The handle http://hdl.handle.net/1887/80328 holds various files of this Leiden University dissertation.

**Author:** Koupriouchina, L.A.
**Title:** Judgmental adjustments in revenue management: exploring the impact of user overrides on the accuracy of system-generated occupancy forecasts at multiple forecasting horizons
**Issue Date:** 2019-11-12
Chapter 5

Conclusion

I cannot teach anybody anything. I can only make them think.

- Socrates -

5.1 General Contribution
The aim of this thesis was to improve our understanding of judgmental adjustments in hotel revenue management forecasting by exploring the impact of user overrides on the accuracy of system-generated hotel occupancy forecasts at multiple forecasting horizons. To achieve this, three empirical studies were conducted. Collectively, the studies demonstrate that the assessment of hotel forecasting accuracy is a non-trivial and complex task because of lack of one ideal measure. Chapter 2 shows that different accuracy measures can lead to conflicting results when analyzed on a single-hotel dataset. Reexamining these inconsistencies in two subsequent chapters using two large datasets with forecasts from multiple hotels confirms this finding. These studies also reveal that the forecasting horizon and the number of overrides have a significant impact on forecast accuracy. In chapter 3 the number of overrides increased the forecast error, while in chapter 4 it decreased the error. This contradiction could be explained by differences in methodological approach. Chapter 3 analyzed the forecasts with and without overrides, while chapter 4 examined only forecasts with overrides. Specifically, chapter 4 identifies the characteristics of the judgmental adjustments (e.g., timing, size, direction) and their impact on the change in forecasting accuracy. Moreover, it shows that the effect of these overrides depends on the business type for which a forecast is made. That is, overrides of group forecasts improve forecast accuracy more than judgmental adjustments of transient forecasts.

Overall, the thesis contributes to the current body of knowledge in a number of ways. First, the notion that many of the existing forecast accuracy measures are not universally applicable, that they can be incomputable, and that they can
produce misleading results, is well established in the generic forecasting literature (Ahlburg, 1992; Armstrong & Collopy, 1992; Chatfield, 1992; Collopy & Armstrong, 1992; Davydenko & Fildes, 2013; Fildes, 1992; Goodwin & Lawton, 1999; Hyndman & Koehler, 2006; Koehler, 2001; Makridakis, 1993). Investigating this issue in a hotel revenue management context, however, is new. A preliminary call made by Schwartz (1999) was not followed up. For this reason, key problems can be observed in the hotel forecasting literature. First, some forecasting studies have continued to use forecast accuracy measures while ignoring their widely acknowledged limitations, such as measures that favor biases (e.g., MAPE), or measures that are scale-dependent (e.g., MAE, RMSE), even when hotel samples comprised hotels with different sizes. Wu, Song, and Shen (2017) reviewed 171 articles on tourism and hotel demand modeling and forecasting, and found that the most widely used measurements of forecasting accuracy were the mean absolute percentage error (MAPE), the root mean square error (RMSE), the root mean square percentage error (RMSPE), and the mean absolute error (MAE); 66, 36, 33, and 19 studies adopted these measures, respectively. However, as mentioned earlier, all these measures have important limitations. MAE and RMSE are scale-dependent and therefore should not be used for comparison of hotels of different size. MAPE favors underforecasting and is sensitive to outliers. Neither MAPE nor RMSPE can be computed when the time series contain zero. A second issue in the hotel forecasting literature is that the majority of studies use forecasting accuracy measures without properly justifying their selection. This thesis confirms these key problems, and provides empirical evidence on the notion that different forecasting accuracy measures can produce different and even contradictory outcomes in hotel revenue management forecasting. This is an important contribution, as hotel revenue management is characterized by a highly complex forecasting context (e.g., multiple booking horizons, continuous reforecasting, etc.). The empirical support for this contribution was established through three separate studies. The findings of the first study, which were published as Koupiouchina et al. (2014), have been adopted by Pereira (2016, p. 20) who confirms that different accuracy measures produce contradictory answers, “not only within a specific hotel’s segment, but also across segments.” Moreover, the three studies contribute to the existing body of knowledge by generating a warning that the use of forecast accuracy measures, without a deep understanding of the characteristics and limitations of each measure, can lead to misjudged hotel forecasting override decisions. Future research should explore how this may fuel sub-optimal pricing, inventory controls, operational planning, and even strategy decisions. A natural progression of this work is to further analyze the consequences of the inconsiderate use of the forecasting accuracy measures: for instance, reduced reliability of the forecasting performance assessment, undetected wrong interventions, failure to learn from critical mistakes, undesired behavior encouraged by incorrectly designed bonuses, negative impacts on the effectiveness of revenue management systems, and other resulting impediments to the hotels’ efforts to optimize revenue.

