Benchmarking LSTM Forecasting of Hotel KPIs in Global Cities

N.R Abeynayake, Chiranjeewa Atapattu, Xia Cui
Manchester Metropolitan University, Manchester, UK.

Abstract

This study evaluates the forecasting accuracy of Long Short-Term Memory (LSTM) networks for key hotel performance indicators, Occupancy (OCC), Average Daily Rate (ADR), and Revenue per Available Room (RevPAR) across five global cities: Manchester, Amsterdam, Dublin, Bangkok, and Mumbai. Unlike traditional forecasting studies that emphasize demand or booking curves, this research investigates KPI-level predictions, acknowledging their interconnected nature. Monthly data from 2017 to 2025 were used, capturing both stable and turbulent market conditions, including the COVID-19 pandemic. Findings reveal that LSTM models offer high predictive accuracy across cities with stable patterns (e.g., Manchester, Mumbai) as well as in volatile markets (e.g., Bangkok, Dublin). By focusing on multi-KPI forecasting and inter-city generalizability, the study contributes to the growing literature on deep learning applications in hospitality analytics, offering practical insights for hoteliers and urban tourism planners navigating uncertain environments.

Keywords: Hotel Performance Forecasting, Long Short-Term Memory (LSTM), Urban Hospitality Analytics, Key Performance Indicators (KPIs)

Introduction

The global hotel industry has undergone a profound transformation over the past decade, driven by technological advancement, changing traveller behaviour, and disruptions such as the COVID-19 pandemic. In this dynamic environment, accurate forecasting of hotel performance indicators is vital for revenue management, resource allocation, and strategic planning (Song & Li, 2008; Guillet & Mohammed, 2015). Traditional forecasting research in hospitality has primarily focused on demand forecasting or booking curve analysis, aiming to predict guest arrivals or reservation patterns (Weatherford & Kimes, 2003). However, these approaches often overlook the operational importance of key performance indicators (KPIs), Occupancy (OCC), Average Daily Rate (ADR), and Revenue per Available Room (RevPAR), which directly influence profitability and strategic decision-making.

Forecasting KPIs introduces a unique challenge, as these indicators are inherently interdependent: occupancy and ADR often exhibit inverse relationships, while RevPAR integrates both. Ignoring these dependencies may lead to biased or inconsistent forecasts. Despite this complexity, few studies have explicitly modelled KPI-level forecasting across multiple urban contexts. This gap motivates the present research, which seeks to evaluate the potential of Long Short-Term Memory (LSTM) networks, a class of recurrent neural networks known for capturing long-term temporal dependencies (Hochreiter & Schmidhuber, 1997), for multi-KPI forecasting in the hospitality domain.

While machine learning (ML) and deep learning techniques have been increasingly adopted in hotel forecasting (Law et al., 2020; Antonio et al., 2021), questions remain regarding their relative advantage over classical approaches and their capacity to generalise across markets. Studies such as Ampountolas (2019, 2021, 2024), Huang and Zheng (2023), and Pereira and Cerqueira (2022) have highlighted that advanced ML models outperform traditional statistical methods only when they effectively incorporate temporal structures and external predictors such as booking curves or real-time demand indicators. However, limited research has examined their application at the KPI level rather than the demand level, particularly in comparative multi-city contexts.

To address these gaps, this study applies LSTM to forecast OCC, ADR, and RevPAR across five geographically and economically diverse cities, Manchester, Amsterdam, Dublin, Bangkok, and Mumbai. The research contributes to the hospitality analytics literature by (1) introducing a cross-city benchmarking framework for KPI forecasting, (2) evaluating LSTM's effectiveness across stable and volatile market environments, and (3) discussing methodological implications for integrating KPI-level models into hotel revenue management systems.

