Journal of Travel & Tourism Marketing, 2015

ISSN: 1054-8408 print / 1540-7306 online
DOI: 10.1080/10548408.2015.1063826

ANALYZING THE USE OF AN ADVANCE BOOKING CURVE IN FORECASTING HOTEL RESERVATIONS

Tony Sze Ming Tse
Yiu Tung Poon

ABSTRACT. Although there is considerable interest in the advance bookings model as a forecasting method in the hotel industry, there has been little research analyzing the use of an advance booking curve in forecasting hotel reservations. The mainstream of advance booking models reviewed in the literature uses only the bookings-on-hand data on a certain day and ignores the previous booking data. This empirical study analyzes the entire booking data set for one year provided by the Hotel ICON in Hong Kong, and identifies the trends and patterns in the data. The analysis demonstrates the use of an advance booking curve in forecasting hotel reservations at property level.

KEYWORDS. Advance bookings model, revenue management, forecasting, booking curve

INTRODUCTION

Revenue management (RM) in the hospitality sector is concerned with revenue maximization through the manipulation of room rates offered to customers (Sanchez & Satir, 2005). Implicit in managing revenue is the pricing of different room types, inventory control, allocation of stock to different distribution channels, and forecasting demand. It is widely accepted that RM is inextricably associated with forecasting room demand (Andrew, Cranage, & Lee, 1990; Anderson & Xie, 2010; Cranage, 2003; Cross, 1997; Cross, Higbie, & Cross, 2009; Helsel & Cullen, 2006; Kiely, 2008; Law, 1998, 2004; Orkin, 1998; Schwartz & Cohen, 2004; Upchurch, Ellis, & Seo, 2004; Weatherford, Kimes, & Scott, 2001; Yuksel, 2007) and the development of a good forecasting system is vitally important for successful RM implementation (Talluri & Ryzin, 2005).

Forecast of future demand is important because the forecast helps make important decisions in RM. Many studies have developed algorithms based on some kind of correlation (usually negative) between price and quantity demanded to help make decisions regarding the acceptance (or rejection) of bookings by guests, and room rates offered by hotels (Badinelli, 2000; Baker & Collier, 1999, 2003; Bitran & Mondschein, 1995; Choi & Cho, 2000; Lai & Ng, 2005; Rest & Harris, 2008). Obviously, the appropriateness of such decisions depends on the accuracy of demand forecasting.

As illustrated by Weatherford and Kimes (2003), many of the forecasting studies in hotel RM have their origin in the airline industry, such as those by Sa (1987) and Lee (1990). Lee (1990),

Tony Sze Ming Tse, PhD, School of Hotel and Tourism Management, The Hong Kong Polytechnic University, 17 Science Museum Road, TST East, Kowloon, Hong Kong SAR, China (E-mail: tony.tse@polyu.edu.hk).
Yiu Tung Poon, PhD, Department of Mathematics, Iowa State University, 100 Enrollment Services Center, Ames, IA 50011, USA (E-mail: ytpoon@iastate.edu).
Address correspondence to: Tony S.M. Tse, School of Hotel and Tourism Management, The Hong Kong Polytechnic University, 17 Science Museum Road, TST East, Kowloon, Hong Kong SAR, China (E-mail: tony.tse@polyu.edu.hk).

for example, differentiates between historical bookings model and advance bookings model in the airline industry, which later serve as the basis of much discussion in hotel forecasting.

In the historical bookings model, the focus is on the final bookings of a particular flight at various points in the past. Researchers study the data which represent final bookings on flight days in the past, assuming that such data reveal cyclical trends and seasonal variations over a period of time. The past behavior of a time series is examined in order to infer something about its future behavior. The historical bookings model uses time series analysis to forecast final bookings for a particular flight on a particular day (or week or month) in the future.

In the advance bookings model, researchers consider the bookings already made on the particular flight for which a forecast of bookings is required. For departed flights, the booking data are already complete; whereas for flights scheduled in the future, the booking data are only partially complete. In trying to predict the incomplete portions of future bookings for a particular flight, the advance bookings model uses the bookings-on-hand on a certain day before departure and uses a sample of booking data from previous departures of the same flight number to estimate the bookings-to-come.

While the application of the advance bookings model in the airline industry is well developed, its application in the hotel industry is limited. This study investigates the analysis and application of the advance bookings model in the hotel industry. Specifically, the advance bookings data of the Hotel ICON for one year are analyzed, and the purpose of the study is to identify and examine the booking patterns and trends, use the advance bookings curve to forecast hotel reservations, and provide insights into the application of the advance bookings model in the hotel industry.

LITERATURE REVIEW

Forecasting in the Airline Industry

This literature review focuses on booking data based on forecasting in the airline industry. Adams and Vodicka (as cited in Weatherford & Kimes, 2003) developed short-term forecasts for just one week before departure using fairly simple estimates and an average of segment class reservations in the airline business. L'Heureux (as cited in Weatherford & Kimes, 2003) discusses the classical "pickup" method to determine the average of reservations picked up between different reading days for departed flights for a particular day of the week to forecast the future pickup between the same reading days for the same flight number on the same day of the week in the future. Sa's (1987) study examines distribution plots of reservations in airline by fare class. The first reservation data retrieved for any flight refers to a period that corresponds to 28 days before departure. From there on, snapshots are taken for every seven days. That is to say, each flight will be analyzed in five 7-day periods, from day 28 to boarding day. Analyses are made in terms of reservations up to a particular period (or bookings-on-hand) and the expected number of reservations still to come (or bookings-to-come). The study observes these two related but distinct variables in terms of shape, means, and standard deviation, and finds that bookings-on-hand do not exhibit good explanatory power when forecasting final bookings via bookings-to-come.

Lee (1990) studies the available booking data of a typical airline database and differentiates between historical bookings model and advance bookings modeling. Based on current and past data-on-hand, the researcher estimates a statistical model and fits a model to the known past observations. Once a model is fitted to the data, the researcher predicts future (unknown) values given the current data-on-hand. Producing an accurate forecast requires the use of estimation modeling and intelligent extrapolation. There are two types of modeling. The first type is the *synthetic booking curve* model, which attempts to describe the shape of the booking curve based on a sample of advance booking data from previous departures of the same flight number. The booking curve is synthesized from an approximation to its shape and other exogenous factors. The second type of advance bookings model is the *time series of advance bookings* model, which expresses total bookings a certain number of days before departure as a time series of total bookings at earlier points in the booking history of all flights with the same flight number and exogenous factors.

Forecasting in the Hotel Industry

From its origin in the airline industry nearly 60 years ago, RM has expanded to other hospitality industries, notably lodging (Anderson & Xie, 2010). Researchers in hotel RM have followed the line of thought in airline RM in the last 25 years. Relihan (1989) and Kimes (1989) were among the earliest researchers in hotel RM to treat booking data as booking curves or threshold curves. Threshold curves represent reservations-on-hand that a hotel expects in relation to the number of days before arrival. If the current reservations-on-hand fall outside the stated range, then the hotel should adjust the room rates either upwards or downwards to maximize revenue.

Weatherford and Kimes (2003) describe seven forecasting methods, and test and compare them using reservations data from Choice Hotels and Marriott Hotels. The seven forecasting methods are simple exponential smoothing; moving average methods; linear regression; logarithmic linear regression; additive method; multiplicative method; and Holt's double exponential smoothing. They conclude that the best forecasting method varies according to property, rate category, and length of stay, and that pickup methods and regression produced the lowest error, while the booking curve and combination forecasts produced fairly inaccurate results. Schwartz (2006) argues that hotel booking is remarkably different from what is described by the traditional approach to modeling consumers' purchase cycle because of the advance-booking nature, and an in-depth understanding of the process and its determinants is essential in sustaining the effectiveness of an RM system. Schwartz (2008) further notes that with the online booking revolution the advance-booking behavior characteristics are not yet fully understood, and the author demonstrates how the shape and pattern of the time-before-the-date-of-stay determines the forecasting model predictions. The implications of consumers making last-minute travel arrangements on marketing strategy and RM are also discussed by Dacko (2004).

