Tub/G 12/12/75

A Forecasting Model For

Hotel Occupancy Rates

by

James R. Freeland

Graduate School of Business
Stanford University

Research Paper No. 291

November, 1975

The author acknowledges the help of Joseph Waddell who supplied some of the data and Max Hall and Gary Koenig who wrote the computer program.

This research was supported in part by a grant from the Atlantic Richfield Foundation.

Abstract

Although in 1973 the average nationwide occupancy rate for hotels and motels was 61%, this rate fluctuates widely based on the time of year and the specific hotel or motel. The purpose of this paper is to present a short range forecasting procedure for occupancy rates which (1) is simple to use; (2) does not require significant computer time or storage; (3) adapts to changes, and (4) is easy to maintain and update. The model developed is an extension of Winter's three factor model to handle two different kinds of seasonality. The proposed model is tested using both hypothetical and actual data. A discussion of how the model output could be used in setting room rates for convention and tourist groups is also presented.

Introduction

An important component of the leisure industry is the hotel and motel business. In 1972 a study of 800 hotels and motels showed a total income and store rental of 2.1 billion dollars with 233,000 available rooms.[1] Although nationwide the average occupancy rate for 1973 was approximately 61%, this rate fluctuates widely based on the time of the year and the specific hotel or motel. The purpose of this study was to develop and test a quantitative forecasting technique for room occupancy rates. The specific model used was a modification of the Winter's three factor model [4]. The modified model was then tested on two sets of data: a hypothetical set and an actual set from a hotel in the Washington, D.C. area.

Nature of the Problem

Most hotels and motels could classify their clientele into one of four categories:

(1) Commercial accounts
(2) Convention
(3) Tour business
(4) Walk-ins

Commercial accounts are those travelers who are employees or guests of commercial firms or government agencies which have agreed to book a minimum number of nights per year in the hotel/motel. In return the hotel/motel agrees to give discounts. The discounted amount is part of marketing policy, but typically it varies between 33-1/3% and 10% off the average rack rate.[2]

---

[1] Based on data contained in [1].

[2] Rack rate is the undiscounted room rate which is normally quoted by the hotel/motel.

Commercial firms are generally approached through the hotel/motel travel desks and through personal calls on executives. The agreements usually cover a period of the next year.

Convention business accounts for a large percentage of hotel business in many metropolitan cities, e.g., San Francisco, Atlanta, etc. Competition among hotels for conventions, particularly large ones, is fierce. The site for a convention is often made a year or more in advance. Many hotels have a full time staff which spends their entire effort on landing conventions. The effort often involves a great deal of "wheeling and dealing" over room rates, facilities provided, fully "comped" (free) inspection visits for those persons making the convention site decision, etc.

Tour business consists of individuals or groups who have some special interest around which a trip might be focused. Travel agents generate most of the tour business. An agent typically contacts a hotel about reserving a block of rooms for the tour group. After some negotiation the agent agrees to buy a certain number of rooms on specific dates at a rate generally below the rack rate. The agent receives his commission in the form of a free room for a certain number of reservations. Tour business reservations are usually made from a few weeks to several months in advance.

Walk-ins include that portion of clientele both business and pleasure which does not fit in the other categories. Walk-ins include that business which makes reservations through a reservation system, travel

Agent or mail. Walk-ins reserve rooms anywhere from a few months to a few minutes ahead. In general, the most likely case for a walk-in is someone without a reservation. Walk-ins, with very few exceptions, pay the full rack rate.

Of course, the percentages of customers from each class will vary across different hotels and motels. For example, a motel located along an interstate will probably have close to 100% walk-ins. On the other hand, a large hotel in downtown San Francisco averages about 10% commercial accounts, 30% conventions, 40% tour business and 35% walk-ins. In addition, for a specific hotel or motel the percentages may vary from month to month.

There are two main reasons that a forecasting model for occupancy rates could be useful:

1) to assist in short range hiring and lay-off decisions, and
2) to provide information for negotiations over price with conventions and tour business.

Because of the nature of hotel/motel business it is normally accepted that there will be fluctuations in the work force size.[3] Typically the work force can be increased with a short (a few days to a couple of weeks) notice. Likewise, the work force can be reduced with a few days notice.