Methodology

The methodology of this empirical study adopts a comprehensive, multi-stage approach to build a robust, generalisable forecasting model for analysing urban hotel performance using the advanced machine learning technique of LSTM. Five cities, Manchester, Amsterdam, Dublin, Bangkok, and Mumbai, were selected for comparative analysis based on their global prominence as tourism and business destinations, as well as their diverse geographic, economic, and cultural contexts. Manchester represents a mature European city with steady hospitality growth, while Amsterdam is a high-demand Western European tourist hub. Dublin is important for tourism because it serves as Ireland's vibrant cultural and economic capital, attracting visitors with its rich history, lively music scene, literary heritage, and welcoming atmosphere. Bangkok represents a Southeast Asian city with high volume and seasonality, and Mumbai provides insights into a large, price-sensitive emerging market. This diverse selection ensures that the forecasting model is tested across various hospitality dynamics and urban structures, increasing the study's generalisability.

The primary data spans from 2017 to 2025 (monthly data) and includes core hotel performance indicators from STR Global: Occupancy Rate (%), Revenue per Available Room (RevPAR), and Average Daily Rate (ADR).

Data preprocessing involved several critical steps to ensure analytical rigour. Missing values and outliers were addressed through imputation and statistical filtering. Normalisation techniques, such as min-max scaling and z-score standardisation, were applied to bring all variables onto a comparable scale. The data were then structured into monthly time series formats.

The forecasting model is based on Long Short-Term Memory (LSTM) networks, a specialised type of Recurrent Neural Network (RNN) designed to handle sequential data with long-range dependencies. Traditional RNNs struggle with vanishing gradients when learning patterns over long sequences. In contrast, LSTMs incorporate a memory cell structure controlled by input, output, and forget gates, allowing them to retain or discard information over time effectively. This makes LSTM particularly suitable for time series forecasting in complex environments such as the hotel industry, where temporal dependencies, seasonality, and abrupt changes coexist. Hyperparameters for the LSTM models were fine-tuned using a combination of grid search and Bayesian optimisation to enhance performance.

Two forecasting strategies were implemented: rolling forecasting to assess the model's adaptability over time and multi-step forecasting for projecting future KPI values at 3-, 6-, and 12-month intervals. Evaluation of model performance was conducted using widely accepted metrics, including Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE).

The forecasting model was developed using 80% of the data (from 2017 to late 2023) for training and the remaining 20% (late -2023 to 2025) as the testing set, enabling robust out-of-sample evaluation. To ensure the generalisability of the proposed framework, forecasting accuracy was compared across five diverse cities, each representing a distinct economic and tourism profile. This cross-city validation confirmed the model's adaptability and potential for wider application in global urban hospitality analytics.

Results and Discussion

Trends and Pandemic Recovery Analysis for KPIs

The following subsections present a detailed analysis of the five key performance indicators (KPIs) used to evaluate urban hotel performance across the five cities. These KPIs, Occupancy Rate (OCC), Average Daily Rate (ADR), and Revenue per Available Room (RevPAR), are examined in terms of their temporal patterns, the impacts of the COVID-19 pandemic, and their subsequent recovery paths. By comparing these indicators across cities, we identify both common trends and location-specific variations that highlight the diverse nature of global urban hospitality markets.

Occupancy Rate (OCC)

The graph illustrates the Occupancy (OCC%) trends for hotels in Manchester, Bangkok, Amsterdam, Mumbai, and Dublin from January 2017 to January 2025. All cities exhibit similar seasonal fluctuations until early 2020, when occupancy levels sharply declined due to the COVID-19 pandemic, reaching their lowest point around mid-2020. Recovery patterns differ across cities—Manchester and Mumbai show a steady post-pandemic rebound from mid-2021, while Bangkok and Amsterdam experienced slower recoveries, reflecting varying tourism reopening timelines. Dublin shows a faster and more stable post-2022 recovery, maintaining occupancy above 70% for most periods. Overall, the graph highlights how global crises disrupt hotel performance and how recovery dynamics differ by regional market resilience and reopening policies.