Of particular interest to this study is the advance booking method. Schwartz and Hiemstra (1997) developed an extrapolative forecasting model by focusing on the shape of past booking curves, and compare four models of forecasting. Chen and Kachani (2007) continue with advance bookings modeling and find that using exponential smoothing, eight weeks of history provide forecasts of the lowest error, while four weeks of history provide the best expected revenue from optimization. Phumchusri and Mongkolkul (2012) use actual booking data to show that booking information and the date of stay are essential in predicting final room demand. Their regression model yields the lowest forecast error, especially for the short-term forecasting, when compared with traditional forecasting methods.

Rajopadhye, Ghalia, Wang, Baker, and Eister (2001), on the other hand, simulate the process of hotel reservation requests using a Poisson distribution to forecast occupancy. They apply the Holt–Winters smoothing method to the simulated booking data and forecast the demand for rooms for each arrival day. The forecast is based on past observations, with recent observations being given more weight and observations further in the past given less weight. The limitation of the study is that it relies on simulated arrivals and simulated cancelations for the analysis, and the model is not supported by empirical study. It is noted that Rajopadhye et al. (2001) forecast unconstrained hotel room demand, which is the number of rooms that can be sold if there are no capacity or pricing constraints. The assumption of unconstrained hotel room demand probably has its origin in the airline industry. It is reasonable and functional for an airline to forecast unconstrained demand because it has the option and flexibility to change the number of flights and type of aircraft to match demand. It is, however, questionable to make such assumptions in the context of the hotel industry. In the case of room demand and supply, a hotel does not have much leeway to change the room inventory to match demand, therefore forecasting unconstrained demand in the hotel industry is not as meaningful as in the case of the airline industry.

In fact, the hotel industry believes that booking pickup is one of the most accurate prognosticators of future performance (Rubicon, 2010), and it has been reported that the average hotel reservation lead time is 36 days for leisure travel and 21 days for business travel (HotelNewsNow, 2012). According to the American Hotel & Lodging Association (AHLA, 2006), knowing the typical percentage of business that books in a certain window prior to arrival is important because appreciating when business actually starts to book will assist in determining how far in advance of the day of arrival restrictions should be placed. Furthermore, today's travelers are increasingly booking their itineraries days in advance, not weeks or months ahead, and the industry is most interested in how to beat the incredible shrinking hotel booking window (Mourier, 2011).

While it seems that there is considerable interest in the advance bookings model among academics and the industry, there has been little research analyzing booking data in detail and making use of the model for forecasting (Zakhary, Gayar, & Atiya, 2008). This is partly due to the fact that (airline) reservations data are confidential (Sa, 1987), and the same can be said of hotel booking data. Discussion with industry practitioners also reveals that common practice in forecasting involves the use of only current bookings-on-hand and tends to ignore previous booking data. This could be due to the use of highly complex algorithms in some advance bookings modeling and the belief that the past booking pattern has little relationship with the future booking pattern.

This study analyzes the trends and patterns in the booking data made available by the Hotel ICON, and explores how the findings can be used in forecasting. The researchers understand that the trends and patterns identified are unique to a particular hotel, and the findings cannot be generalized. Nonetheless, the analysis demonstrates a more in-depth understanding of advance bookings modeling, and explores a forecasting method which is based entirely on bookings-on-hand data.

ADVANCE BOOKINGS MODEL

Defining Booking Data

The advance bookings model is based on booking data of a hotel. Hotel booking data are the number of rooms reserved by guests over a period of time, for example, 90 days or 60 days before a certain day, and we call this Day 0. In a normal situation, reservations increase as the hotel approaches Day 0. The reservations pickup rate varies from hotel to hotel, and from period to period, depending on macro factors and micro factors. If macro factors such as the economy in a particular source market and its exchange rate are favorable, the bookings by guests from that source market are likely to be strong. If micro factors such as hotel pricing and promotions are effective, the bookings received are likely to increase quickly. In other words, the bookings-on-hand are a good indicator or reflection of all the macro and micro factors. Therefore it can be argued that the bookings-on-hand subsume all the factors affecting room occupancy, and such effects will carry through over the entire booking period. If a booking trend based on the advance bookings-on-hand can be estimated or established, it could be a good indicator of final room occupancy.

A characteristic of using booking trend as an indicator of final room occupancy is that the resulting forecast is specific to a particular hotel, and not the total demand for all hotels in a destination. Of course one can aggregate the forecasts of all hotels to arrive at total demand in a destination if needed.

Advance Bookings Modeling

The advance bookings model considers the increase in the number of reservations over a relatively short period of time, usually days, for a particular day in a hotel. Researchers assume that the number of reservations in a hotel increases gradually and use additive or multiplicative models to forecast room occupancy in the future. The following methods have been applied in advance bookings models (Weatherford & Kimes, 2003):

- Linear regression methods assume that there is a correlation between the current number of reservations-on-hand (Day n) and the final number of reservations (Day 0):

$$\text{Forecast@Day } 0 = \text{a} + \text{b} \times \text{bookings@Day n}$$

- Logarithmic linear regression methods also assume a different relation between the current number of reservations-on-hand (Day n) and the final number of reservations (Day 0):

$$\text{Log}\left(\text{Forecast@Day } 0\right) = \text{a} + \text{b} \times \text{Log}\left(\text{bookings@Day n}\right)$$

- The additive method adds the current bookings to the average historical pickup in bookings from the reading day (Day n) to the actual night of the stay (Day 0):

$$\text{Forecast@Day } 0 = \text{bookings@Day n} + \text{average pickup}\left(\text{Day n to Day } 0\right)$$

- The multiplicative method multiplies the current bookings by the average historical pickup ratio in bookings from the reading day (Day n) to the actual night of the stay (Day 0):

$$\text{Forecast@Day } 0 = \text{bookings@Day n} \times \text{average pickup ratio}\left(\text{Day n to Day } 0\right)$$

We note that all of the above methods use only the bookings-on-hand on a certain day (bookings@Day n) and ignore the booking data before Day n, and try to estimate the bookings-to-come based on some other historical data. This practice assumes that future bookings on a particular day depend only on the current number of bookings-on-hand, not on how the bookings-on-hand were generated. In statistics, this is known as the Markov property, which states that the probability of any particular future behavior of the process, given its current state, is not altered by additional knowledge concerning its past behavior (Lee, 1990). The practice of assuming the Markov property in advance bookings modeling has its origin in the airline industry (Bertsimas & Boer, 2005; Haerian, Homem-de-Mello, & Mount-Campbell, 2006; Lee, 1990; Littlewood, 1972). Haerian et al. (2006), for example, developed a booking process model which assumes that requests for seat reservations follow a homogeneous Poisson process, meaning that the Poisson parameter is constant and therefore time independent.

This study analyzes the empirical booking data of a hotel, identifies the trends and patterns identified in the booking data, and uses its advance bookings curve to forecast hotel reservations.

RESEARCH DATA AND METHOD

Research Data

Hotel ICON is a 262-room teaching and research hotel owned by The Hong Kong Polytechnic University, Hong Kong, operated on a commercial basis (Tse, 2012). The booking data provided by the hotel covers the 12-month period September 1, 2011–August 31, 2012. There are 366 days during the 12-months period, and for each day, there are 91 data corresponding to final occupancy (on Day 0) and the number of rooms reserved ranging from 1 day to 90 days prior to arrival. Mathematica is the software used to analyze the data.