Second, in this research, the importance of the time dimension was explored. Prior research either ignored or simplified the role of the time dimension in hotel revenue management forecasting research. Only a few studies analyzed forecasts at multiple forecasting horizons, and in the best case, they included few arbitrarily selected fixed data points, while forecasting is a dynamic process and re-forecasting happens continuously. This is the first time, to our knowledge, that the time variable has been considered as a predictor in the analysis of overrides in a hotel context, providing a more accurate representation of the dynamic reforecasting process. Despite its exploratory nature, chapter 3 identified that the relationship between time and forecast accuracy is complex, with both linear and non-linear components. Moreover, chapter 4 offered insights into the impact of the timing of judgmental adjustments on forecasting accuracy. It was found that early forecast adjustments are associated with lower improvement in accuracy than late forecast adjustments. Furthermore, an interaction effect of the adjustment size on the relationship between time and forecast improvement was observed - time was more important for larger adjustments. To date only two studies (i.e., Claveria, Monte, & Torra, 2015; Pereira, 2016) have evaluated forecast accuracy for multiple forecasting horizons.

While previous work acknowledges the importance of judgmental interventions in forecasting, there are also conflicting views on the impact of these interventions. Some research in this field has demonstrated that although experts can add value, they are also prone to biases, often make unnecessary adjustments, ignore warnings that their adjustments reduce accuracy, and continue making large adjustments even in a presence of high probability of a significant loss (Goodwin & Wright, 1994; Lawrence et al., 2006; J. S. Lim & Oconnor, 1995; Petropoulos et al., 2016). By investigating user interventions in a multiple forecasting horizon context, support for these conflicting views is provided by demonstrating that judgmental adjustments can add value to system-generated forecasts (depending on their characteristics).

Third, this research extends knowledge of the impact of judgmental adjustments on forecast accuracy at a disaggregated level. An interesting finding was that the overrides impact forecast accuracy in a complex manner, depending on the business type. Judgmental forecast adjustments to group segment were associated with larger improvements in accuracy than adjustments to transient segment. Moreover, the size of judgmental adjustments was positively associated with improvements in accuracy and, for the adjustments to the group segment, the size of adjustment had a more profound effect than for the adjustments to the transient segment. These findings are important, as little is known about group forecasting despite its importance for hotel performance. Prior research has mostly reported findings at
the total hotel level (Yuksel, 2007) or only at a group level (Kimes, 1999). However, segmentation is a key characteristic of profitable revenue management, which facilitates tailoring the strategy to match the needs of various customers and to increase bookings, revenue and profit. The empirical findings of this study provide a new understanding of judgmental adjustments at the segment level. The work by Sierag et al. (2017, p. 46) followed our suggestions and concluded that, “As a whole, the study finds support for the work of Koupiouchina et al. (2014) who argue that research in forecasting should take place at a more granular level”. The study challenged previous findings that positive adjustments, “were much less likely to improve accuracy than negative adjustments” (Fieldes et al., 2009, p. 3). In fact, it was demonstrated that the selection of the accuracy measure for this evaluation was also critical: different measures lead to opposite and contradictory conclusions.

Finally, an important contribution is the empirical work itself, which demonstrates how to analyze large real-world datasets on hotel occupancy forecasts through multilevel regression. By examining extensive datasets, this study responds to calls for more work in forecasting based on non-simulated, real-world data (Bodea, Ferguson, & Garrow, 2009; Sierag et al., 2017). This study led to the realization that the analysis of high volumes of real forecasting data is not only time-consuming, but requires several pre-conditions: powerful computers for analysis, good and clear communication with the data provider about the functionality of the RMS, and a solid understanding of the structure and nature of the extracted data. Using a multilevel modeling approach, this study addressed the need for more rigorous and fruitful research in hospitality “by controlling for the contextual and potentially confounding variables that exist at multiple levels” (Wong, 2016, p. 7). This study provides important empirical evidence of the current state of real-world judgmental forecast adjustments considering a wide variety of different hotel types, sizes, and geographical regions over a plurennial forecasting period.

