Forecasting Daily Accommodation Occupancy for Supply Preparation by a Sharing Economy Platform

Ta-Chun Kuan, Sz-Wei Wu, Cheng-Che Liao, Mahsa Ashouri, Galit Shmueli, and Che Lin
National Tsing Hua University
Hsinchu 30013, Taiwan

***Abstract*—Sharing economy platforms for accommodation sharing offer a more flexible supply-demand model compared to traditional hotels that own a fixed number of rooms. However, this flexibility can cause uncertainty and become a challenge if not managed dynamically. An important task for such platforms is forecasting future daily occupancy in a certain area, with sufficient lead time so that they can reach out to new hosts and secure more rooms in time for peak demand. We developed such a forecasting solution for AsiaYo, the largest Chinese language online accommodation sharing platform. We evaluate and compare various forecasting algorithms, including statistical and machine learning methods, using two years of data from AsiaYo on occupancy in different cities. The empirical results show that the occupancy is highly dependent on the weekday, city, and holidays. We show the strengths and weaknesses of different methods in terms of required accuracy level, computation time, and flexibility.**

I. Introduction

Accommodation sharing platforms such as AirBnB have been growing in popularity globally. In contrast to traditional hotels, the platforms do not own or run the accommodation, but rather help facilitate between B&B-type property owners and guests. AsiaYo is a start-up that provides an online platform for travelers bounded to Asia to select their desirable place to stay. As a third party OTA (Online Travel Agency) platform, they do not possess rooms but instead must collaborate with hosts. AsiaYo competes with other similar OTAs to provide optional channels for hosts reaching out to more customers.

Due to fluctuations in daily demand in different locations, the platform often experiences accommodation shortages, which might lead to lost revenues and potential guest dissatisfaction. It is therefore useful if they can forecast such shortages with sufficient lead time. We address this challenge by developing a forecasting system that provides daily occupancy forecasts in different cities, with a lead time of 28 days, thereby allowing AsiaYo to reach out to new potential B&B owners to increase supply on dates with forecasted shortages.

More formally, our goal is to forecast 28-days ahead daily room occupancy (F_{t+28}) in each city, using historical bookings data. We use a roll-forward approach, where models are updated daily with the new data. Figure 1 illustrates this process.

This work was partially supported by Taiwan Ministry of Science and Technology, research grants 106-3114-E-007-007 and 107-2218-E-007-045.

Our paper contributes to the literature on occupancy demand forecasting by evaluating and comparing a variety of time series forecasting methods in the new realm of shared economy time series. Such data differ from more traditional hotel occupancy data and are affected by a variety of new sources of heterogeneity. We consider methods that range in terms of parametric-to-nonparametric, computational intensity, scalability, and interpretability. We show the strengths and weaknesses of different methods in terms of required accuracy level, computation time, and flexibility.

The paper proceeds as follows: Section II describes related research on forecasting tourism demand. Section III introduces the data from AsiaYo and the pre-processing steps we have taken. Section IV describes four types of time series forecasting methods, explaining the different variations we considered. We describe and discuss the results in Section V. Section VI concludes and considers limitations and future directions.

II. Forecasting Tourism Demand Literature

Time series forecasting is extremely popular in forecasting tourism demand. Witt and Witt focused on forecasting tourism demand, and argued that methods like spatial models and time series models are the ones mostly used by researchers for tourist demand forecasting [1]. Various data are used for such purpose, trying to capture information as early as possible. Khadivi and Ramakrishnan used Wikipedia Usage Trends (WUTs) to forecast tourism demand at the trip planning stage and showed that using WUT time series improve the accuracy of tourism demand forecasts [2]. Tourism can be domestic or international. Mamula forecast the international tourist flows from Germany to Croatia and noted the importance of modeling and forecasting tourism demand and all its components, especially when there is a lack of systematic research on the international tourism markets [3]. Tourism demand forecasting is often based on data from hotels. Weatherford and Kimes