Table 1 presents the set of data made available to the researchers. The figures in row 90 are the reservations 90 days before Day 0, the figures in row 3 are the reservations 3 days before Day 0, and the figures in row 0 are the final occupancies. For example, for August 31, 2012, the number of rooms reserved 90 days out is 32, the number of rooms reserved 3 days out is 207, and the final occupancy is 238.

TABLE 1. Number of Rooms Booked 1 – 90 Days Before Each Day Between September 1, 2011, and August 31, 2012

Number of days before Day 0	September 1, 2011	September 2, 2011	September 3, 2011	…	August 29, 2012	August 30, 2012	August 31, 2012
90	12	11	9	…	31	31	32
89	12	11	9	…	31	31	32
88	12	11	9	…	31	31	33
:	:	:	:	…	:	:	:
3	206	248	245	…	174	189	207
2	207	247	244	…	183	206	214
1	213	245	244	…	201	214	232
0	207	243	244	…	206	220	238

DATA ANALYSIS

Distribution Analysis

Using the analysis method by Sa (1987), we examine the distribution of 90-days bookings in each of the 366 days in seven-day intervals. For each day in the data set, we first retrieve the booking data corresponding to seven days before Day 0, and from there on, snapshots are taken for every seven days. In other words, the bookings each day are analyzed in 12 seven-day intervals and there are 12 views of bookings-on-hand corresponding to 7, 14, 21, 28, 35, 42, 49, 56, 63, 70, 77, and 84 days before Day 0. We conduct the same analysis for bookings still to come in terms of percentage of final room occupancy, and obtain the corresponding 12 views of bookings-to-come. Figure 1 shows the distribution plots of bookings-on-hand and bookings-to-come in seven-day intervals. The bookings-on-hand and bookings-to-come distribution plots for each interval are put side by side for comparison. We find that the distributions by and large are normal during the six weeks period before Day 0, and as expected, the central tendency of bookings-on-hand moves to the right and the central tendency of bookings-to-come moves to the left as Day 0 approaches.

Analysis of Trends and Patterns

We further study the distributions by looking at the means and standard deviations of the bookings-on-hand and bookings-to-come, which are shown in Table 2.

While we expect to see the overall trend of bookings-on-hand increase as Day 0 approaches, the following patterns can also be observed:

1. The pickup is rather modest between 84 days and 56 days before Day 0, and the pickup is much faster between 56 days and 7 days before Day 0.
2. Almost 80% of the bookings appear in the two months before Day 0.
3. The booking window of 28 days before Day 0 accounts for more than half (54.4%) of the bookings.
4. The standard deviation is relatively large in the early days, meaning that the bookings received then fluctuate considerably.

We repeat the same analysis for each day of the week and find that the same trend appears every day of the week (Appendix). By plotting all the bookings-on-hand against the number of weeks before Day 0, we can visualize the trends in the booking data (Figure 2). We notice that the booking trends on Friday and Saturday stand out among the seven days of the week. Not only do Friday and Saturday exhibit higher final occupancy and bookings-on-hand throughout the 90 days, the percentage of bookings-to-come is consistently lower than the averages. This is an indication of the pattern that bookings on Friday and Saturday are stronger and pickup quicker.

FIGURE 1. Distribution Plots of Bookings-On-Hand and Percentage of Bookings-To-Come