5.2 Limitations of the Research

A review of previous work indicated a number of limitations. Most studies concentrated on the examination of new forecasting methods using simulated data, rather than an evaluation of real business forecasts and their judgmental adjustments. When real business data were used, many studies investigated only one hotel or one hotel chain. Moreover, most of the analyses were performed within a limited timeframe (e.g., several weeks/months), and studies often concentrated on one segment only (e.g., Kimes, 1999). Importantly, the majority of the studies did not justify the selection of the accuracy measures. Even inappropriate measures were used (e.g., MAPE: sensitive to outliers, favoring under-forecasting, MAE: not suitable for comparisons of the series with different scales) without fully disclosing the implications. Also, as demonstrated by Davydenko & Fieldes (2013) some studies applied asymmetrical measures for analyzing the size of judgmental adjustments, thereby undermining previous findings in the literature about the negative relationship between the size of judgmental adjustments and forecasting accuracy. This dissertation overcomes many of these important limitations by exploring the real-business forecast data of 1,752 hotels for two segments over approximately a 3-year period. It analyzed data using four different forecast accuracy measures, justifying its selection and emphasizing the limitations of each measure. Also, it applies Davydenko & Fieldes’ (2013) suggestion, which helps to restore symmetry in the adjustment-size measure, and, therefore, makes the findings about the relationship between the size of judgmental adjustments and forecasting accuracy more robust.

However, the findings of this study are still subject to at least three limitations. The first of these relates to the generalizability of the results. Although the volume of extracted data was extensive, the data were only available for a limited time period (24 months for study 1, and 36 months for studies 2 and 3). Extending the time periods in the analysis would have reinforced the change patterns identified, or perhaps uncovered others (i.e., periodical cycles over several years, etc.). Moreover, as previous-year data were not available, the relative forecast accuracy measures could not be computed as they were based on two-year data, requiring one more previous year of data for studies 2 and 3.

Second, even though an extensive and comprehensive dataset was obtained, data were obtained from a single revenue management system (i.e., IdeaS revenue management system – G2), and, therefore, the conclusions may not apply to other revenue management systems. Inclusion of other revenue management systems in the study (e.g., Duetto, Infor EzRMS, Rainmaker) would have increased robustness. This choice was made for practical purposes – it took several weeks of specialized training to understand how the RMS forecasting module worked and several more months to learn how to combine and interpret extracted data. Moreover, different systems have different functionalities and data structures. Considering the limited empirical evidence on hotel revenue management forecasting overrides, it was decided to opt for an in-depth revelatory study of one revenue management system using data from more than a thousand hotels around the world.

Third, and perhaps the most important limitation, lies in the absence of specific methods developed for hotel revenue management forecasting research. Approaches and methods of this study were inspired by product SKU-forecast research. SKU forecasting in manufacturing has some important differences with RM forecasting in hotels. Therefore, great care must be taken when adopting the SKU research methods in a hotel Revenue Management context. For example, a lack of a well-calibrated and well-tested methodology for analyzing the multiple overrides over the booking horizon was identified. In revenue management systems, the status of the
override can change from active to reverted, and a reverted override can be (re) activated later. While override status was included as a variable in the multilevel models, it was not analyzed at a granular level because the time dimension of the changes in the override statuses could not be precisely analyzed i.e., it was uncertain how long specific overrides remained active in the system, and how critical the timing of changes in their statuses were. As the hotels in the datasets varied in the way the majority of the hotel demand was received (e.g., some received it within the last 90 days before arrival, whereas for others it was spread out across 365 days), it was difficult to determine how uncertainty in the time dimension of the override status changes would affect the results at the individual hotel property level.

5.3 Directions for Future Research

This thesis opens up at least six interesting and novel areas for future research. The following research suggestions provide possible directions to extend the results of this study, and thus to increase the body of knowledge on the hotel revenue management forecasting adjustment process.

1. This research incorporated a limited number of hotel-level variables (e.g., geographical region, hotel type, size of hotel). Future research could include additional hotel-related variables. For example: Are there any common characteristics that top performers (in terms of forecasting) share? How long have they used (the same version of) the RMS? Do they operate in highly competitive and volatile markets, or in fairly stable markets? Are these hotels predominantly chain-affiliated or independent? Moreover, according to a review conducted by Lawrence et al. (2006), judgmental forecast accuracy also depends on the characteristics of time series (e.g., trend, seasonality, randomness, and discontinuities). For instance, previous research showed that subjects reacted differently to downward trends than to upward trends. Andreassen and Kraus (1990), found that noise also impacts the forecaster’s ability to detect a trend, and this was confirmed in later work. Including these and similar variables would improve our understanding of how organizational characteristics and market contexts impact judgmental adjustments, which would in turn affect hotel revenue management forecasting performance.