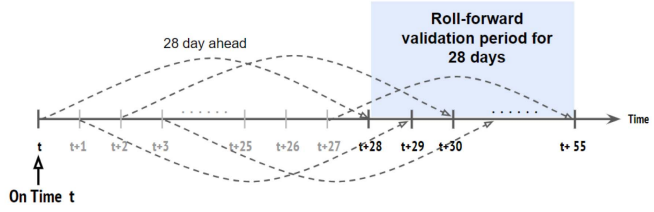

Fig. 1. Forecasting daily occupancy with 28-day lead time

978-1-7281-4117-6/19/$31.00 ©2019 IEEE
DOI 10.1109/DataCom.2019.00030

collected data on eight hotel rate categories, seven length of stay categories and two hotels and applied seven different methods of forecasting on their data to find the most accurate procedure with the lowest error for forecasting the number of arrivals [4]. They argued that the pickup method and regression procedures are the best models. The pickup method is one of the most common methods used in reservation-based data. Its main idea is to estimate the increments of bookings (to come) and then aggregate these increments to obtain a forecast of total demand to come [5]. Using demand data for five star hotels in Ankara, Turkey, Yuksel [6] introduced a novel way to include intangibles and human judgment in additional to numerical data for decision-making. He showed that applying this forecasting and adjustment process to monthly hotel data produces demand forecasts that help management avoid crises arising from demand fluctuations. Our study is situated in the tourism forecasting literature, but the sharing economy context makes it different because B&B room supply is dynamic over time, variable across cities, and more uncertain compared to occupancy at a specific hotel or aggregated in an entire country. Given the literature and these challenges, we consider a variety of popular time series forecasting methods and algorithms, including exponential smoothing, linear regression, ARIMA, and neural networks (see Section IV).

III. Data Description and Pre-Processing

AsiaYo provided us with two years of bookings data, with orders starting from January 1, 2016 to December 31, 2017. Each booking includes a booking ID, time stamp of booking, and information about the B&B location (country and city), check-in date and check-out date. We first transformed the transactions data by converting the check-in and check-out information into daily time series of total room occupancy in each city (see Figure 2). The series all exhibit strong day-of-week seasonality, but different cities have different magnitudes, trends, auto-correlation, and extreme peaks.

A. Data Exploration

Comparing length-of-stay and advance booking statistics in the different cities reveals significant differences between stays in Taiwan and in Japan. Since the data does not have the guests' location information, based on conversations with AsiaYo we made the assumption that most stays in Taiwan are domestic and most stays in Japan are international. Under this assumption, it is reasonable that bookings in Taiwan cities are typically for shorter periods and with shorter advance booking duration (domestic tourism) compared to Japan (international tourism).

B. Derived Variables

We created several derived variables to capture some of the observed patterns: day-of-week indicators (DOW) and Holiday indicators are based on information extracted from the date and external calendars. We also created lagged occupancy predictors (Occ_{t-k}), reflecting total bookings for k days before the day of prediction t. Another important derived variable is

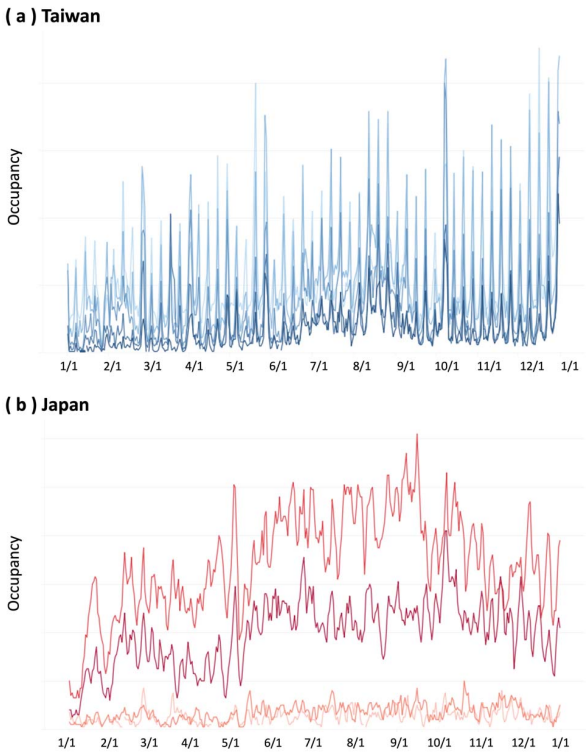

Fig. 2. Daily occupancy in different cities in Taiwan (a) and Japan (b). All series exhibit strong day-of-week seasonality, but different magnitudes, trends, autocorrelation, and peaks.

the 28-day ahead cumulative occupancy for day $t+28$ on the day of forecasting t ($Cum28_t$), which reflects "total bookings for stays on day $t+28$, 28 days in advance". For example, if on May 1 (t) we are forecasting occupancy on May 29 (Occ_{t+28}), then $Cum28_t$ is the total bookings for stays on May 29 as of May 1. In contrast, Occ_t is the total bookings for stays on May 1.

All these predictors are available at the time of prediction t, which is 28 days prior to a given date. We developed our forecasting solution for the top 5 cities in Taiwan and top 4 cities in Japan.

C. Data Partitioning

All methods were trained on data from 2017/01/01 to 2017/11/06 and predictive performance was evaluated on the validation (holdout) period 2017/12/04 to 2017/12/31. The validation window was set to 28 days due to the need for 28 forecasts and the short overall series (which we needed for training). We set the seasonal naive method as our benchmark, where the naive forecast is the value from 28 days ago ($F_{t+28} = Occ_t$).