In regards to the second reason many hotels and motels are fully occupied a significant portion of the year. If a hotel can reasonably be expected to be nearly sold out, then giving a discount to a convention or tour group should be examined more carefully. If a hotel could predict its walk-in and commercial accounts business, a rational pricing policy could be pursued in negotiations with conventions and tour operators. The example below illustrates this issue.

<u>Example</u>

Suppose two groups are interested in booking 40 rooms each in a 400 room hotel during a particular week. Group 1 wants to check in on Friday

---

[3] Typically, labor unions for hotel workers also recognize that this fluctuation is a characteristic of reality.

afternoon and out on Monday morning. Group 2 plans to arrive Monday afternoon and leave Friday morning. Pertinent cost and demand data appears in Tables 1 and 2.

	Conventioneer Room	Walk-in Room
Room Rate	X	$40
Variable Rental Cost	$5	$5
Contribution to Profit from Beverage and Food	$5	$2

Table 1: Cost Data for Example

Day	Expected Walk-Ins
Friday	250
Saturday	190
Sunday	290
Monday	300
Tuesday	340
Wednesday	330
Thursday	300

Table 2: Demand Data for Example

The expected profit without the group's meetings is

Group			
1	37(250 + 190 + 290)	=	$27,010
2	37(300 + 340 + 330 + 300)	=	$46,990

The room rate per night for group 1 which would yield the same expected profit as without the group is given by solving the below equation for x:

$$27010 = 3(140x) + 37\left[\sum_{i=1}^{3}\left\{260\int_{260}^{\infty} f(y_i)\,dy_i + \int_{-\infty}^{260} y_i f(y_i)\,dy_i\right\}\right]$$

where $f(y_i)$ is the density of a normal distribution for the random variable room demand on night i (i=1 is Friday, i=2 is Saturday, etc.). It is assumed that the mean of the room demand is given in Table 2 and the standard deviation for each night is equal to 30. For group 1, x = \$3.54.

Similarly the room rate for group 2 per night is \$15.37.

The particular data used is hypothetical, but does represent reasonable values. The point is that a rational pricing strategy depends on being able to forecast non-convention sales of room rates.

The Forecasting Model

Figure 1 shows the actual demand history for a hotel in Washington, D.C. on Sunday for 1969 and 1970. An analysis of the complete data set indicates that there is no apparent trend in the data, but there is definitely a seasonal pattern. Closer analysis reveals that there is both a day of the week seasonal effect and a week of the year seasonal effect.