2. A richer theoretical framework including variables at different levels should be developed and tested. According to I.-P. C. Chen (2010), the interaction between a revenue manager’s judgment and RMS is not the only element that affects the accuracy of forecasting. Other influential factors should be taken into consideration. The author distinguishes between the external environment (e.g., the economy, social trend, or competitors’ action) and internal factors (e.g., RMS, Human judgment, Model selection, Forecasting method, Data collection, Goal of forecasting). The author also states that “professional’s judgment may be affected by work experience, available information, and related knowledge within the area, as well as by motive, belief, and forecasting system” (I.-P. C. Chen, 2010, p. 16). It is suggested to advance research on judgmental adjustment in hotel revenue management forecasting, with multilevel modeling including variables at several levels and their interactions. This is similar to Wong (2016) who proposed a research agenda in which multilevel methods could help to investigate the linkage between service providers, employees, and customers.

3. The mechanisms of the decision-making processes and users’ behaviors associated with each override are virtually unknown. A greater focus on psychological and behavioral aspects of forecast overrides could lead to a better understanding of how and why certain decisions are made. Further research might explore, for example, research questions such as: How do Revenue Managers review forecasts? How do they make a final decision on whether to override? How do they determine the size of override? Do they periodically monitor the override, and adjust or revert it if necessary? How do they evaluate their forecasting performance? Other more fundamental research questions that might be explored are: Why do some revenue managers override more than others? How do perceptions and trust in the RMS forecast impact forecast adjustment behavior? Are there personality traits that explain judgmental adjustment behavior? Currently, very little research has been done to clarify these issues. Schwartz and Cohen (2004b) conducted a study involving 57 experienced hotel revenue managers who were exposed to simulated forecasting software. Gender of the revenue manager was found to be significant for the level of expressed confidence in the forecast accuracy; the women expressed about 10% less confidence in their forecasts than the men. Another variable influencing the confidence according to the study was experience. “For every additional year of experience, the level of confidence in forecast increases by 0.37%” (Schwartz & Cohen, 2004b, p. 64). Since confidence in the own forecast/expertise may have a direct impact on the propensity to override, it may be worthwhile hypothesizing whether, for example, personality affects propensity to adjust system-generated forecasts through its effect on the level of confidence. Insights from the Human-Computer Interaction (HCI) literature, which specifically deals with the ways humans interact with information and technologies (Rogers, 2012), could help to identify relevant theories, and to formulate these hypotheses related to the interaction between RM professionals and the RMS, the reasons behind their forecast overrides, and how trust, behavior, and attitude affect accuracy and...
consequently financial performance. In addition, the theory of ecological rationality (Gigerenzer et al., 1999), which suggests that the mind adapts its limitations to match the structures of information available in the environment, may be useful.

4. In the Human-Computer Interaction field a systematic research has been done to select relevant cognitive theories that could be useful for the interface design. The research on this topic in the hotel RMS context is virtually non-existent. According to Schwartz and Cohen (2004a, p. 85), “the nature of the user interface influenced the way the revenue managers adjusted the computers’ forecasts.” Since the user interface impacts the behavior related to the judgmental adjustments, the impact of the revenue management system interface on the user forecast adjustment behavior deserves additional attention.

5. Forecast adjustments are not the only way in which users can influence the system. Apart from forecast overrides, revenue managers can perform price overrides or overbooking levels overrides; they can activate or modify length of stay controls, register special events, and perform other system activities that may influence decisions suggested by the revenue management system. Future work should explore the attributes of these overrides and their impact. A natural progression of this study is to analyze pricing overrides and their impact on, for example, Average Daily Rate (ADR) or Average Daily Rate Index (ARI). In a similar vein, further research can explore whether the overrides of overbooking levels can lead to improvements in occupancy. It would be interesting to investigate whether some of the conclusions related to forecast overrides are, to some degree, applicable to the price and overbooking level overrides.

6. Building on the previous suggestion, it can be proposed, that although forecast adjustments have a direct impact on forecast accuracy, there may also be cascade effects in the complex system of interrelated key revenue management decisions (e.g., capacity allocation, length of stay controls, pricing). The RMS under analysis in this thesis consisted of four interconnected modules: (a) forecasting, (b) optimization, (c) control, and (d) monitoring. The results of the forecasting module are fed into an optimization module which determines the optimal mix of demand for each day. Once the optimal business mix is defined by the optimization module, its results are used in the control module responsible for the proposal of the controls that are necessary to ensure that an optimal mix of demand is achieved. This control module is connected to the monitoring module; the circle is completed through the connection of the monitoring module with the forecast module. The process is dynamically repeated multiple times for the same arrival date. But, apart from the forecast overrides, users can perform price overrides or override the overbooking levels. Future work should also study all overrides as part of a complex, dynamic, and interconnected system, where one decision simultaneously impacts many others. If the debate on the impact of user influences on revenue management system is to be moved forward, a more holistic understanding and examination of major user influences and their impact on topline (e.g., RevPAR, TrevPAR), bottom line (e.g., GOPPAR) and competitive set performance (e.g., MPI, ARI, RGI) is needed. With regard to the latter, a deeper understanding could be gained of the flow-through effects of user overrides on operational decisions and expenses (e.g., payroll).