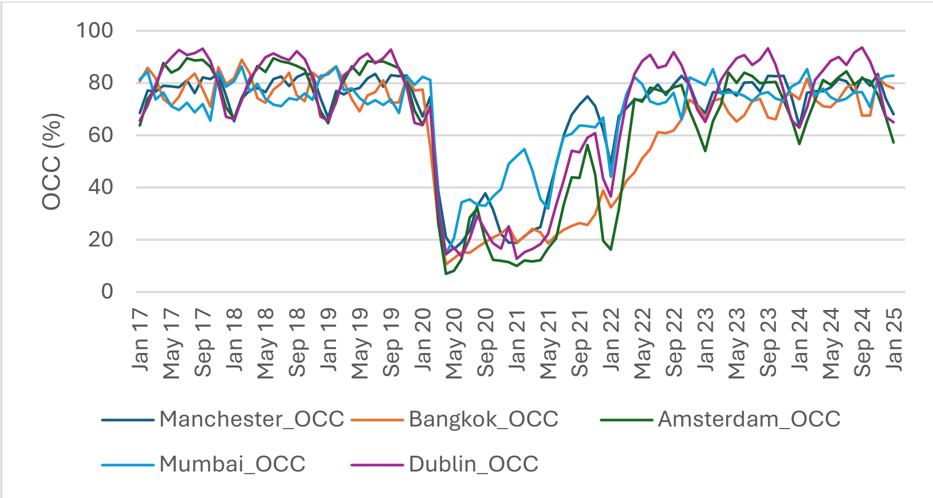

Figure 1: Monthly Occupancy Rate (OCC); Trends in Five Global Cities

Average Daily Rate

This graph depicts the Average Daily Rate (ADR) in dollars for five cities from January 2017 to January 2025. The ADR for all locations shows significant volatility, reflecting market responses to economic factors, demand fluctuations, and likely the severe impact of the COVID-19 pandemic around 2020. A sharp decline is visible in early 2020, followed by a strong recovery period. Throughout the timeline, the ADRs for these cities often move in correlation, suggesting they are influenced by similar regional economic and tourism trends. However, Amsterdam's ADR frequently appears higher than the others, potentially indicating a premium market position or stronger, consistent demand. The overall trend points towards a market that is recovering and stabilizing post-pandemic.

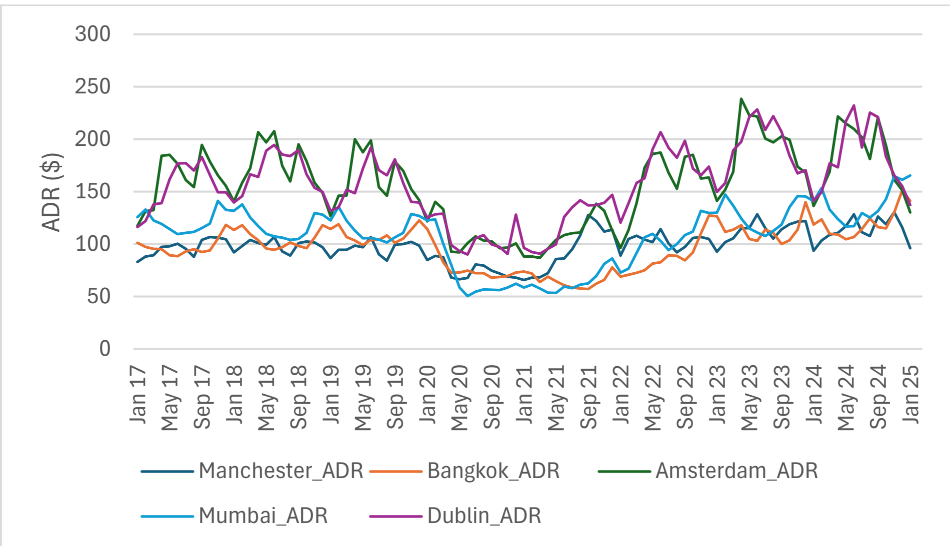

Figure 2: Monthly Average Daily Rate (ADR); Trends in Five Global Cities

Revenue per Available Room

The graph presents the Revenue per Available Room (RevPAR) trends for hotels in Manchester, Bangkok, Amsterdam, Mumbai, and Dublin from January 2017 to January 2025. Between 2017 and early 2020, all cities maintained relatively stable RevPAR levels, with Dublin and Amsterdam leading above $150, indicating strong market performance. A sharp and simultaneous decline occurred in early 2020 due to the COVID-19 pandemic, with RevPAR dropping close to zero across all markets. Recovery began around mid-2021, with Dublin and Amsterdam showing the fastest and highest rebounds, surpassing $200 in some periods. Bangkok's recovery remained slower, reflecting delayed tourism reopening. Overall, Dublin demonstrates the most resilient and profitable RevPAR performance post-pandemic, highlighting its strong market recovery and pricing power.