IV. Forecasting Methods

We examined four types of time series forecasting methods from the statistics and data mining approaches: exponential

smoothing, linear regression, auto-regressive integrated moving average (ARIMA) models, and neural networks. These are the most popular methods used in practice and in competitions such as the M-competitions [7], [8][1] We studied in depth multiple variations of linear regression models, due to their simplicity, explainability, scalability, and ability of incorporating external information [9], [10]. We benchmarked these methods against seasonal naive forecasts.

The following four forecasting methods were trained and validated by daily roll-forward approach with training period from 2017/1/1 to 2017/12/3 and validation period from 2017/12/4 to 2017/12/31.

A. Exponential Smoothing

Because daily occupancy data is highly seasonal, we consider Holt-Winter's exponential smoothing method that captures seasonality. This algorithm expands on simple exponential smoothing (SES), which is a weighted average of all the past values, with higher weight for recent values. The k-step ahead SES forecasting equation is given by:

$$F_{t+k} = \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + ... \quad (1)$$

where $0 \leq \alpha \leq 1$ is the learning rate. SES is useful in series that have only a level and error, but no trend or seasonality.

Holt-Winter's exponential smoothing captures local trend and local seasonality patterns by using three updating equations for Level, Trend, and Seasonality, respectively:

$$L_t = \alpha(y_t - S_{t-M}) + (1-\alpha)(L_{t-1} + T_{t-1}) \quad (2)$$
$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \quad (3)$$
$$S_t = \gamma(y_t - L_{t-1}) + (1-\gamma)S_{t-M} \quad (4)$$

where α, β, γ are smoothing parameters that control how fast the algorithm learns level, trend, and seasonal patterns over time, respectively. At time t, the k-step ahead forecast is generated by combining the estimates of these three components via the forecasting equation:

$$F_{t+k} = L_t + kT_t + S_{t+k-M} \quad (5)$$

The exponential smoothing family also includes algorithms that capture multiplicative trend and seasonality, algorithms that capture only trend, only seasonality, etc. [9]. We used an automated selection algorithm (function ets in R) that selects among the entire exponential smoothing family the most suitable algorithm for forecasting F_{t+28} for each city.

[1]In the M3 Competition [7], where 3003 time series were to be forecast, the two most popularly used methods by contestants were exponential smoothing and ARIMA. Some contestants ensembled a linear regression model with other methods such as naive forecasts (random walk). In the recent M4 Competition [8], with 100,000 series, several machine learning algorithms were used including neural nets, yet pure ML methods performed poorly, with none of them being more accurate than the combination benchmark. Of the 17 most accurate methods, 12 were combinations of mostly statistical approaches and one combined ML and statistical models.

B. Linear Regression Models

We modeled Occ_t as a function of the derived variables in a few different ways. The different approaches attempt to capture the strong day-of-week seasonality.

- **Basic Model:** Occ_{t+28} is the outcome variable and is modeled as a function of the predictors *DOW* (to capture day-of-week seasonality), *Holiday* (to capture holidays), running index $t = 1, 2, 3, \ldots$ (to capture a linear trend), and $Cum28_t$ (the cumulative bookings for day $t+28$ on the day of forecasting).
- **Lagged Model:** The Lagged model is similar to the basic model, with the addition of predictors that capture lagged values of cumulative occupancy in the previous 7 days ($Cum28_{t-1}, Cum28_{t-2}, \ldots, Cum28_{t-7}$) and lagged values of the Occupancy series in the previous 7 days ($Occ_{t-1}, Occ_{t-2}, \ldots, Occ_{t-7}$).
- **Differencing Model:** In this model, we first take lag-7 differences of the original Occupancy series ($D_t = Occ_t - Occ_{t-7}$) and then use a linear regression model for the outcome variable D_{t+28} as a function of all the predictor variables.
- **Separate Models:** An alternative way to handle the strong day-of-week effect is to train separate models. We therefore trained 3 linear regression models: one for Fridays, a second for Saturdays, and a third for the remaining days (Sundays-Thursdays). In each of these models, the outcome is Occ_{t+28} and the predictors are *Holiday*, t (trend), and cumulative occupancy on day t ($Cum28_t$).

C. Neural Network

We used the predictors in the above basic linear regression model as input predictors of a feed-forward neural network with the following attributes (using R's $nnetar$ function):

- Single hidden layer
- Automated selection of optimal number of non-seasonal lags (p); seasonal lags set to P=1 (default)
- $(p+P+1)/2$ nodes in the hidden layer (default)
- Sigmoid activation function
- Ensemble of 5 runs of the algorithm

D. ARIMA Models

A non-seasonal ARIMA(*p, d, q*) model is given by:

$$y_t' = c + \phi_1 y_{t-1}' + \phi_2 y_{t-2}' + ... + \phi_p y_{t-p}' + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \ldots + \theta_q\epsilon_{t-q} + \epsilon_t \quad (6)$$

where

- y_t' is the d-differenced Occ_t series at time t (D_t)
- ϵ_t is white noise at time t
- p = order of the autoregressive part
- d = degree of first differencing involved
- q = order of the moving average part

We also considered a seasonal-ARIMA model of the form ARIMA$(p,d,q)(P,D,Q)_m$.