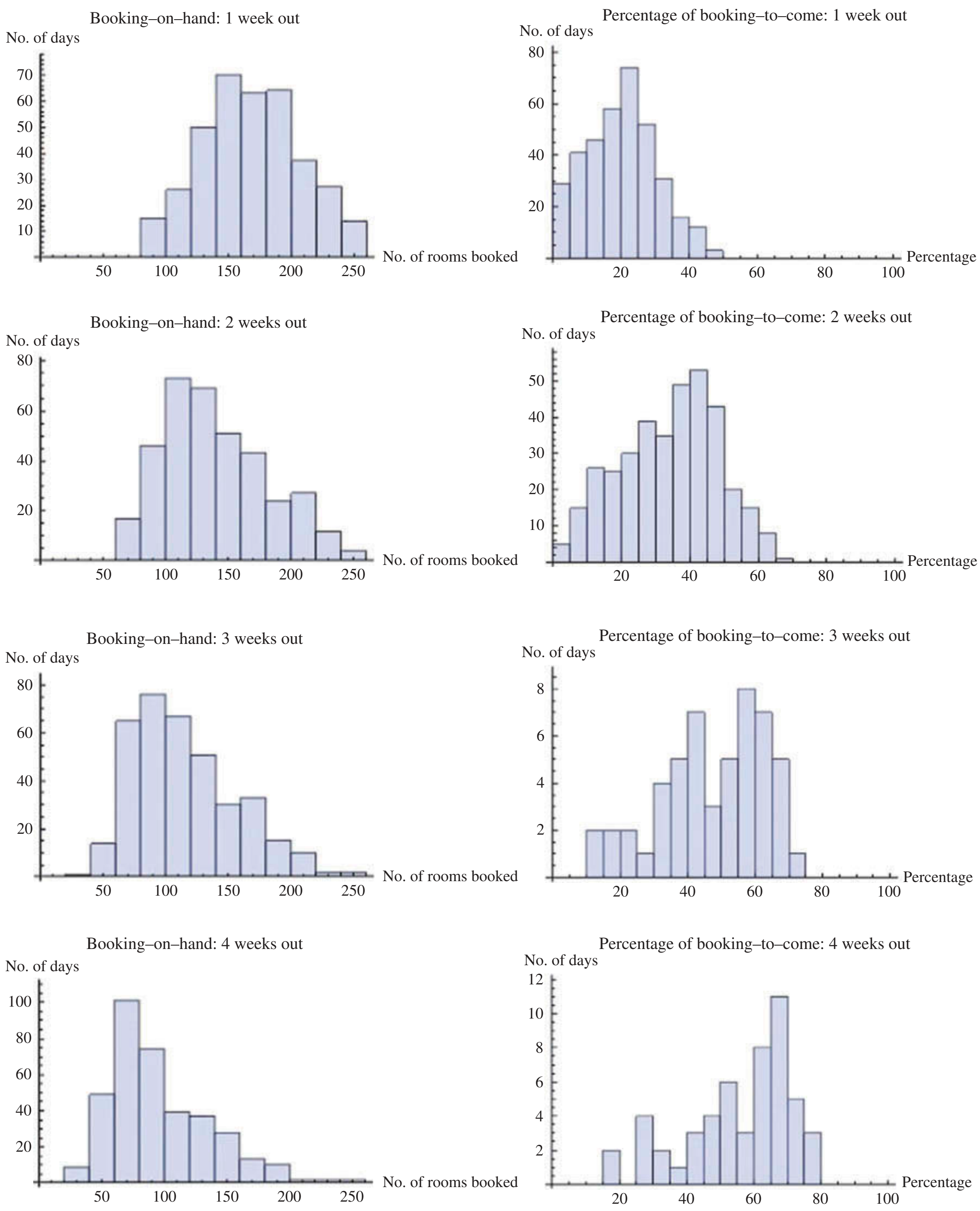


FIGURE 1. Continued

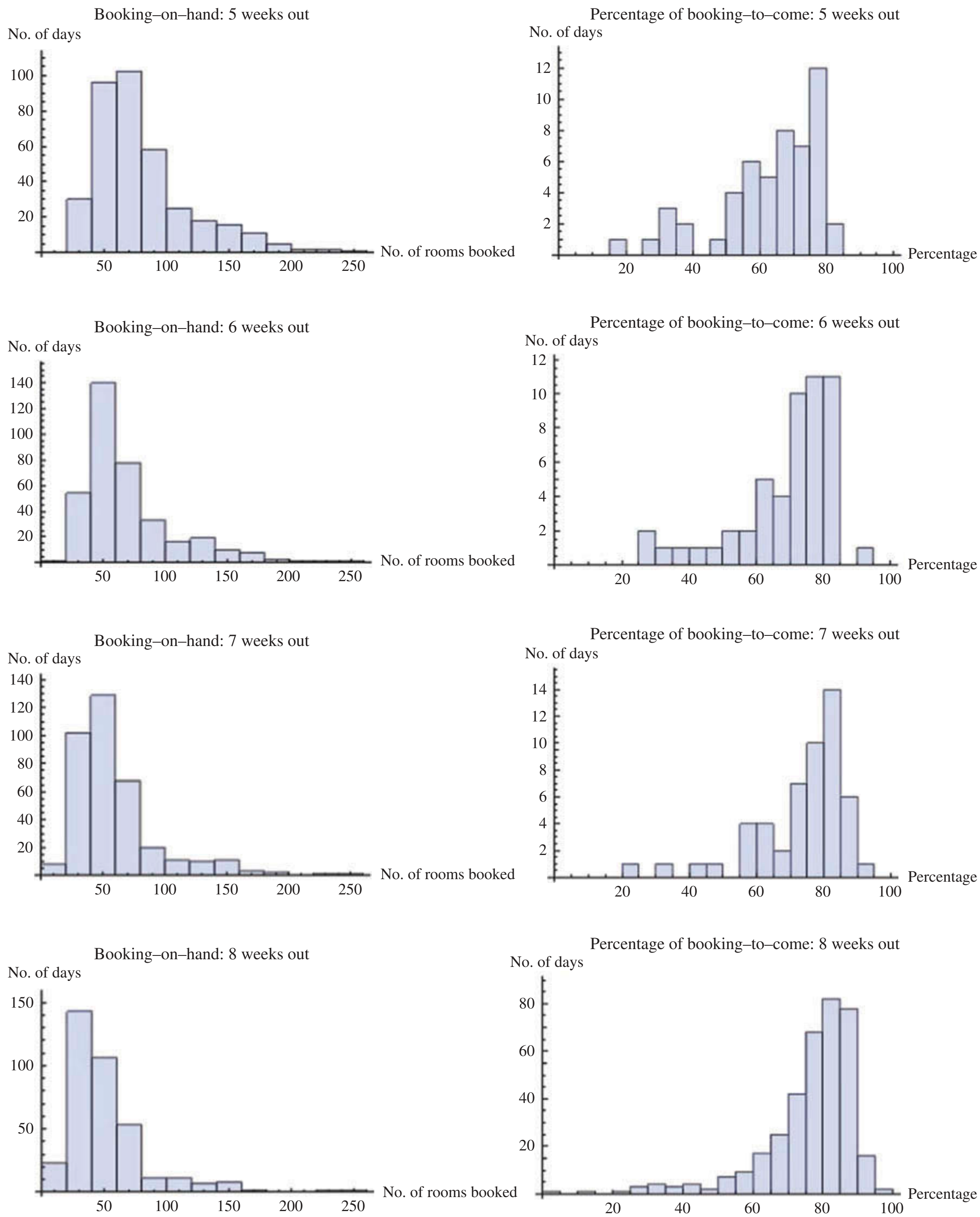


FIGURE 1. Continued

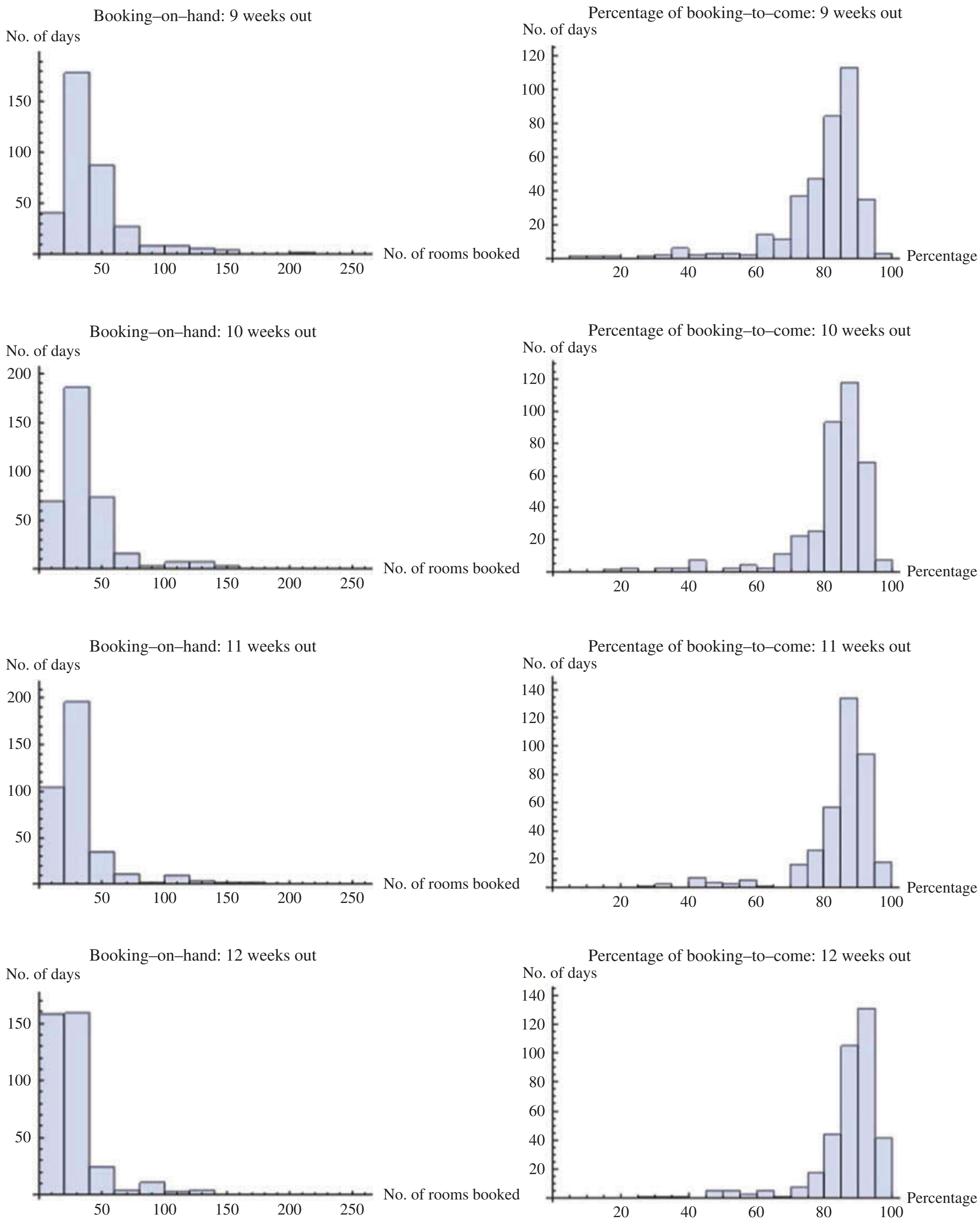
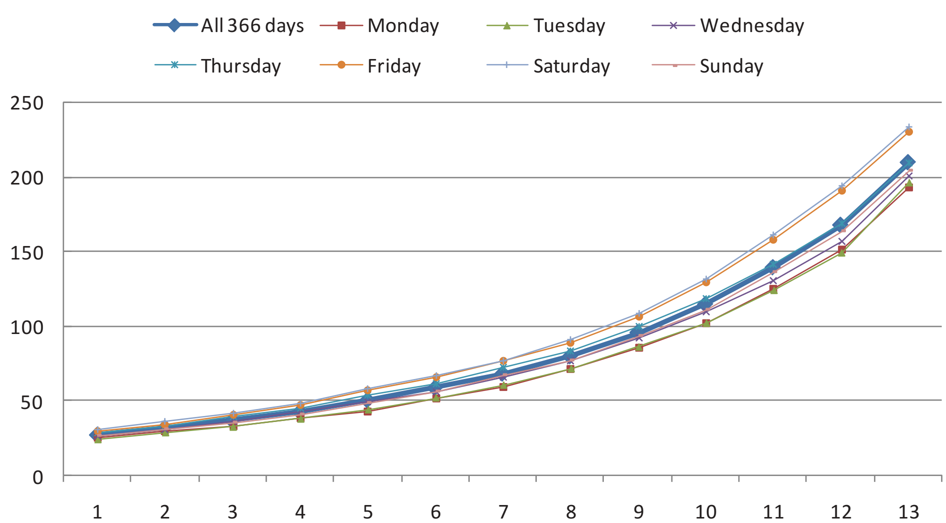