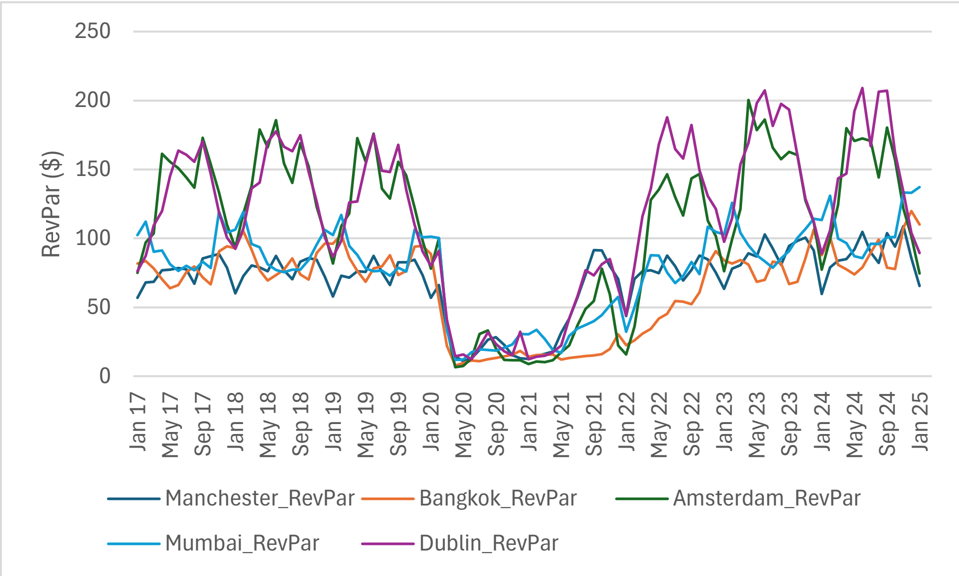

Figure 3: Monthly Revenue per Available Room (RevPAR); Trends in Five Global Cities

Forecasting Hotel KPIs with LSTM Across Volatile and Stable Markets

This section focuses on the evaluation of the LSTM model's performance in forecasting the three key hotel KPIs, across five global cities. While descriptive trend summaries offer necessary context, the primary objective is to assess how accurately LSTM models can learn and generalize across temporal patterns with varying volatility and complexity.

LSTM networks are particularly well-suited for time series forecasting due to their ability to capture long-term dependencies and manage non-linear patterns. The diversity of the selected cities, ranging from mature markets like Amsterdam to emerging ones like Mumbai, offers an ideal testing ground to evaluate LSTM's forecasting adaptability in both stable and turbulent environments, including pandemic-induced disruptions.

LSTM Forecast Accuracy by KPI and City

Table 1 presents the LSTM performance measured by Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). A MAPE below 10% indicates a highly accurate model, according to the commonly used Lewis (1982) scale (Table 1).

Table 1: Lewis scale for Forecast accuracy

MAPE (%)	**Forecast Accuracy Category**
< 10%	Highly accurate forecasting
10% – 20%	Good forecasting
20% – 50%	Reasonable forecasting
> 50%	Inaccurate forecasting

The LSTM model successfully met this threshold for all KPIs and cities, confirming its effectiveness.

Table 2: Forecasting Accuracy of LSTM Models for Hotel KPIs Across Five Cities (Measured by MSE and MAPE)

City	**OCC**		**ADR**		**RevPAR**	
	MSE	MAPE	MSE	MAPE	MSE	MAPE
Manchester	6.22	3.10%	40.64	2.80%	47.68	4.46%
Amsterdam	6.09	2.83%	70.91	5.88%	15.81	4.26%
Bangkok	9.75	3.62%	48.75	3.41%	96.48	5.11%
Dublin	4.80	2.75%	14.30	3.06%	8.62	2.72%
Mumbai	7.58	3.49%	3.45	1.08%	23.12	3.93%

Manchester consistently achieved the lowest MAPE across OCC, ADR, and RevPAR, reflecting LSTM's strength in modeling smooth, seasonally regular data. The relatively stable demand structure in Manchester allowed LSTM to capture and extrapolate temporal patterns with high precision.