We used the *auto.arima()* function in R, which uses a variation of the Hyndman-Khandakar algorithm [11] that combines

unit root tests, minimization of the AICc and MLE to select p, d, q, P, D, Q, and estimates the resulting ARIMA model.

V. Results and Discussion

To evaluate the performance of different algorithms, we examined two types of plots: (1) time plots of the actual vs. forecasted time series along with the time plots of the forecast errors series, and (2) side-by-side box plots of the forecast errors. We examined these separately for the training and validation (holdout) periods. Figure 3 displays these plots for City A in Taiwan (TW City A) and City A in Japan (JP City A).

We found that performance of different algorithms was better for different cities. For example, the neural net (R package *nnetar*) worked best for TW City A, providing the most precise forecasts both in the training and validation periods; In contrast, for JP City A the neural net performed best on the training period, but the linear regression (R package *lm*) performed better on the validation period (Figure 3).

Given the very strong day-of-week patterns, we examined several approaches for improving the linear regression models. One approach included day-of-week indicator variables, a second included lags, a third used differencing operations, and a fourth approach modeled Fridays and Saturdays separately from other days. Comparing these different approaches, the best results on training and validation periods were obtained by only using day-of-week indicator variables (basic), and those were comparable to modeling weekdays vs. Fridays vs. Saturdays separately (separate), although the separate models achieved lower extreme errors (Figure 4). We note that although the differencing model ("Diff") performed best on the training period, its performance on the validation period was highly variable and inferior (see bottom-most validation error line chart).

Finally, the above results were compared for all other top cities in Taiwan and Japan, resulting in similar conclusions (Figures 5 and 6): For Taiwan cities, neural net performed best, and linear regression was second best (but with larger extreme errors) on training and validation periods. For Japan cities, neural net and linear regression performed equally well on training and validation periods. An important difference between the two algorithms in all Taiwan cities is that linear regression tended to over-forecast in the validation period, whereas the neural net under-forecasted. This is important due to the different impact and cost of over- vs. under-forecasting.

Our results have one important limitation: Given the data available to us, we used 11 months (Jan-Nov) for training and the remaining month of December for validation. This limits the generalizability of our results in two ways: First, although we also examined performance on the training set, true holdout results are only available for December. December also poses a special challenge as a validation period, due to the New Year holiday that has unusually high peaks on Dec 29-30. Second, with only one year of data, it is very difficult to model occupancy on holidays, since each holiday seems to have its own magnitude and behavior. Hence, with more years