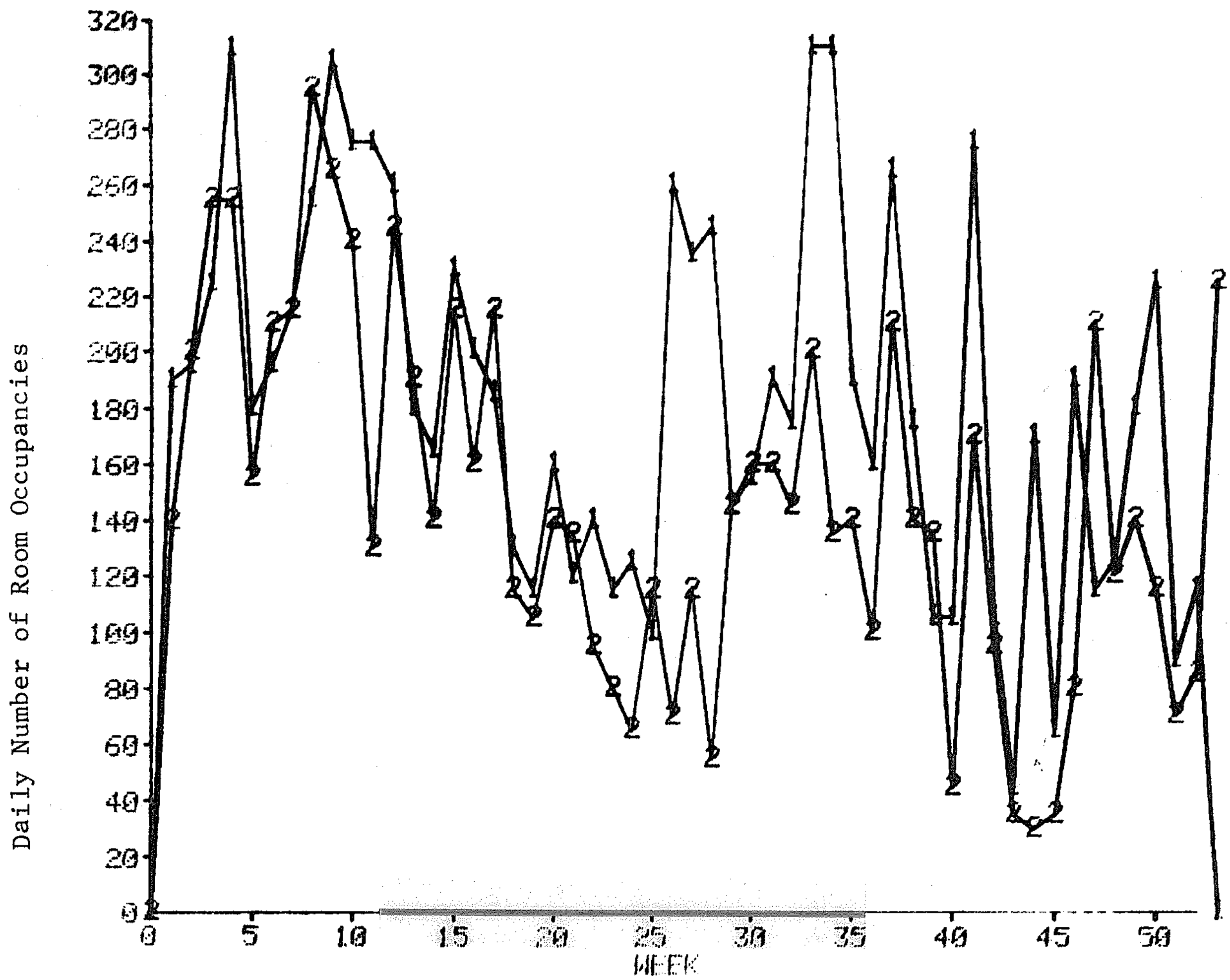

Figure 1: Hotel Occupancies on Sunday for 1969 (symbol 1) and 1970 (symbol 2)

Discussions with hotel and motel management people netted the following list of desired characteristics for a forecasting model to be useful:

(1) simple to use,

(2) not require significant computer time or storage,

(3) adaptive to changes

(4) easy to maintain and update.

A search of existing techniques with these characteristics plus the ability to handle two different seasonal factors yielded no candidates. Winter's [4] three factor model seems to possess the desired characteristics except it does not handle two different seasonalities. Therefore, it was decided to try to modify Winter's basic approach.

Winter's model is an exponential smoothing approach that accounts for three factors: a constant portion, a linear trend, and a seasonal cycle. While the model is heuristic in nature, it has intuitive appeal and has been used with considerable success [2]. Since the hotel data displayed no trend, the following model was hypothesized:

$$x_{ij} = a\, w_j\, d_i + \epsilon_{ij} \qquad (1)$$

where x_{ij} = demand for rooms on day i of week j, a is the base signal, w_j is the week of the year factor, d_i is the day of the week factor, and ϵ_{ij} is the random error component. Note that it is assumed that the seasonal factors are multiplicative. It is assumed that $\sum_{j=1}^{52} w_j = 52$ and $\sum_{i=1}^{7} d_i = 7$ because there will be 52 weeks in a year and 7 days in a week. Week 1, day 1 of each year is March 1. Thus, the 365th day is February 28 and during a leap year the 366^{th} day is February 29. To forecast for either of these two days the last week of the year factor, w_{52} is used with the appropriate day of the week factor.

The basic idea behind the model to be developed is to get estimates of a, w_j and d_i and then use equation (1) to forecast in the future. As actual occupancy data becomes available each of the model's factors are updated by smoothing the old estimate with a new one. The smoothing constants specify how rapidly the new estimates change with respect to new information.