TABLE 2. Distribution of Bookings-On-Hand and Bookings-To-Come for 366 days

Days before Day 0	Bookings-on-hand		Percentage of Bookings-to-come	
	Mean	Standard deviation	Mean	Standard deviation
84	27.4	23.2	87.0	10.4
77	32.1	24.9	84.8	11.1
70	37.1	27.0	82.4	12.1
63	42.8	29.5	79.6	13.1
56	50.6	32.6	76.0	14.2
49	58.8	34.4	72.1	14.8
42	68.4	37.1	67.5	15.8
35	80.0	39.0	62.0	16.4
28	96.0	40.8	54.4	16.5
21	114.8	41.3	45.5	15.8
14	139.5	42.0	33.7	14.6
7	167.9	38.5	20.1	10.7
0	209.8	36.4	0	0

FIGURE 2. Booking Trends: Overall and By Day of Week



The findings that there are clear trends and patterns in the booking data lead us to question the common industry practice of using current bookings-on-hand and ignoring the booking data beforehand in forecasting. This practice assumes that future bookings on a particular day depend only on the current number of bookings-on-hand, not on how the bookings-on-hand were generated. Discussion with experienced marketing management professionals in the hotel industry reveals that the justification for ignoring past booking data in forecasting is that the *pickup rates at different times are different and thus the booking data in early days or weeks bear no relevance in forecasting*. We shall illustrate with an example below that this justification does not hold and, in fact, the booking data in earlier weeks can sometimes be very useful in forecasting.

Consider the 90-days bookings-on-hand data for October 3, 2011. We look at the increase of room reservations each week before Day 0

TABLE 3. Pickup of bookings for October 3, 2011

Days before Day 0	Week number	Bookings-on-hand	Increase of bookings from previous week
84	1	14	14
77	2	25	11
70	3	25	0
63	4	28	3
56	5	29	1
49	6	33	4
42	7	39	6
35	8	59	20
28	9	82	23
21	10	120	38
14	11	162	42
7	12	210	48
0	13	241	31

(Table 3), and plot them against the number of the week in Figure 3.

It is interesting to note that while the pickup rates in different weeks are indeed different, the data roughly follow a linear trend between weeks 5 and 12 (see Figure 3). Mathematically, it means that the bookings-on-hand at week t can be closely approximated by a quadratic function by regression on the booking data. Extending the regression to the full data set, we obtain a quadratic function of the form

$$Y(t) = at^2 + bt + c,$$

where Y(t) represents the forecast bookings-on-hand at (90 – t) days before Day 0, and a–c are coefficients obtained by regression.

Establishing the Booking Trend

Based on the above finding and using regression on the booking data for October 3, 2011, the best fit quadratic function is given by

$$Y(t) = 0.0487t^2 - 2.0295t + 34.315,$$

where Y(t) represents the forecast bookings-on-hand at (90 – t) days before Day 0, and we call this the booking trend approximation. Figure 4 shows that the booking trend approximation fits the entire bookings-on-hand data for October 3, 2011 quite well.

The fact that the booking trend approximation fits the entire bookings-on-hand data for October 3, 2011, quite well gives us the confidence to try to use the quadratic function for forecasting purposes.

Using the Booking Trend for Forecasting

Assuming that we are now 30 days before October 3, 2011, we are trying to forecast the final occupancy. We noted earlier from the distribution plots that the bookings during the first four weeks constitute only about 20% of the final occupancy, with a relatively large standard

FIGURE 3. Weekly Increase of Room Reservations for October 3, 2011

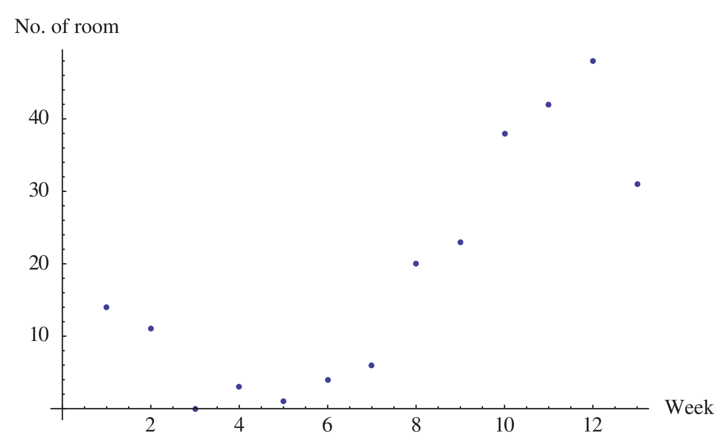
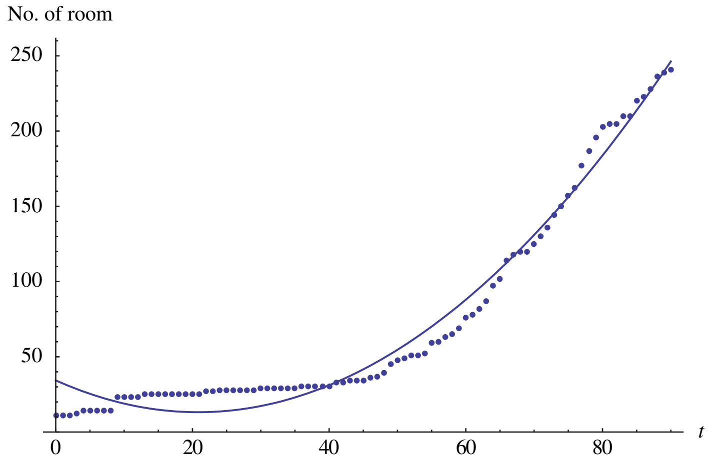


FIGURE 4. Bookings-On-Hand Data for October 3, 2011, and the Booking Trend Approximation



deviation, therefore the data in the first four weeks are not expected to behave consistently with the rest of the data. Hence, we try to derive another quadratic function with data from t = 30 to t = 60. We use t = 30 because we want to avoid the abnormality in the data in the first four weeks, and we use t = 60 because we are now (90 – t) or 30 days before Day 0. Applying this reasoning to the bookings-on-hand data for October 3, 2011, we have the quadratic function

$$Y(t) = 0.0696t^2 - 4.825t + 112.249,$$

which is shown in Figure 5.

In order to illustrate how well this quadratic function estimates the final occupancy, we plot the curve Y(t) together with the entire bookings-on-hand data set in Figure 6. The close approximation means that the quadratic function obtained from the data from t = 30 to t = 60 can provide a good forecast.