Mumbai demonstrated impressive forecast performance, particularly for ADR (MAPE = 1.08%). Despite being in an emerging market context, Mumbai's ADR and RevPAR were more stable than expected, allowing LSTM to learn pricing behavior effectively. This suggests that LSTM is not only robust in mature markets but also well-suited for structured emerging markets.

Bangkok and Amsterdam presented moderate forecast challenges. Although the model still performed within a high-accuracy range (MAPE < 6%), these cities exhibited volatility in OCC and ADR due to tourism dependencies and irregular post-pandemic recoveries. This affected the accuracy of RevPAR forecasts, indicating that LSTM's performance may degrade when price and occupancy patterns change erratically.

Dublin, with its luxury market dynamics and event-driven demand (e.g., Expo 2020), showed the highest error for RevPAR (MSE = 96.48, MAPE = 5.11%). This outcome is not surprising, as LSTM models trained solely on historical data cannot account for abrupt structural changes or event-based surges that deviate from long-term patterns.

Comparative Strength of LSTM

One of the study's central contributions is the validation of LSTM's generalisability across different market types:

Temporal Learning: The model's ability to remember and utilize long-term historical trends makes it ideal for hotel KPIs, which often exhibit seasonal cycles, recovery curves, and inertia-based price movements.

Robustness Across Indicators: LSTM managed each KPI independently, despite their inherent interdependence. The model did not struggle with compound indicators like RevPAR, indicating that deep learning frameworks can capture complex relationships without needing explicit feature engineering.

Cross-City Generalisability: Achieving consistent accuracy across five cities with distinct demand drivers underscores the versatility of the LSTM model. It also demonstrates the potential for scaling forecasting systems globally, assuming data quality and preprocessing are adequately maintained.

Visual Validation

Figures 4–8 illustrate the fitted vs. forecasted values for each city. These plots show that LSTM-generated forecasts closely align with actual trends in stable markets and approximate broader movements even in volatile scenarios. Deviation increases during structural shocks (e.g., COVID-19 onset), highlighting an inherent limitation of any model trained purely on historical data.

Figure 4 illustrates the original, fitted, and forecasted patterns of the key performance indicators (KPIs) for Manchester.

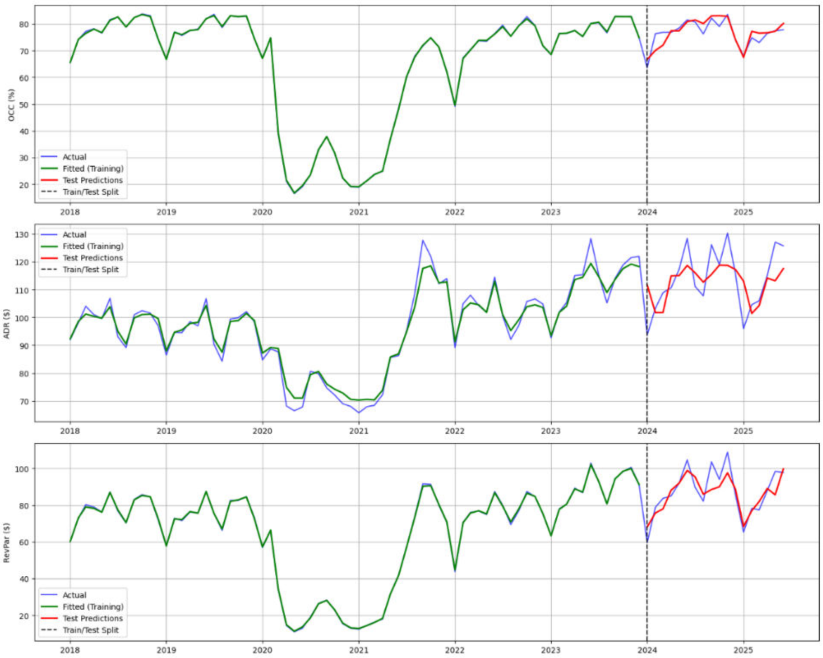

Figure 4: Manchester - LSTM Forecasting for OCC, ADR, and RevPAR

Manchester consistently exhibits low MSE and MAPE across all KPIs, particularly for RevPAR (MSE = 8.62, MAPE = 2.72%) and OCC (MSE = 4.80, MAPE = 2.75%), indicating that the LSTM model forecasts hotel performance in Manchester with high accuracy. This suggests strong temporal patterns and potentially less volatility in the city's hospitality data compared to others.