Initially it is assumed that smoothing constants α_1, α_2, and α_3 are fixed, and estimates of the base signal, day of the week factors, and week of the year factors, denoted by $\hat{a}$, $\hat{d}_i$ (i=1,...,7), and $\hat{w}_j$ (k=1,...,52) respectively, exist. If a forecast of daily demand is desired for the week, h weeks in the future, and this week (h weeks in the future) is the j^{th} week of the year, then equation (2) below is used:

$$\hat{x}_{ij} = \min\{c, \hat{a}\, \hat{d}_i\, \hat{w}_j\} \text{ for } i=1,\ldots,7 \qquad (2)$$

where c = maximum number of rooms available.

Periodically, as the actual number of occupancies becomes known, the estimates of a, d_i and w_j are revised. For example, in the testing of the model in the next section $\hat{a}$ and $\hat{d}_i$ (i=1,...,7) are updated weekly while $\hat{w}_j$ is updated once per year. Specifically, after week t's actual demand becomes available, $\hat{a}$ is updated using (3):

$$\hat{a}^{\,new} = \alpha_1 \left[\sum_{i=1}^{7} (1 - \alpha_1)^{7-i} \left(\frac{x_{it}}{\hat{d}_i\, \hat{w}_t} \right) \right] + (1 - \alpha)^7 \hat{a} \qquad (3)$$

where x_{it} is the actual demand on day i of week t,

$\hat{a}$ = estimate of a before week t,

$\hat{a}^{\,new}$ = estimate of a after week t, and

$0 < \alpha_1 < 1$.

Note that (3) results from successive updating of $\hat{a}$ using

$$\hat{a}^{new} = \alpha_1\left(\frac{x_{it}}{\hat{d}_i\hat{w}_t}\right) + (1-\alpha_1)\hat{a}$$

Estimates of the day of week factors would then be revised using (4):

$$\hat{d}_i^{new} = \alpha_2\left(\frac{x_{it}}{\hat{a}_{new}\hat{w}_t}\right) + (1-\alpha_2)\hat{d}_i \qquad \text{for } i=1,\ldots,7 \qquad (4)$$

where $\hat{d}_i$ = estimate of d_i before week t,

$\hat{d}_i^{new}$ = estimate of d_i after week t, and

$0 < \alpha_2 < 1$.

Finally, the estimate of week of the year factor for week t is revised using (5):

$$\hat{w}_t^{new} = \alpha_3\left[\sum_{j=1}^{7}(1-\alpha_3)^j\left(\frac{x_{jt}}{\hat{a}_{new}\hat{d}_j^{new}}\right)\right] + (1-\alpha_3)^7\hat{w}_t \qquad (5)$$

where $\hat{w}_t$ = estimate of w_t before week t

$\hat{w}_t^{new}$ = estimate of w_t after week t.

Each week the $\hat{d}_i$ values are normalized so that $\sum_{i=1}^{7}\hat{d}_i = 7$, and at the end of each year the $\hat{w}_j$ values are normalized so that $\sum_{j=1}^{52}\hat{w}_j = 52$. Note, that each time a forecast is made the most current estimates of a, d_i, and w_j are utilized.

Testing the Model

A computer program was written which carries out all the calculations required. Figure 2 is a printout from a run. The model was then tested on two sets of data. In each case there were four years of data. The first year was used to initialize the a, d_i, and w_k factors. These factors and the smoothing constants, α_1, α_2 and α_3, which were input, were used to update the forecast parameters for the first two years. Forecasts were then prepared for the next two years for different forecast horizons. The actual occupancy levels were then compared to the forecast and errors were calculated. Figure 3 is a flow chart of the basic program.

```
HOTEL

# OF TOTAL YRS OF DATA?4
# OF YRS OF INITIALIZATION?2
DAY OF WEEK YR STARTS (1=SUN) AND YEARLY DATA FILE NAME:
  YEAR 1    ?1,H1
  YEAR 2    ?1,H2
  YEAR 3    ?1,H3
  YEAR 4    ?1,H4
YEARLY ERROR FILE NAMES:
  YEAR 3    ?ERROR1
  YEAR 4    ?ERROR2
WANT TO SAVE A, D, AND W FACTORS ON A FILE?NO
MAX # OF ROOMS?600
USING SAME INIT. DATA AS LAST RUN?NO
ALPHAS FOR A, D, AND W FACTORS?.05,.05,.05
FORECASTING HORIZON IN WEEKS?10
*******************************************
FOR FIRST 43   WEEKS
AVERAGE ERROR = -.22755     WITH STD. DEV. =  41.9672
****************************************************
RESULTS FOR YEAR 4
AVERAGE ERROR = -4.06891    WITH STD. DEV. =  40.6754
CUMULATIVE RESULTS
AVERAGE ERROR = -2.33019    WITH STD. DEV. =  41.2782
REPUN?NO

DONE
```

