.7	271,886.80	246,796.09	242,220.70	241,089.37	245,813.77
0.95	0.8	273,317.07	251,056.41	246,418.44	245,229.25	250,093.76
0.95	0.9	277,695.17	255,290.98	250,581.72	249,369.03	254,256.01
0.95	0.95	281,780.01	257,416.66	252,697.08	251,494.77	256,382.53

Table 11 shows results trending in the same direction as in Tables 9-10. The MA method gives the best profit expectation for all θ_1 and θ_2.

Table 12 The expected profit when $\lambda_1 = 70$, $\lambda_2 = 90$, $r_1 = 2,434.4$ and $r_2 = 472.5$

θ_1	θ_2	MA	EM	N1	N2	N3
0.7	0.7	241,761.59	222,781.08	218,186.42	212,326.89	220,015.90
0.7	0.8	244,882.09	227,383.63	222,694.43	216,623.98	224,524.72
0.7	0.9	248,138.16	231,945.56	226,989.34	220,899.45	228,800.58
0.7	0.95	250,744.10	234,156.00	229,179.78	222,987.81	231,010.48
0.8	0.7	260,636.48	237,702.52	232,832.54	226,974.50	234,661.69
0.8	0.8	262,841.88	242,317.23	237,340.54	231,271.59	239,170.51
0.8	0.9	267,155.24	246,893.54	241,635.45	235,547.06	243,446.37
0.8	0.95	269,700.32	249,103.98	243,825.89	237,635.42	245,656.27
0.9	0.7	282,027.10	253,555.27	248,722.65	241,806.63	249,498.05
0.9	0.8	284,433.33	258,157.82	253,230.66	246,103.72	254,006.87
0.9	0.9	288,259.81	262,734.13	257,525.57	250,379.20	258,282.74
0.9	0.95	290,731.86	264,956.72	259,716.01	252,467.56	260,492.63
0.95	0.7	289,725.89	261,666.50	255,963.62	249,014.30	257,799.44
0.95	0.8	292,168.64	266,460.70	260,471.63	253,311.39	262,308.26
0.95	0.9	296,086.41	271,037.01	264,766.53	257,586.86	266,584.13
0.95	0.95	298,327.19	273,247.45	266,956.97	259,675.22	268,794.02

As same as the earlier results, the MA method still produces the best profit expectation for all θ_1 and θ_2. The simulation study shows that the MA method produces a higher expected profit than the EM, N1, N2, and N3 methods. The MA gives the highest profit expectation, 11.71%, 14.41%, 14.41%, 17.76%, and 13.81% more than the EM, N1, N2, and N3 methods, respectively. Furthermore, the hypothesis testing found that the MA method gave the highest expected profit and differed from the EM, N1, N2, and N3 methods at a significance level of 0.05. So, it can be concluded that applying the MA method to the two-class overbooking model will give airlines the highest profit expectations.

7. Conclusions

The booking demand data is simulated to compare the performance of meta-analysis (MA) applied to unconstraining data with the other unconstraining methods N1, N2, N3, and EM. We find that the MA method gives both classes an approximate average booking demand close to the actual booking demand. The estimates variate follows all parameters in the two-class overbooking model. In contrast, when using EM, N1, N2, and N3 unconstraining methods, the estimated average booking demand is not close to the average actual booking demand in almost all study cases. In the two-class overbooking model, meta-analysis applied to unconstraining data provided the highest profit expectations.

Acknowledgements

This research was funded by King Mongkut's University of Technology North Bangkok, Contract no. KMUTNB-62-NEW-10.

References

Dempster AP, Laird NM, Rubin, DB. Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc Ser B Methodol. 1977; 39(1): 1-22.

Hedges LV, Olkin I. Statistical methods for meta-analysis. Florida: Academic Press; 2014.

Little RJ, Rubin DB. The analysis of social science data with missing values. Sociol Methods Res 1989; 18(2-3): 292-326.

Littlewood K. Forecasting and control of passenger bookings. J Revenue Pricing Manag. 1972; 4(2): 111-123.

McLachlan G, Krishnan T. The EM algorithm and extensions. New Jersey: John Wiley & Sons; 1997.

Pölt SP. From booking to demand: the process of unconstraining. Proceeding of the AGIFORS Reservations and Yield Management Study Group, New York, USA; 2000.

Saleh R, Chatterjee H. Estimating lost demand with imperfect availability Indicators. Proceedings of the AGIFORS Reservations and Yield Management Study Groups, Montreal, Canada; 1997.

Skwarek DK. Competitive impacts of yield management system components: forecasting and sell-up models. PhD [dissertation]. Cambridge (MA): Massachusetts Institute of Technology; 1996.

Somboon M, Amaruchkul K. Combined overbooking and seat inventory control for two-class revenue management model. Songklanakarin J Sci Technol. 2016; 38(6): 657-665.

Somboon M, Amaruchkul K. Applied two-class overbooking model in Thailand's passenger airline data. Asian J Shipp. Logist. 2017; 33(4): 189-198.

Suriya K. The impact of low cost airlines to airline industry: an experience of Thailand. J Ekon Malays. 2009; 43: 3-25.

Weatherford L, Pölt S. Better unconstraining of airline demand data in revenue management systems for improved forecast accuracy and greater revenues. J Revenue Pricing Manag. 2002; 1(3): 234-254.

Weatherford L. The history of forecasting models in revenue management. J Revenue Pricing Manag. 2016; 15(3): 212-221.

Zeni RH. Improving forecast accuracy in revenue management by unconstraining demand estimate from censored data. PhD [dissertation]. Newark (NJ): Rutgers University; 2001.