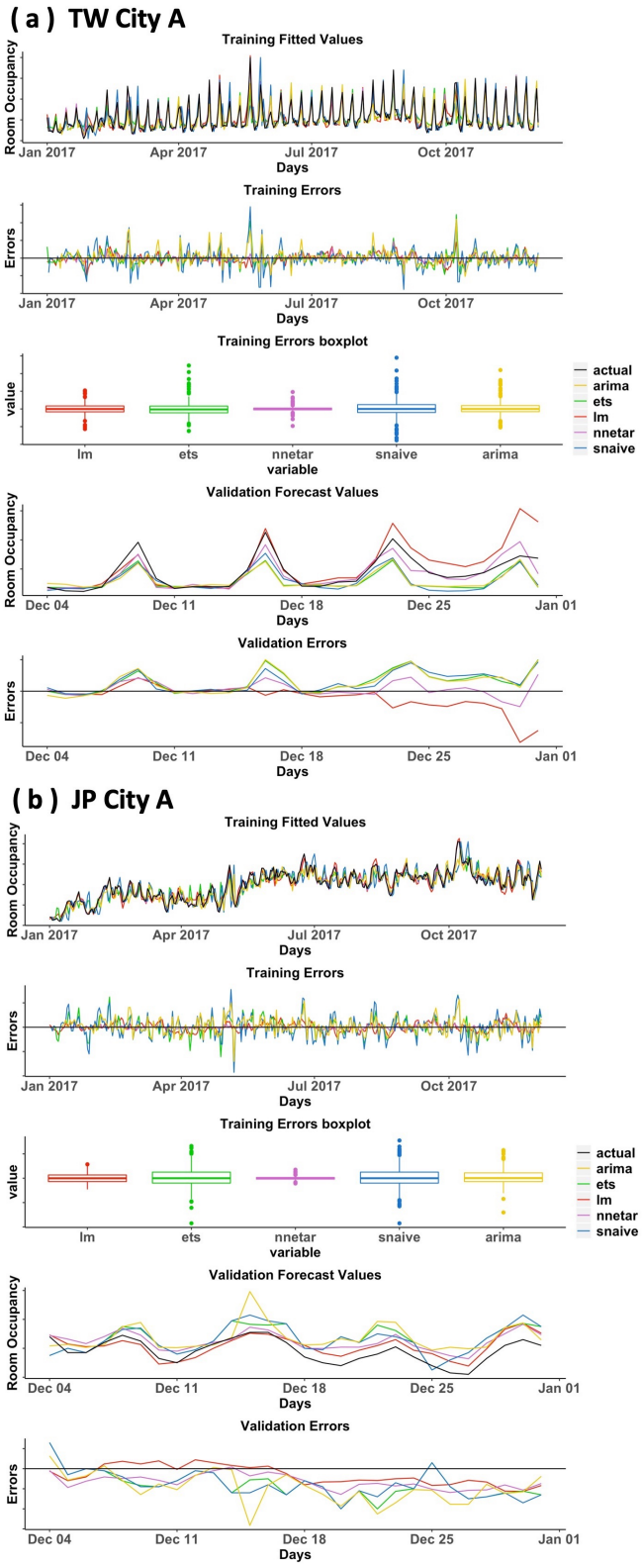

Fig. 3. Performance of different algorithms for TW City A (a) and JP City A (b). For each city, top 3 panels show training performance of all algorithms and bottom 2 panels show validation performance. For TW City A, the neural net (*nnetar*) gave the most precise forecasts in both training and validation periods; In contrast, for JP City A, *nnetar* performed best on the training period, but linear regression (*lm*) performed better on the validation period.

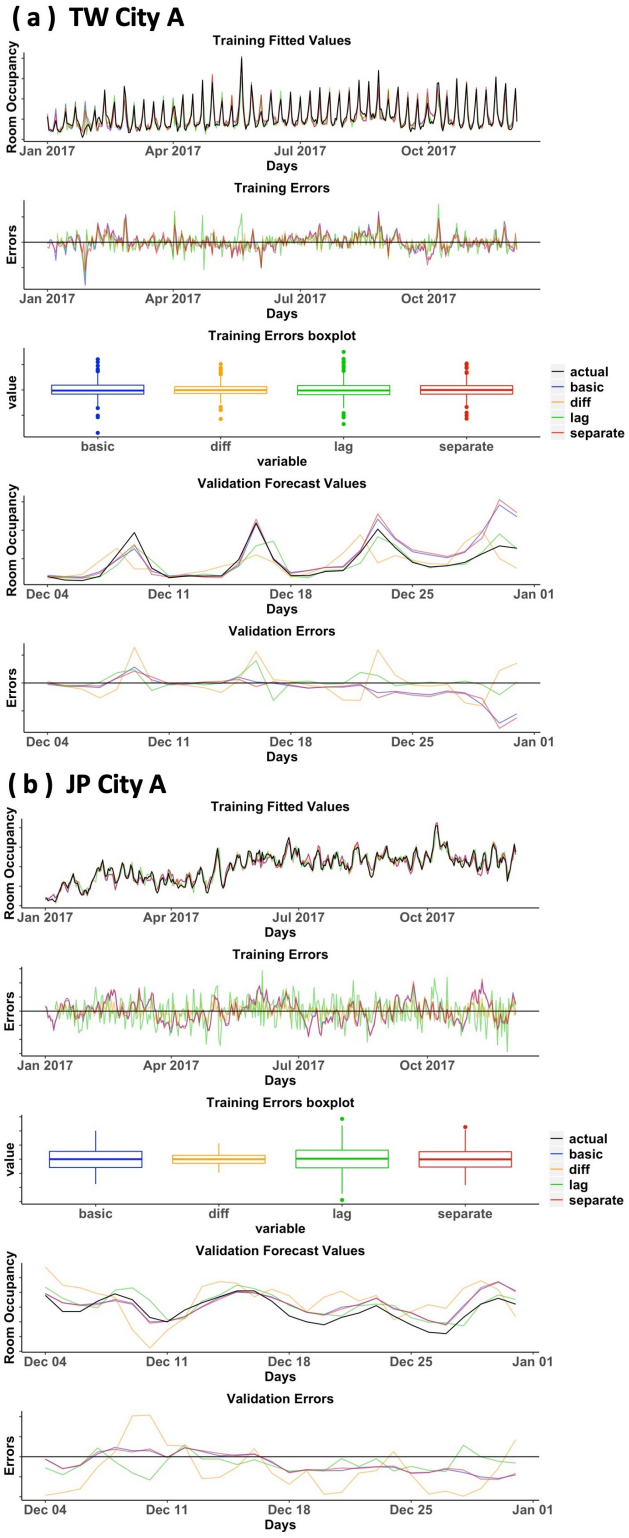

Fig. 4. Performance of different linear regression approaches for capturing day-of-week seasonality for TW City A (a) and JP City A (b). For each city, top 3 panels show training performance of all linear models and bottom 2 panels show validation performance. On average, *basic* approach performed best in both cities on training and validation periods.

of data we can better evaluate performance as well as improve forecasts on holidays.