5.4 Implications

The findings of this study have a number of implications for practice, and are of interest to hotels as well as to RMS software providers.

This study has illustrated potential deficiencies and dangers of the inconsiderate use of forecasting accuracy measures. Different error measures generate contradictory answers and the forecasting accuracy can be misjudged and, as a consequence, potentially undermine decision-making in other important hotel management areas: pricing, inventory control, operational planning, distribution, and strategy. The findings also demonstrate the need to further educate the Revenue Managers about the pitfalls and biases of each accuracy measure, so that they are in a position to carefully select the forecasting accuracy measures applicable for their conditions. This cautious choice should be based on the theoretical nature of error measures, together with the practical implications of the forecast, the nature of the derived decisions, and the data characteristics.

Given the importance of forecasting in the hotel revenue management optimization cycle, a more fruitful approach may be to expand the common set of narrow accuracy measures approach with a more structured, comprehensive and consistent framework of forecasting quality assessment. This framework should incorporate multiple forecast accuracy measures, as no single measure can be preferred. One of the important elements of this framework could be the collaborative development and implementation of a set of automated or semi-automated forecast quality monitoring procedures, including feedback mechanisms that allow users to learn from their past decisions and actions. To make these feedback mechanisms meaningful, additional efforts will be required from the hotels and their Revenue Managers. For example, it would be extremely helpful if Revenue Managers kept a consistent log in RMS of the reasons for overrides, especially for large and frequent overrides, making it possible to systematically assess these reasons and incorporate the outcomes of this analysis in the feedback loop.
RMS providers, on the other hand, could add automated procedures to continuously monitor judgmental adjustments introduced by the Revenue Managers, and analyze and report on various important aspects such as size, direction, frequency, timing, segments to which they are applied, type of override (demand or wash), and so forth. This analysis could include measurements of override effectiveness and the impact on forecasting performance, categorized by type of override and by reason. This enhancement to the RMS could be used to generate alarms, as well as periodic reports to users, system administrators and solution providers. Moreover, by leveraging the immense number of exchanges with the thousands of hotels operating in different conditions, RMS software providers could further shape their software with a deeper understanding of the users’ behavior. These insights could influence how new versions of the software are designed and launched, how the users are trained (e.g., informing them on how to better make override decisions, which pitfalls to avoid in interpretation of the forecasting performance), and how customers’ requests are handled (e.g., in case of complaints or requests to fine-tune the forecasting module in order to improve forecasting accuracy).

Another practical implication is that both forecasting horizons and aggregation levels should be taken into account when designing a comprehensive forecasting quality assessment framework. Current industry practice is to evaluate forecasting performance at the total hotel level only 30 days before arrival. There are two major concerns with this approach. First, as highlighted in this study, the relationship between time and forecasting accuracy is complex. It is therefore suggested to monitor forecasting accuracy at multiple time horizons, taking into account hotel-specific time frame(s) at which forecasts influence important managerial decisions. Second, decisions in revenue management are segment-based, and reporting accuracy at the hotel level only could potentially distort the real picture (e.g., ‘compensating’ under-forecasting in one segment over-forecasting in another segment). Therefore, a reasonable approach would be to use more granular, segment-level data for segment-level forecast accuracy monitoring.

In sum, a well-designed, systematic and appropriately implemented framework could help Revenue Managers make better judgments about forecasting performance, avoiding misleading interpretations, encouraging helpful override behavior, and minimizing override behaviors that are damaging for the forecasting performance. Regular forecast monitoring could also help to compensate for the ‘black box’ effect related to an incomplete understanding by the Revenue Managers of how the algorithms behind RMS really function. Due to the high numbers of forecasts, and the dynamic and complex nature of forecasting and re-forecasting processes, this framework, however, cannot be developed in isolation. Close cooperation and continuous communication between hotels and RMS software providers is a key condition for the continuing success of hotel revenue managers.

Conclusion

REFERENCES