Figure 2: Computer Printout for Forecasting Procedure

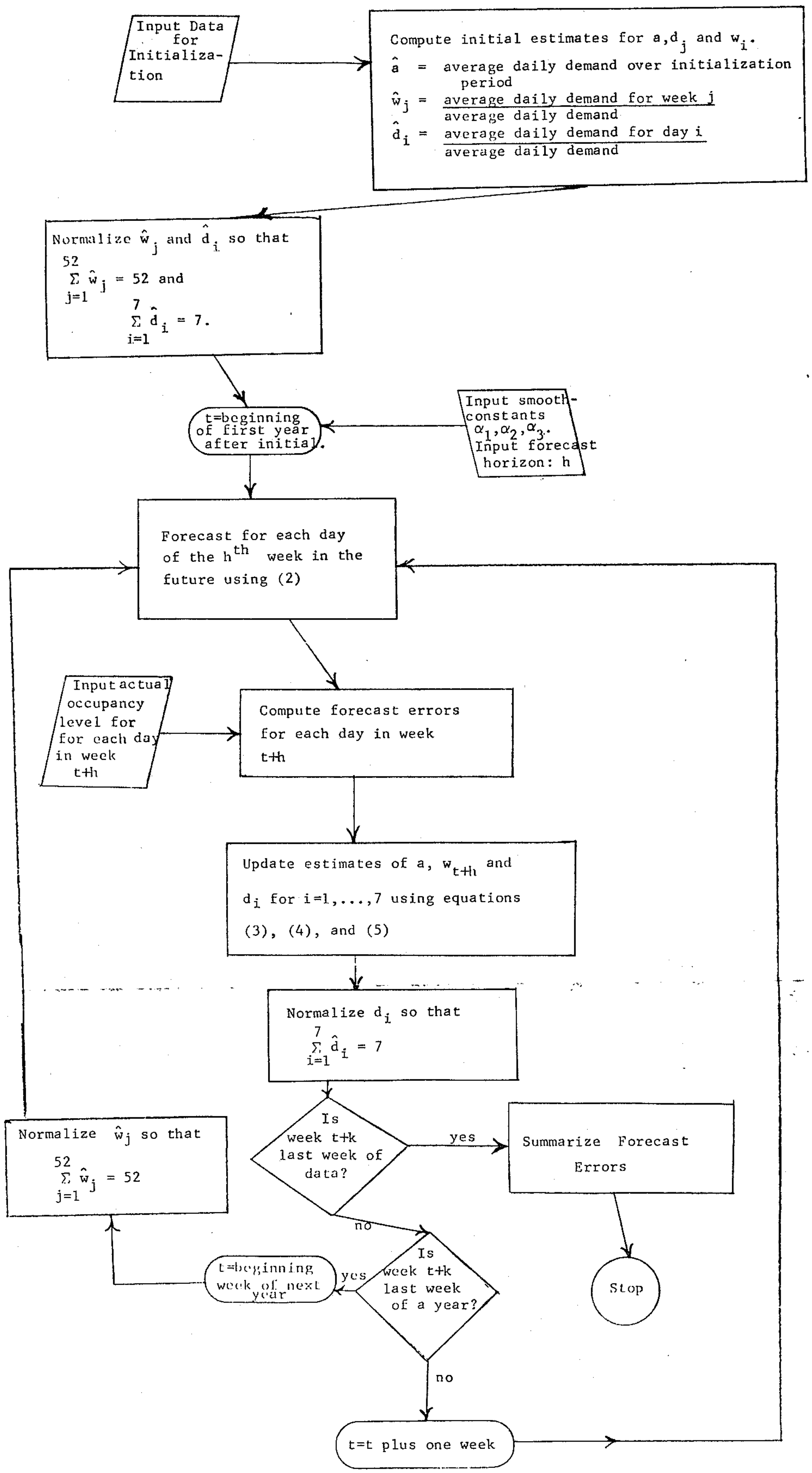

Figure 3: Flow Chart of Computer Program

The first set of data was generated by using the model given in (1) with an error term, ϵ_{ij}, normally distributed with mean equal to zero and standard deviation equal to 40 rooms. The base amount and the day of week factors were fixed at reasonable values and w_j was generated from a normal distribution with mean equal to 1.0 and standard deviation .15.

The second set of data was actual occupancy levels for a hotel in the Washington, D.C. area. The actual data had been tabluated as multiples of five so that some of the forecast error is due to this.

Different smoothing constant values were tried for each data set and for different forecast horizons. Tables 3 and 4 summarize the average error and the standard deviation of error for year 4 for both sets of data.

For the hypothetical data set the smallest standard deviation of errors for a time horizon of 10 weeks was found with smoothing constants for the base amount, weekly factors and day of week factors all set equal to .1. This value, 43.15 rooms, is quite good since the standard deviation of the underlying process with all seasonalities removed is 40. For a 20 and 40 week time horizon the smallest standard deviation of errors resulted with all smoothing constants equal to .1. The actual standard deviation values are again reasonable. It seems non-intuitive that the standard deviation decreases as the time horizon increases; however, the amount of difference between the standard deviation for 10 weeks and 40 weeks is not significant. Also, the underlying process for the hypothetical data set is very stable so that if the estimates of the models parameters are "good" then the error should not increase with a longer time horizon.