While exponential smoothing and operations such as differencing allow capturing non-parametric trends and seasonality, our "basic" and "separate models" regression models included a linear trend term. While for most series a linear trend seems reasonable, for a few Japan cities perhaps a quadratic trend is a better approximation. This is one direction for future work.

We note the importance of evaluating and comparing performance by considering more than just single metrics such as RMSE. As our figures show, it is important to know not only the magnitude of the errors, but also *when* they occur. Large errors on holidays have a different meaning and implication than on non-holidays. Over-forecasting has different implications than under-forecasting (in a sharing economy context, under-forecasting might be more damaging as guests' trust will be lost if their bookings are eventually canceled). The side-by-side box plots highlight not only variability of the forecast errors, but also extreme errors - for example, in Figure 3 we see that ARIMA, ETS and seasonal naive have several extreme positive and negative forecast errors, unlike linear regression and the neural network. This visualization in combination with the line plots can help communicate and discuss with the domain experts issues of extreme forecast errors.

In terms of personal data use, our results show the level of accuracy achievable with very basic data: the time stamp and daily bookings in each city. We did not use any user-specific data nor any other type of personal information on the guest or host beyond the check-in and check-out dates and the visited city. It is possible that further information (such as the guest's origin) can help improve accuracy. Moreover, the occupancy data provides information on supply, but not on demand. An interesting future direction is to incorporate data on demand, such as search volume for accommodation in each city on certain days, similar to [12] and [13].

Another future direction, especially useful in the case of many time series, is applying the approach suggested by [10] for clustering a large collection of time series using Model-Based partitioning (MOB) [14]. Each resulting cluster has series with similar temporal patterns and external information, thereby requiring only one linear model per cluster for forecasting all the series. Using the AsiaYo data, we can use the above approach by integrating external information such as weather or holiday dates to cluster the series and forecast series in each cluster using a single linear model instead of individual models for every single series.

VI. Conclusion

We developed, evaluated, and compared several forecasting algorithms and approaches in order to create an automated system for forecasting daily occupancy in many different cities with a 28-day lead time. The solution must be scalable, flexible, and simple in order for a sharing economy company to adopt it for forecasting occupancy in each of many cities on a regular basis.