FIGURE 5. Booking Trend Approximation With Booking Data From t = 30 to t = 60 for October 3, 2011

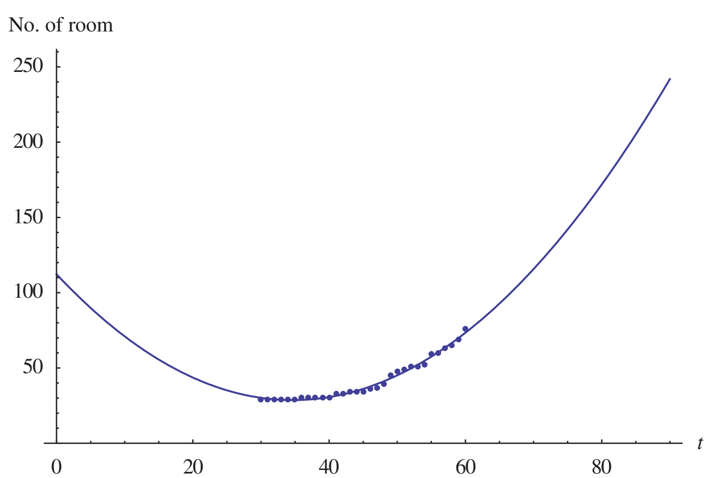
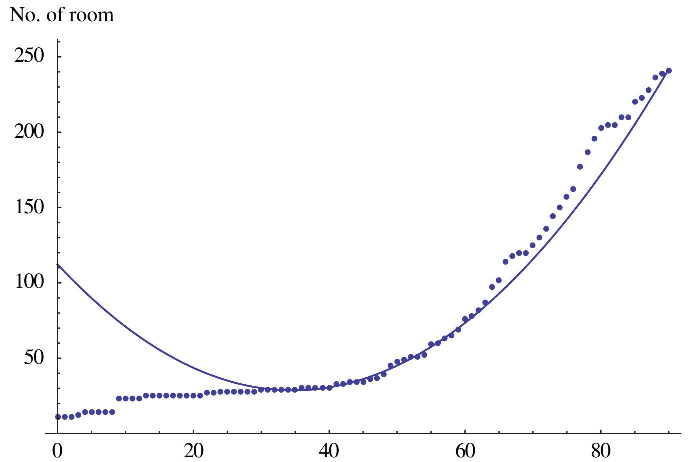


FIGURE 6. Entire Bookings-On-Hand and Booking Trend Approximation With Booking Data From t = 30 to t = 60 for October 3, 2011



We then repeat the regression procedure for each of the 366 days, using the booking data between t = 30 and t = 60, thus obtaining a quadratic function Y(t) for each day. For the purpose of forecasting, if the forecast value of Y(90) is greater than the room capacity (262) of the hotel, we will take the forecast to be 262.

Following such a procedure, we find that among the 366 days there are 79 days in which the forecast value is within 10% of the actual occupancy, and 39 days in which the forecast value is within 5% of the actual occupancy. Since one would expect the accuracy to increase if we use data closer to Day 0, we repeat the forecasting for booking data two weeks and one week out respectively. We also calculated the mean absolute percentage error (MAPE) for these forecasting horizons. The analysis confirms that the accuracy of the forecasting increases (or the error reduces) as booking data closer to Day 0 are used, and the findings are summarized in Table 4:

TABLE 4. Accuracy of Forecasting Using Different Data Sets

	Using data four weeks out t = 30 to t = 60	Using data two weeks out t = 46 to t = 76	Using data one week out t = 54 to t = 84
Error within 5%	30	94	131
Error within 10%	68	168	225
Mean absolute percentage error	33.9	15.3	9.7

We would like to point out that in the above regression procedure, care must be taken to avoid overfitting. Overfitting can occur if the degree of the fitting polynomial is too high or too much booking data are used in the regression. For example, Schwartz and Hiemstra (1997) applied regression with polynomials of degree four on booking data starting at 120 days before arrival. The results were mixed and unimpressive. To demonstrate the possible consequence of such overfitting, we calculate the MAPE of the regression with the following two models: (1) degree four polynomial using data from t = 54 to t = 84, and (2) degree two polynomial using data from t = 1 to t = 84.

1. Although a higher degree polynomial has more flexibility (it can go up and down more often) in fitting the data, it can also be easily affected by a few abnormal data, thereby affecting the forecast. The model with degree four polynomial, using data from t = 54 to t = 84, produces a MAPE equal to 18.1, almost twice as much as our model.
2. We have already observed that the booking data for the first few weeks account for a very small percentage of the total booking and have a larger variation. Trying to fit the curve to these data may also lead to overfitting. The model with quadratic polynomial, using data from t = 1 to t = 84,

TABLE 5. Accuracy of Forecasting Using Different Regressions and Data Set

Degree of regression polynomial	4	2	2
Booking data used	t = 54 to t = 84	t = 1 to t = 84	t = 54 to t = 84
Mean absolute percentage error	18.1	16	9.7

produces a MAPE equal to 16, about 65% higher than that of our model.

Table 5 summarizes the results found in (1) and (2) above.

It should be pointed out that the accuracy could be further improved if the outliers were removed from the data set. Judgment should be exercised when deciding whether to (1) remove the outliers altogether or (2) wait until booking data closer to Day 0 are available. One should not make forecasts based on a booking trend which looks unreasonable, such as a downward-sloping curve.

CONCLUSION

While advance bookings modeling is an accepted forecasting method in the hotel industry, there has been little research analyzing booking data in detail. This is partly due to the fact that hotel reservations data are confidential and difficult to obtain. This study analyzes the trends and patterns in the empirical booking data made available by the Hotel ICON, and explores how the findings can be used for forecasting final occupancy.

We first examined the data based on 90-days bookings-on-hand and bookings-to-come. The clear trends and patterns identified in the data set lead us to believe that we should make more and better use of the booking data despite the fact that the data are several weeks out and the pickup might vary from day to day and week to week. We question the industry practice of using only the current bookings-on-hand and ignoring previous booking data in forecasting. This common practice assumes that future bookings on a particular day depend only on the current number of bookings-on-hand, not on how the bookings-on-hand were generated. The analysis of the Hotel ICON booking data set shows that the assumption may not hold, and the trends and patterns identified in the booking data in earlier weeks can sometimes be very useful in forecasting. In fact, in many cases, the previous booking data is found to follow quite closely a booking trend which is identified as a quadratic function.

After identifying the booking trend as a quadratic function, we illustrate how the booking trend can be used for forecasting final occupancy with some degree of accuracy. Accuracy improves as we use booking data closer to the day for which we try to forecast the occupancy. An important point is that this forecasting method can be adopted by any hotel on the understanding that the regression function might not be quadratic in other situations and the booking window might vary. We also demonstrated the problem of overfitting in the regression model. Since the booking windows vary according to different sectors (HotelNewsNow.com, 2012) in the hotel industry and are shrinking fast (Mourier, 2011), one should explore the optimal data set interval to be used in the regression for specific markets and sectors.

Admittedly, the trends and patterns identified in the analysis are unique to a particular hotel and cannot be generalized. Nonetheless, the analysis demonstrates a more in-depth understanding of advance bookings modeling, and explores a forecasting method which is entirely based on bookings-on-hand data. This study shows that advance bookings data analysis gives hotel management a better appreciation of when business actually starts to book and helps them determine how far in advance of the day of arrival restrictions should be placed.

One advantage of using advance bookings data to forecast room demand is that the data are always available. In fact, in the case of a new hotel when there is little historical data and an absence of data going back several years, booking data such as 60 days out and 90 days out are the only reliable data available. It is up to the hotel to capture the reservations pickup data, conduct the analysis, and make its own forecast. Anderson and Xie (2010) stress the need for further development in the application of property-level RM and pricing, and this study, to a certain extent, has addressed this need.