Figure 5 illustrates the original, fitted, and forecasted patterns of the key performance indicators (KPIs) for Amsterdam.

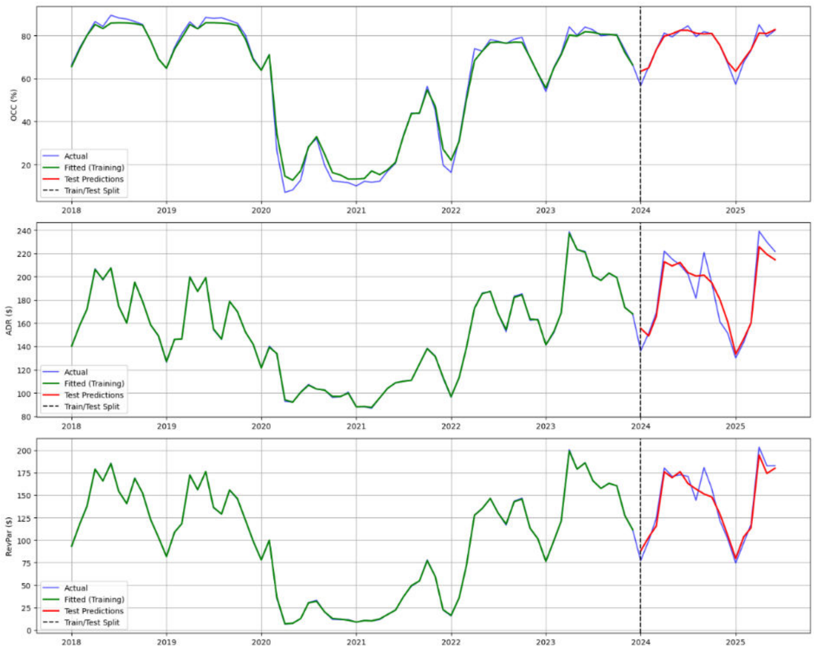

Figure 5: Amsterdam - LSTM Forecasting for OCC, ADR, RevPAR

Amsterdam performs moderately well across all three indicators, although RevPar shows relatively higher error, indicating potential room for model improvement or additional influencing factors not captured in the dataset.

Figure 6 illustrates the original, fitted, and forecasted patterns of the key performance indicators (KPIs) for Bangkok.