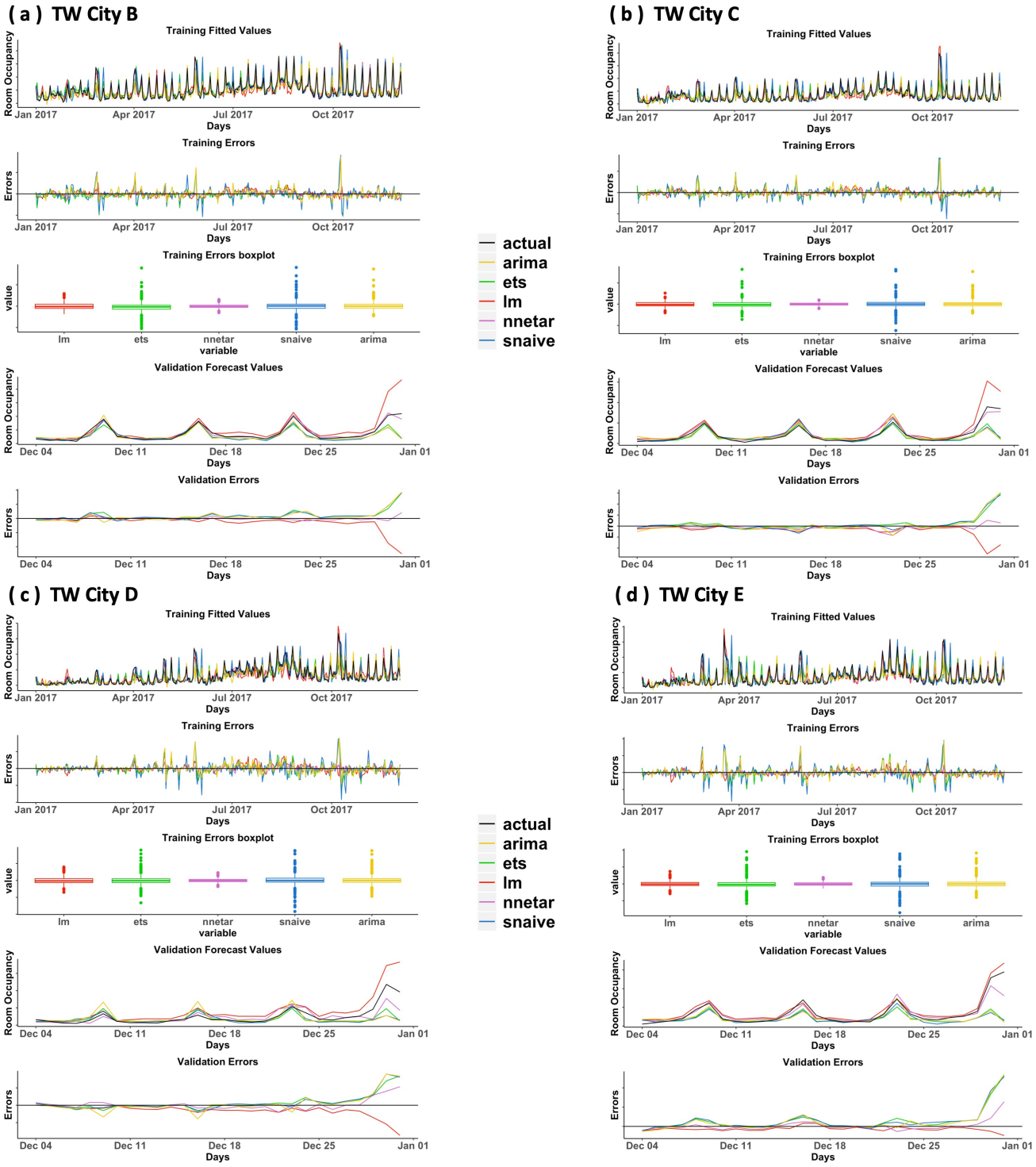

Fig. 5. Performance of different algorithms for 4 other Taiwan major cities. For each city, top 3 panels show training performance of all algorithms and bottom 2 panels show validation performance. Neural net performed best, linear regression is second best (but with larger extreme errors) on training and validation periods for all cities.