With the convenience and competitive pricing information made available by the Internet, more consumers tend to wait and make their booking decisions at the last minute. This behavior has made booking windows narrower and the ability to read advance booking data accurately is more important than ever. The forecasting method explored in this study does not replace conventional methods but supplements them by contributing to the understanding of the changing behavior of booking data, particularly within a relatively narrow booking window. We intend to compare the new method with conventional methods in advance bookings modeling when more data are made available. In the future, we intend to study the behavior of advance booking data by room type and booking channel. We believe that analyzing the data by room type would help differentiate the behavior of different types of guest, and analyzing the data by booking channel would help to understand online distribution, which has become increasingly important.

FUNDING

The work described in this paper was fully supported by SHTM's Hotel ICON Research Fund of The Hong Kong Polytechnic University (Project Account Code: Z0G5).

REFERENCES

AHLA, 2006. *Revenue management*. American Hotel & Lodging Association. Retrieved February 1, 2013 from, http://www.ahla.com/uploadedFiles/AHLA/Members_Only/_Common/technology_primers_pdf/88119NEI02ENGE.pdf.

Anderson, C. K., & Xie, X. (2010). Improving hospitality industry sales: Twenty-Five years of revenue management. *Cornell Hospitality Quarterly*, *51*(1), 53–67. doi:10.1177/1938965509354697

Andrew, W. P., Cranage, D. A., & Lee, C. K. (1990). Forecasting hotel occupancy rates with time series models: An empirical analysis. *Journal of Hospitality and Tourism Research*, *14*(2), 173–181.

Badinelli, R. (2000). An optimal, dynamic policy for hotel yield management. *European Journal of Operational Research*, *121*, 476–503. doi:10.1016/S0377-2217(99)00046-6

Baker, T., & Collier, D. (1999). A comparative revenue analysis of hotel yield management heuristics. *Decision Sciences*, *30*(1), 239–263. doi:10.1111/deci.1999.30.issue-1

Baker, T., & Collier, D. (2003). The benefits of optimizing prices to manage demand in hotel revenue management systems. *Production and Operations Management*, *12*(4), 502–518. doi:10.1111/j.1937-5956.2003.tb00217.x

Bertsimas, D., & Boer, S. (2005). Simulation-based booking limits for airline revenue management. *Operations Research*, *53*(1), 90–106. doi:10.1287/opre.1040.0164

Bitran, G., & Mondschein, S. (1995). An application of yield management to the hotel industry considering multiple day stays. *Operations Research*, *43*(3), 427–443. doi:10.1287/opre.43.3.427

Chen, C., & Kachani, S. (2007). Forecasting and optimisation for hotel revenue management. *Journal of Revenue and Pricing Management*, *6*(3), 163–174. doi:10.1057/palgrave.rpm.5160082

Choi, T. Y., & Cho, V. (2000). Towards a knowledge discovery framework for yield management in the Hong Kong hotel industry. *International Journal of Hospitality Management*, *19*(1), 17–31. doi:10.1016/S0278-4319(99)00053-5

Cranage, D. (2003). Practical time series forecasting for the hospitality manager. *International Journal of Contemporary Hospitality Management*, *15*(2), 86–93. doi:10.1108/09596110310462931

Cross, R. (1997). Launching the revenue rocket: How revenue management can work for your business. *The Cornell Hotel and Restaurant Administration Quarterly*, *38*(2), 32–43.

Cross, R., Higbie, J., & Cross, D. (2009). Revenue management's renaissance: A rebirth of the art and science of profitable revenue generation. *Cornell Hospitality Quarterly*, *50*(1), 56–81. doi:10.1177/1938965508328716

Dacko, S. (2004). Marketing strategies for last-minute travel and tourism: Profitability and revenue management implications. *Journal of Travel & Tourism Marketing*, *16*(4), 7–20. doi:10.1300/J073v16n04_04

Haerian, L., Homem-de-Mello, T., & Mount-Campbell, C. (2006). Modeling revenue yield of reservation systems that use nested capacity protection strategies. *International Journal of Production Economics*, *104*, 340–353. doi:10.1016/j.ijpe.2006.02.006

Helsel, C., & Cullen, K. (2006). *Defining revenue management: Top line to bottom line*. McLean, VA:

Hospitality Sales and Marketing Association International.

HotelNewsNow.com. (2012, August 20). Revenue managers adapt to longer lead time. Retrieved from http://www.hotelnewsnow.com/articles.aspx/7918/Revenue-managers-adapt-to-longer-lead-times

Kiely, T. (2008). Revenue maximization: An examination of the influences of heuristics and biases on the yield management decision process in Dublin hotels. *TTRA Conference, 2008*, Helsinki, Finland.

Kimes, S. E. (1989). Yield management: A tool for capacity-considered service firms. *Journal of Operations Management*, *8*(4), 348–363. doi:10.1016/0272-6963(89)90035-1

Lai, -K.-K., & Ng, W.-L. (2005). A stochastic approach to hotel revenue optimization. *Computers & Operations Research*, *32*, 1059–1072. doi:10.1016/j.cor.2003.09.012

Law, R. (1998). Room occupancy rate forecasting: A neural network approach. *International Journal of Contemporary Hospitality Management*, *10*(6), 234–239. doi:10.1108/09596119810232301

Law, R. (2004). Initially testing an improved extrapolative hotel room occupancy rate forecasting technique. *Journal of Travel & Tourism Marketing*, *16*(2–3), 71–77. doi:10.1300/J073v16n02_06

Lee, A. O. (1990). Airline reservations forecasting: probabilistic and statistical models of the booking process (*Flight Transportation Laboratory Report* R90-5). Cambridge: Massachusetts Institute of Technology.

Littlewood, K. (1972). *Forecasting and control of passenger bookings*. Proceedings of the 12th Annual AGIFORS Symposium, Netanya, Israel.

Mourier, J. F. (2011). How to beat the incredible shrinking hotel booking window? Retrieved June 26, 2013, from http://www.tnooz.com/2011/09/19/how-to/how-to-beat-the-incredible-shrinking-hotel-booking-window/

Orkin, E. (1998). Wishful thinking and rocket science: The essential matter of calculating unconstrained demand for revenue management. *The Cornell Hotel and Restaurant Administration Quarterly*, *39*(4), 15–19. doi:10.1177/001088049803900404

Phumchusri, N., & Mongkolkul, P. (2012). *Hotel room demand forecasting via observed reservation information*. Proceedings of the Asia Pacific Industrial Engineering and Management System Conference 2012, Phuket.

Rajopadhye, M., Ghalia, M., Wang, P., Baker, T., & Eister, C. (2001). Forecasting uncertain hotel room demand. *Information Sciences*, *132*, 1–11. doi:10.1016/S0020-0255(00)00082-7

Relihan III., W. J. (1989). The yield-management approach to hotel-room pricing. *The Cornell Hotel and Restaurant Administration Quarterly*, *30*(1), 40–45. doi:10.1177/001088048903000113

Rest, J. P. I., & Harris, P. J. (2008). Optimal imperfect pricing decision-making: Modifying and applying Nash's rule in a service sector context. *International Journal of Hospitality Management*, *27*(2), 170–178. doi:10.1016/j.ijhm.2007.01.001

Rubicon. (2010). Recovery signs: booking pace. Retrieved August 16, 2012 from, http://www.hotelnewsnow.com/Articles.aspx/3428/Recovery-signs-Booking-pace

Sa, J. (1987). *Reservations forecasting in airline yield management* (MIT Flight Transportation Lab Report R87-1). Cambridge, MA.