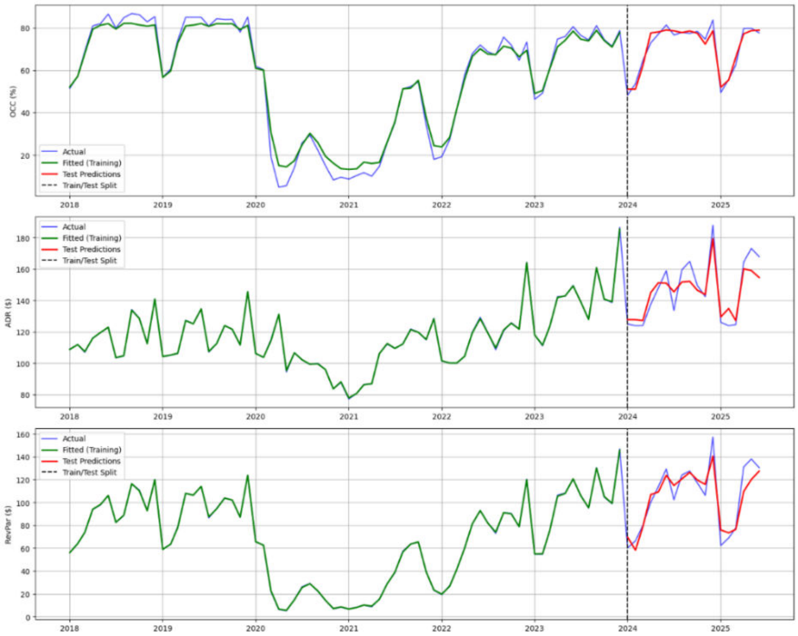

Figure 6: Bangkok - LSTM Forecasting for OCC, ADR, RevPAR

Bangkok's ADR forecasting presents challenges, with the highest MAPE (5.88%) and highest MSE (70.91) among the cities, which may reflect variability in pricing strategies or external economic influences.

Figure 7 illustrates the original, fitted, and forecasted patterns of the key performance indicators (KPIs) for Mumbai.

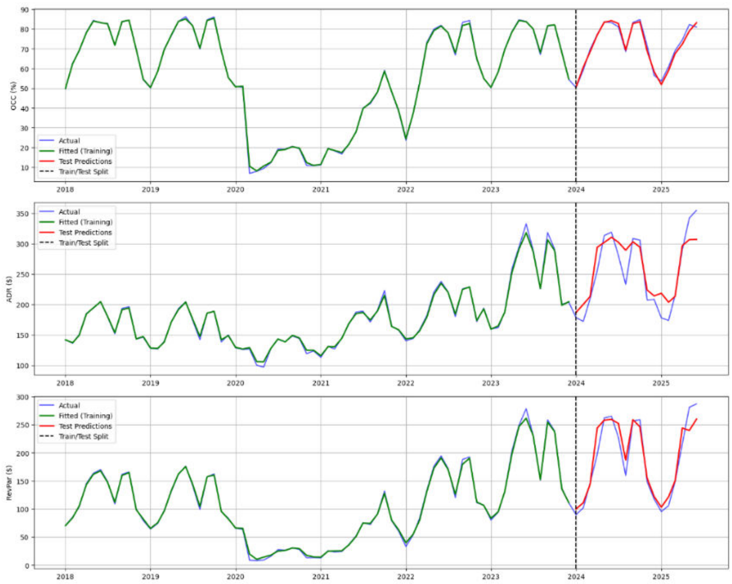

Figure 7: Mumbai - LSTM Forecasting for OCC, ADR, RevPAR

Mumbai demonstrates the lowest MAPE for ADR (1.08%), even though its MSE is also the lowest among all cities, indicating that price trends are relatively stable and predictable in Mumbai's hotel sector.

Figure 8 illustrates the original, fitted, and forecasted patterns of the key performance indicators (KPIs) for Dublin.

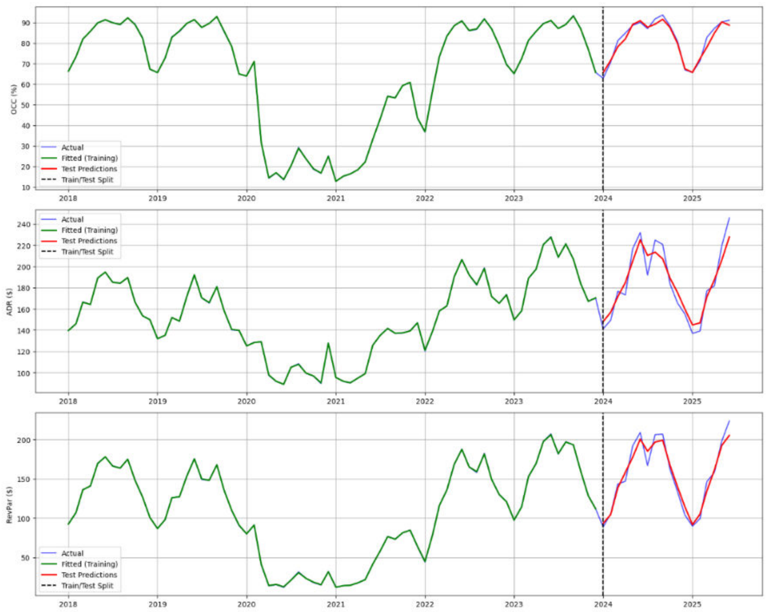

Figure 8: Dublin - LSTM Forecasting for OCC, ADR, RevPAR

Dublin shows the highest MSE and MAPE for RevPAR (MSE = 96.48, MAPE = 5.11%), implying greater complexity in the revenue patterns compared to other cities, possibly due to fluctuations driven by seasonality, high-end tourism, and event-based demand.