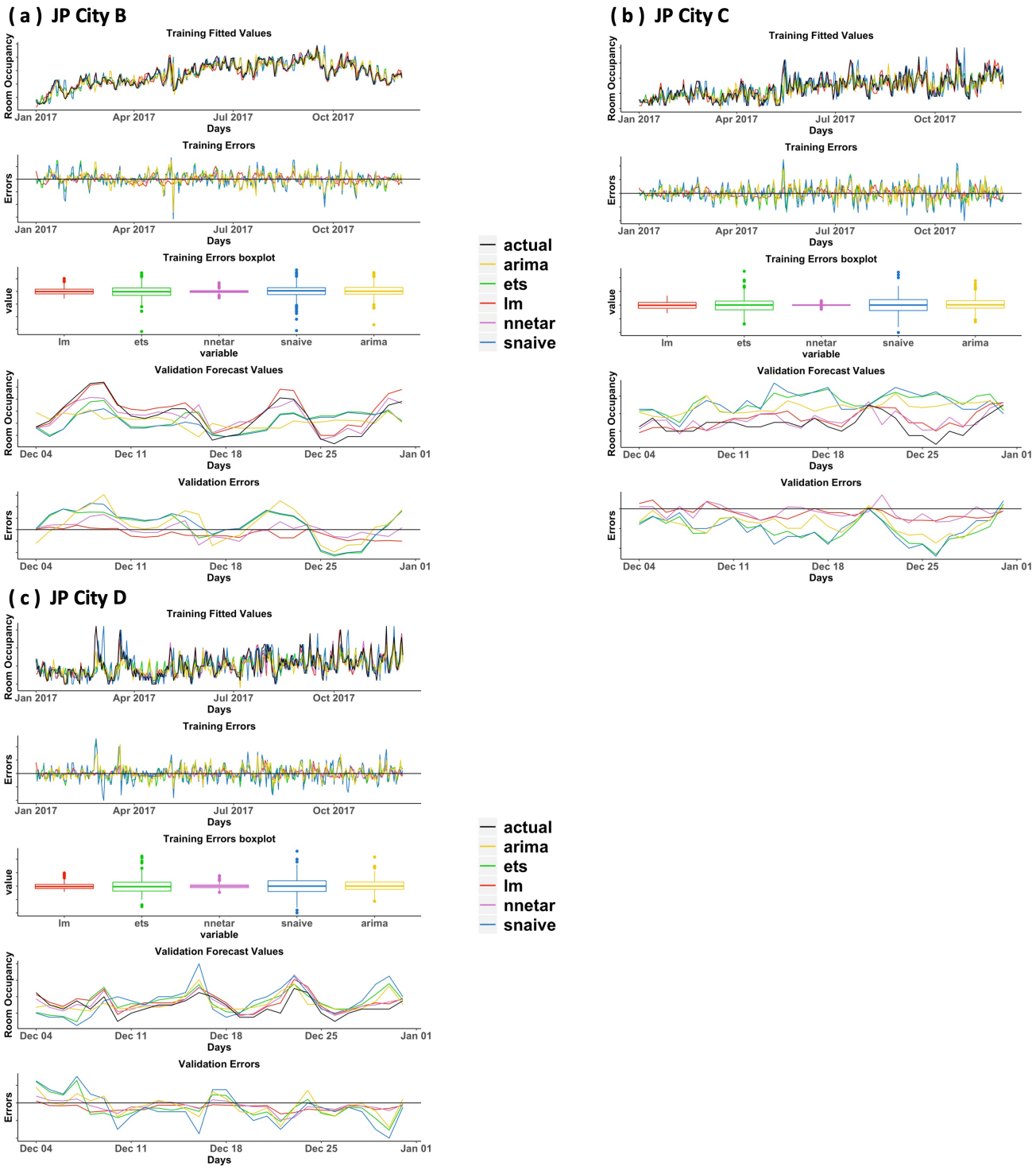

Fig. 6. Performance of different algorithms for 3 other Japan major cities. For each city, top 3 panels show training performance of all algorithms and bottom 2 panels show validation performance. Linear regression (*lm*) and neural net perform best (and are nearly identical) on training and validation periods for all cities.

We found that for cities in Taiwan, the dominating signal is the weekly seasonality (Fridays vs. Saturdays vs. other weekdays). A single-layer neural network performed best on the training and validation periods, with a "basic" linear regression performing similarly, but with larger magnitude extreme errors. We were able to improve the linear regression results by modeling separately Fridays/Saturdays/Sun-Thurs. For cities in Japan, the basic linear regression performed best on the validation period, while the neural network seems to overfit - it performed best on the training period but worse on the validation period. We note that the regression and neural net performances surpassed not only the seasonal naive benchmark, but also traditional forecasting methods such as ARIMA and exponential smoothing. One possible reason is the very strong seasonality which in combination with holidays "confuses" these extrapolation methods; another is the ability to incorporate external information such as holidays and cumulative occupancy into linear regression and neural nets.

Our results showed that shared-economy occupancy is highly dependent on the weekday, city, and holidays. We showed the strengths and weaknesses of different methods in terms of required accuracy level, computation time, and flexibility.

ACKNOWLEDGEMENTS

We thank CK Cheng, Wayne Lai, and Auphie Chen from AsiaYo for their support and valuable feedback.

REFERENCES

[1] S. F. Witt and C. A. Witt, "Forecasting tourism demand: A review of empirical research," *International Journal of forecasting*, vol. 11, no. 3, pp. 447–475, 1995.

[2] P. Khadivi and N. Ramakrishnan, "Wikipedia in the tourism industry: forecasting demand and modeling usage behavior," in *Twenty-Eighth IAAI Conference*, 2016.

[3] M. Mamula, "Modelling and forecasting international tourism demand-evaluation of forecasting performance," *International Journal of Business Administration*, vol. 6, no. 3, p. 102, 2015.

[4] L. R. Weatherford and S. E. Kimes, "A comparison of forecasting methods for hotel revenue management," *International journal of forecasting*, vol. 19, no. 3, pp. 401–415, 2003.

[5] A. Zakhary, N. E. Gayar, and A. F. Atiya, "A comparative study of the pickup method and its variations using a simulated hotel reservation data," *ICGST international journal on artificial intelligence and machine learning*, vol. 8, pp. 15–21, 2008.

[6] S. Yüksel, "An integrated forecasting approach to hotel demand," *Mathematical and Computer Modelling*, vol. 46, no. 7, pp. 1063–1070, 2007.

[7] S. Makridakis and M. Hibon, "The m3-competition: results, conclusions and implications," *International journal of forecasting*, vol. 16, no. 4, pp. 451–476, 2000.

[8] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The m4 competition: Results, findings, conclusion and way forward," *International Journal of Forecasting*, vol. 34, no. 4, pp. 802–808, 2018.

[9] G. Shmueli and K. C. Lichtendahl, *Practical Time Series Forecasting with R: A Hands-On Guide*, 2nd ed. Axelrod Schnall Publishers, 2016.

[10] M. Ashouri, G. Shmueli, and C. Y. Sin, "Tree-based methods for clustering time series using domain-relevant attributes," *Available at SSRN 3282849 (accepted in Journal of Business Analytics - Taylor&Francis)*, 2018.

[11] Y. Khandakar and R. J. Hyndman, "Automatic time series forecasting: the forecast package for R," *Journal of Statistical Software*, vol. 27, no. 03, 2008.

[12] B. Pan, D. C. Wu, and H. Song, "Forecasting hotel room demand using search engine data," *Journal of Hospitality and Tourism Technology*, vol. 3, no. 3, pp. 196–210, 2012.

[13] D. C. Wu, H. Song, and S. Shen, "New developments in tourism and hotel demand modeling and forecasting," *International Journal of Contemporary Hospitality Management*, vol. 29, no. 1, pp. 507–529, 2017.

[14] A. Zeileis, T. Hothorn, and K. Hornik, "Model-based recursive partitioning," *Journal of Computational and Graphical Statistics*, vol. 17, no. 2, pp. 492–514, 2008.