Sanchez, J., & Satir, A. (2005). Hotel yield management using different reservation modes. *International Journal of Contemporary Hospitality Management*, *17*(2), 136–146. doi:10.1108/09596110510582332

Schwartz, Z. (2006). Advanced booking and revenue management: Room rates and the consumers' strategic zones. *International Journal of Hospitality Management*, *25*(3), 447–462. doi:10.1016/j.ijhm.2005.02.002

Schwartz, Z. (2008). Time, price, and advanced booking of hotel rooms. *International Journal of Hospitality & Tourism Administration*, *9*(2), 128–146. doi:10.1080/15256480801907885

Schwartz, Z., & Cohen, E. (2004). Subjective estimates of occupancy forecast uncertainty by hotel revenue managers. *Journal of Travel & Tourism Marketing*, *16*(4), 59–66. doi:10.1300/J073v16n04_08

Schwartz, Z., & Hiemstra, S. (1997). Improving the accuracy of hotel reservations forecasting: Curves similarity approach. *Journal of Travel Research*, *36*(1), 3–14. doi:10.1177/004728759703600102

Talluri, K. T., & Ryzin, G. J. V. (2005). *The theory and practice of revenue management*. New York, NY: Springer.

Tse, T. (2012). The experience of creating a teaching hotel: A case study of hotel ICON in Hong Kong. *Journal of Hospitality & Tourism Education*, *24*(1), 17–25. doi:10.1080/10963758.2012.10696658

Upchurch, R., Ellis, T., & Seo, J. (2004). Applying the hierarchical cluster analysis procedure upon the process of yield management – a comparative study. *Journal of Travel & Tourism Marketing*, *16*(4), 47–58. doi:10.1300/J073v16n04_07

Weatherford, L., & Kimes, S. (2003). A comparison of forecasting methods for hotel revenue management. *International Journal of Forecasting*, *19*(3), 401–415. doi:10.1016/S0169-2070(02)00011-0

Weatherford, L., Kimes, S., & Scott, D. (2001). Forecasting for hotel revenue management: Testing aggregation against disaggregation. *The Cornell Hotel and Restaurant Administration Quarterly*, *42*(4), 53–64. doi:10.1016/S0010-8804(01)80045-8

Yuksel, S. (2007). An integrated forecasting approach to hotel demand. *Mathematical and Computer Modelling*, *46*(7–8), 1063–1070. doi:10.1016/j.mcm.2007.03.008

Zakhary, A., Gayar, N., & Atiya, A. (2008). A comparative study of the pickup method and its variations using a simulated hotel reservation data. *The International Journal of Artificial Intelligence and Machine Learning*, *8*(special issue), 15–21.

SUBMITTED: July 18, 2013
FINAL REVISION SUBMITTED: June 19, 2014
ACCEPTED: June 26, 2014
REFEREED ANONYMOUSLY

APPENDIX

Bookings-on-hand and bookings-to-come for Monday

Days before Day 0	Bookings-on-hand		Percentage of bookings-to-come	
	Mean	Standard deviation	Mean (%)	Standard deviation
84	25.3	20.9	86.9	10.0
77	29.4	21.8	84.8	10.4
70	33.1	23.6	82.8	11.6
63	38.0	24.9	80.2	12.4
56	43.3	25.6	77.4	12.7
49	51.6	27.7	73.1	14.0
42	59.7	29.7	68.7	15.1
35	71.4	31.9	62.9	15.6
28	86.0	34.6	55.5	16.1
21	102.0	35.7	47.2	15.5
14	125.1	37.2	35.3	14.5
7	151.3	34.6	21.6	11.3
0	193.2	35.0	0	0

Bookings-on-hand and bookings-to-come for Tuesday

Days before Day 0	Bookings-on-hand		Percentage of bookings-to-come	
	Mean	Standard deviation	Mean (%)	Standard deviation
84	24.6	22.0	87.4	10.1
77	29.2	24.2	85.1	11.2
70	32.9	25.6	83.2	12.1
63	38.3	27.6	80.4	12.8
56	44.5	29.5	77.2	13.7
49	52.4	32.4	73.1	15.0
42	60.9	35.5	68.7	16.3
35	71.8	36.8	63.2	16.5
28	86.3	39. 1	55.9	16.8
21	102.2	39.8	47.6	16.3
14	124.1	41.5	36.6	14.9
7	149.1	36.9	23.6	10.8
0	195.9	41.7	0	0

Bookings-on-hand and bookings-to-come for Wednesday

Days before Day 0	Bookings-on-hand		Percentage of bookings-to-come	
	Mean	Standard deviation	Mean (%)	Standard deviation
84	26.2	20.8	86.9	9.5
77	31.1	23.9	84.6	10.9
70	35.9	25.1	82.1	11.5
63	42.2	27.5	79.0	12.4
56	48.9	30.6	75.7	13.6
49	56.3	32.3	72.0	14.0
42	66.0	36.3	67.2	15.4
35	76.6	37.9	61.9	16.0
28	92.1	40.3	54.3	16.3
21	109.1	40.9	45.7	15.7
14	130.4	42.7	34.9	15.2
7	156.3	37.3	21.7	11.4
0	200.4	40.0	0	0

Bookings-on-hand and bookings-to-come for Thursday

Days before Day 0	Bookings-on-hand		Percentage of bookings-to-come	
	Mean	Standard deviation	Mean (%)	Standard deviation
84	28.4	22.5	86.5	10.1
77	33.3	24.6	84.2	10.9
70	39.2	26.6	81.3	11.8
63	44.7	29.0	78.9	12.5
56	53.5	33.4	74.7	14.3
49	61.8	36.2	70.7	15.0
42	71.8	40.2	66.0	16.3
35	83.8	42.0	60.4	16.7
28	99.6	44.0	53.0	17.0
21	118.0	43.7	44.0	16.5
14	141.6	43.6	32.9	15.2
7	168.9	37.6	19.5	10.9
0	209.6	34.7	0	0

Bookings-on-hand and bookings-to-come for Saturday

Days before Day 0	Bookings-on-hand		Percentage of bookings-to-come	
	Mean	Standard deviation	Mean (%)	Standard deviation
84	30.5	27.1	87.0	11.4
77	35.5	28.3	84.9	11.9
70	41.7	31.5	82.2	13.3
63	48.3	34.7	79.4	14.6
56	58.3	38.4	75.2	16.0
49	66.5	39.5	71.7	16.5
42	77.0	41.0	67.1	17.0
35	90.5	42.6	61.3	17.7
28	108.6	43.1	53.6	17.8
21	131.5	40.7	43.8	16.2
14	161.0	37.9	31.3	14.2
7	194.1	30.5	17.1	9.9
0	234.0	21.5	0	0

Bookings-on-hand and bookings-to-come for Friday

Days before Day 0	Bookings-on-hand		Percentage of bookings-to-come	
	Mean	Standard deviation	Mean (%)	Standard deviation
84	29.3	25.3	87.3	10.8
77	34.2	26.4	85.1	11.2
70	40.1	29.4	82.6	12.5
63	46.8	33.0	79.8	13.8
56	56.5	37.7	75.7	15.4
49	65.5	38.7	71.7	15.6
42	76.2	40.3	67.1	16.0
35	88.7	42.8	61.7	16.7
28	106.5	43.7	54.0	16.9
21	129.1	42.9	44.1	16.2
14	157.3	41.6	31.9	14.7
7	191.2	35.1	17.1	9.9
0	230.3	27.9	0	0

Bookings-on-hand and bookings-to-come for Sunday

Days before Day 0	Bookings-on-hand		Percentage of bookings-to-come	
	Mean	Standard deviation	Mean (%)	Standard deviation
84	26.9	24.2	86.9	11.1
77	31.5	25.5	84.7	11.6
70	36.2	27.4	82.4	12.6
63	41.6	29.2	79.7	13.6
56	49.3	30.9	76.0	14.0
49	57.0	32.3	72.3	14.6
42	66.9	34.3	67.4	15.7
35	77.5	36.3	62.2	16.3
28	93.2	36.3	54.6	15.9
21	111.0	36.6	45.9	15.2
14	136.6	35.7	33.3	13.5
7	164.0	32.3	19.8	9.6
0	204.3	29.5	0	0