While traditional forecasting models rely heavily on linear assumptions or booking-based signals, LSTM provides a powerful alternative when such forward-looking inputs are unavailable. This is particularly valuable in post-crisis recovery periods, where booking curves are unreliable and volatility is high. However, in extremely dynamic contexts (e.g., Dublin's RevPAR), supplementing LSTM with external variables such as event calendars or macroeconomic data could further improve performance.

Conclusion

This study demonstrates the effectiveness of advanced machine learning technique, Long Short-Term Memory (LSTM) networks, in forecasting hotel performance across diverse urban settings. By analysing three key performance indicators, Occupancy (OCC), Average Daily Rate (ADR), and Revenue per Available Room (RevPAR), for five globally significant cities (Manchester, Amsterdam, Dublin, Bangkok, and Mumbai), the research reveals both the strengths and challenges of forecasting in varying market contexts.

Manchester and Mumbai exhibited high forecast accuracy, reflecting more stable and predictable hospitality trends, while Dublin and Bangkok showed greater volatility, driven by seasonal tourism, luxury market dynamics, and international travel dependencies. Amsterdam, though traditionally a strong performer, displayed lingering post-pandemic effects, particularly in RevPAR recovery.

The study confirms that LSTM models, supported by robust time series decomposition and well-structured training/testing strategies, offer a reliable and scalable approach for urban hospitality forecasting. These findings can guide hoteliers, policymakers, and tourism strategists in enhancing operational planning, investment decisions, and post-crisis recovery efforts. Moreover, the model's adaptability across different urban profiles underscores its relevance for global application in data-driven hospitality management.

Acknowledgements:

I would like to express my sincere gratitude to the Manchester Statistical Society (MSS) for providing funding support, which made this project possible. I am also thankful to Manchester Metropolitan University (MMU) for offering the academic environment and resources that enabled me to carry out this research effectively. A special acknowledgement is extended to STR (Smith Travel Research) for generously providing access to their data, which formed the foundation of this analysis.

References

Ampountolas, A. (2019). Forecasting hotel demand uncertainty using time series Bayesian VAR models. Tourism Economics, 25(5), 734–756.

Ampountolas, A. (2021). Modeling and forecasting daily hotel demand: A comparison based on SARIMAX, neural networks, and GARCH models. Forecasting, 3(3), 580–595.

Ampountolas, A. (2024). Addressing complex seasonal patterns in hotel forecasting: A comparative study. Journal of Revenue and Pricing Management, 1–10.

Ampountolas, A., & Legg, M. (2024). Predicting daily hotel occupancy: A practical application for independent hotels. Journal of Revenue and Pricing Management, 23(3), 197–205.

Antonio, N., de Almeida, A., & Nunes, L. (2021). Application of machine learning in hotel demand forecasting: A systematic review. International Journal of Hospitality Management, 94, 102835.

Guillet, B. D., & Mohammed, I. (2015). Revenue management research: Current state and future directions. International Journal of Contemporary Hospitality Management, 27(4), 526–560.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

Huang, L., & Zheng, W. (2023). Hotel demand forecasting: A comprehensive literature review. Tourism Review, 78(1), 218–244.

Law, R., Li, G., Fong, D. K. C., & Han, X. (2020). Tourism demand forecasting: A deep learning approach. Annals of Tourism Research, 80, 102775.

Lewis, C.D. (1982). Industrial and Business Forecasting Methods. London: Butterworths.

Pereira, L. N., & Cerqueira, V. (2022). Forecasting hotel demand for revenue management using machine learning regression methods. Current Issues in Tourism, 25(17), 2733–2750.

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—A review of recent research. Tourism Management, 29(2), 203–220.

Weatherford, L. R., & Kimes, S. E. (2003). A comparison of forecasting methods for hotel revenue management. International Journal of Forecasting, 19(3), 401–415.