Table 3: Summary of Year 4 Forecast Errors for Data Set No. 1

Time Horizon (Weeks)	α_1	α_2	α_3	Mean	Error Standard Deviation
10	.1	.1	.1	-4.68	43.15
20				-3.14	42.45
40				-8.51	41.96
10	.1	.1	.3	-5.19	47.76
20				-3.49	46.00
40				-10.89	44.15
10	.1	.1	.5	-4.95	53.86
20				-4.33	53.02
40				-13.79	48.15
10	.1	.3	.1	-3.53	45.03
20				-3.30	44.71
40				-7.68	43.94
10	.1	.3	.3	-3.98	47.85
20				-3.55	46.98
40				-9.91	45.38
10	.1	.3	.5	03.84	51.66
20				-3.78	51.35
40				-12.37	48.29
10	.3	.1	.1	-6.21	48.61
20				-3.11	46.81
40				-9.47	47.83
10	.3	.1	.3	1.90	50.40
20				-.43	47.60
40				-10.69	50.84
10	.3	.1	.5	-6.32	60.16
20				-2.99	56.44
40					
10	.3	.3	.1	-5.12	50.12
20				-3.63	48.38
40				-11.98	54.53
10	.3	.3	.3	-5.22	54.06
20				-3.73	51.52
40				-9.67	51.98
10	.3	.3	.5	-4.96	59.07
20				-3.38	56.05
40				-11.04	55.19

Table 4: Summary of Year 4 Forecast Errors for Data Set No. 2

Time Horizon (Weeks)	α_1	α_2	α_3	Error Mean	Error Standard Deviation
10	.1	.1	.1	-6.17	58.80
20				-6.96	71.40
40				-25.99	81.34
10	.1	.1	.3	-5.06	65.58
20				-4.89	76.24
40				-26.84	88.65
10	.1	.1	.5	-1.99	72.38
20				-1.83	81.40
40				-27.23	97.51
10	.1	.3	.1	-4.67	61.39
20				-6.79	77.03
40					
10	.1	.3	.3	-2.87	66.49
20				-6.34	78.51
40				-27.26	90.75
10	.1	.3	.5	-.82	71.96
20				-4.31	81.54
40				-29.87	95.28
10	.3	.1	.1	-8.23	63.30
20				-7.88	69.70
40				-28.75	88.43
10	.3	.1	.3	-7.01	64.63
20				-6.51	77.32
40				-31.13	91.16
10	.3	.1	.5	-4.44	69.47
20				-4.11	80.98
40				-31.70	96.24
10	.3	.3	.1	-7.43	65.90
20				-8.97	79.72
40				-26.51	94.63
10	.3	.3	.3	-6.33	66.19
20				-8.61	79.55
40				-28.70	96.48
10	.3	.3	.5	-3.99	69.26
20				-6.46	83.25
40				-31.25	97.73

The primary conclusion from the study of the procedure's performance on the hypothetical data is that if the underlying process follows (1), then the proposed method forecasts quite well.

From Table 4 the smallest standard deviation for a 10 week time horizon at the hotel in the Washington, D.C. area was 58.8 rooms. A more exhaustive search over different values for the smoothing constants showed that when $\alpha_1 = \alpha_2 = \alpha_3 = .05$ a minimal standard deviation of errors (50.77 rooms) resulted. Given that the actual hotel has approximately 310 rooms, this level of precision is not overly impressive. However, there are some ameliorating factors. First, the occupancy data had been recorded in multiples of five rooms. Therefore, some amount of the error is due to the manner in which the data was recorded. In addition, not all the occupancies which were due to conventions and tour groups could be removed from the input data. This would decrease the effectiveness of the forecast procedure because the procedure is designed only for walk-in occupancies. Conventions and tour groups rent blocks of rooms and destroy the assumption of randomness called for in a short range forecasting procedure.

To gauge the accuracy of the forecasting procedure it was compared to a "naive" method. For a particular day of a given week the naive method computed a forecast by averaging the actual occupancies of the previous two years for the same day of week and week of the year. In this way the seasonalities are accounted for. At present the hotel was actually using a procedure similar to the naive method for estimating future occupancies. An analysis of forecast errors for the naive method showed an average error of -23.78 rooms and a standard deviation of errors of 102.97 rooms for the fourth year. These results would be the same for any horizon time of less than one year because it relies on the actual occupancy levels in the previous year.

In light of these results the proposed procedure appears more favorable. In addition, a more detailed analysis of the forecast errors for the proposed method showed very little autocorrelation between values for any lag of more than one day. Also, the errors seem to be normally distributed. This seems to suggest that the hypothesis of the underlying process is appropriate, and that there is an inherent high degree of variation in the demand pattern.

For the second data set the forecast errors increase with the horizon time. With a 40 week horizon time there is a very large bias in the forecasts. These results suggest that the proposed model may not be appropriate for time horizons longer than 20 weeks or so.

Summary

This paper has presented a short range forecasting procedure for estimating hotel room occupancy rates. The procedure uses a form of exponential smoothing to update estimates of seasonal factors for the day of the week and the week of the year. While the method is not likely to be as accurate or to predict changes in the demand pattern as an econometric forecasting model, the procedure has several advantages. It is relatively simple to use, little time and cost is involved in setting up and using the procedure, and it does adapt (at a rate chosen by the user) to shifts in parameters of the demand process.

The proposed forecasting procedure was tested with a hypothetical time series and a time series from a hotel in the Washington, D.C. area. For the actual hotel occupancy data the procedure performed significantly better than a simple procedure similar to what is currently being used. The tests indicated that forecast accuracy decreases as the time horizon increases, so that the procedure is not likely to be useful in those situations where conventions or tour groups are booked far in advance.

As with any model the output is only as good as the input. It is important that the input is the actual demand for hotel rooms from walk-in clientele. Historical records often show only the actual number of occupancies and when all rooms are occupied the true demand is often hard to get.

References

1. Harris, Kerr, Forster & Company, "Trends in the Hotel/Motel Business," 1973 Edition.

2. Groff, Gene K., "Empirical Comparison of Models for Short Range Forecasting," Management Science, Volume 20, Number 1, (September, 1973), pp. 22-31.

3. Winter , Peter R., "Forecasting Sales by Exponentially Weighted Moving Averages," Management Science, Volume 6, Number 3, (April, 1960), pp. 324