

A BIBLIOMETRIC REVIEW OF STATISTICAL MODELING TECHNIQUES IN HOSPITALITY RESEARCH: MALAYSIAN AUTHOR'S PERSPECTIVE (2005-2025)

KEDIN, N. A.^{1*} – BAKHTIAR, M. F. S.² – TALIB, A. H.³

¹ Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Selangor, Malaysia.

² Faculty of Hotel and Tourism Management, Universiti Teknologi MARA, Selangor, Malaysia.

³ Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Melaka, Melaka, Malaysia.

*Corresponding author
e-mail: noradila_kedin[at]uitm.edu.my

(Received 18th October 2025; revised 28th November 2025; accepted 05th December 2025)

Abstract. The applications of statistical modelling techniques as methodological approach to address complex research questions in hospitality research shows a significant evolvement. However, there remains limited systematic evidence on the types of techniques employed and their evolution of time. This study aims to provide a bibliometric review of statistical modelling techniques applied in Malaysian hospitality research, with a focus on identifying most frequently used methods, examining their developments across time. Using Scopus database, bibliometric data were retrieved on 15th August 2025, cleaned and analysed with VOS viewer to visualize keywords co-occurrence and Microsoft Excel to visualize the yearly trends. The results indicate that structural equation modelling (particularly PLS-SEM) dominate the applications, supporting by factor analysis (EFA and CFA), followed by regression analysis. Findings also reveals that the methodological shift from traditional regression approaches to more complex SEM applications in recent years. Furthermore, the findings also reveals the machine learning techniques starts to gain interest among Malaysia authors. This study offers important and valuable insights into the methodological contributions in hospitality research for both academic and industry stakeholders.

Keywords: *bibliometric review, statistical modelling, hospitality research, Malaysian authors*

Introduction

The hospitality and tourism industry is one of the important sectors in Malaysia as it contributes significantly to the economic development and social impact (Kadir and Karim, 2012). A significant increase in revenue (10.5%) contributed by the tourism and hospital industry in 2024 reflects both local and international demand on this industry. Over time, hospitality industry continues to evolves with higher expectation from all parties, led to constant challenges such as changing consumer preferences, adapting technological advancements and quality, restructuring operations and managing customer experiences (Dutta, 2024). Academic research in hospitality that focuses on understanding customer behaviours, service quality, operational efficiency and market trends will help the industry provide evidence-based insights to improve challenges and develop strategies for sustainable growth and competitiveness (Balasubramanian and Hanafiah, 2022). Accordingly, 'Hospitality education differs from other disciplines such as Science Technology, Engineering and Math (STEM) in that hospitality is social science, a discipline that studies the interactions of people'. Hospitality is concerned

with human interactions, guest behaviour, service dynamics, and organizational management – hallmark of social science inquiry. As the hospitality research field demonstrates significant growth, statistical modelling techniques have become prevalent tools among researchers to uncover patterns test theories and predict outcomes in fields like customer satisfaction, booking trends, consumer preferences, sustainability and technology adoption. Statistical modelling in hospitality involves using mathematical and statistical techniques to analyse, understand and create representations through models of the real hospitality field (Assaf and Tsionas, 2019)

Varieties of statistical modelling techniques are widely applied in the social sciences to address a complex research question. The most fundamental statistical modelling techniques is regression analysis encompassing techniques such as linear regression, logistic regression, Probit regression, Poisson regression, Hierarchical regression, Stepwise regression, and Polynomial regression. Regression analysis is a technique used to determine the relationship between one dependent with one or more independent variables (Cohen et al., 2013). Multivariate analysis techniques such as MANOVA, MANCOVA, Discriminant Analysis, and Canonical Correlation on the other hand provide tools to examine a relationship among multiple dependent variables simultaneously. For research with more advanced modelling frameworks, Structural Equation Modelling (SEM) is the best options since SEM has ability to cater latent constructs and more complex relationships through techniques such as Path Analysis, CFA, Covariance-Based SEM, Partial Least Square SEM (PLS-SEM). Other techniques include Survival Analysis (e.g., Cox Proportional Hazards Model, Kaplan-Meier) and Time Series Analysis (e.g., ARIMA, SARIMA, Vector Auto Regression (VAR) which support modelling and forecasting. In addition to these techniques is factor analysis which is a fundamental exploratory tool used in data reduction and construct validation. It serves a preliminary step to validate construct before structural modelling. Factor analysis techniques include Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). Alongside with these traditional statistical modelling is a technique that driven by digitalization, big data and predictive insights and data driven decision making.

Although bibliometric analyses have been conducted to examine publication trends and thematic developments in hospitality and tourism research, there is limited focus on the statistical modelling techniques employed within the field. Therefore, this study aims to fill these gaps by conducting a bibliometric review of statistical modelling techniques applied in hospitality research. Bibliometrics analysis provides strong and influential approach in mapping research trends, identifying significant authors and journals, and uncovering methodological shifts in a given field. This study is driven by two research questions which are: (1) which statistical modelling techniques are most applied in hospitality research among Malaysian Authors? (2) How have the applications of these techniques evolved over time in Malaysian hospitality research? By addressing these questions, this study will offer beneficial insights for academicians, researchers and organizations seeking to understand methodological trends and upcoming directions in hospitality research.

Materials and Methods

This study carried out a bibliometric review approach to analyse publications in Scopus Scientific database focusing on statistical modelling techniques used hospitality

field. Bibliometric review is a quantitative technique that summarize, evaluate and interpret wide-ranging and large quantities of bibliometric data into logical structure and trends on a research topic or field (Van Leeuwen, 2006). Gaining insights and connections between journal citations by summarizing the main topics in the current research and recognizing new issues are the main objective of bibliometric review approach (Low and Siegel, 2020).

Data source and search strategies

Publication data was collected on 14th August 2025 retrieved mainly from Scopus database considering it is one of the reliable databases that provides extensive coverage, stringent quality control and structured bibliometric data accessible in various fields including hospitality and tourism. The Scopus search string was created to capture broad publications related to statistical modelling techniques focusing on traditional and modern statistical method in hospitality research with at least one Malaysian author affiliated with Malaysian institution for the past 20 years. This means collaborative publications with other countries were kept in the dataset. The Scopus search string used throughout the study can be referred in *Figure 1*. Only publications that used English language were counted in to make sure the consistency in the keyword used and interpretation later on. There's no restrictions and limitations in 'document type', 'source type' and 'subject area' to make sure no important publications missed. A total of 1460 publications were retrieved from Scopus database using these search strategies.

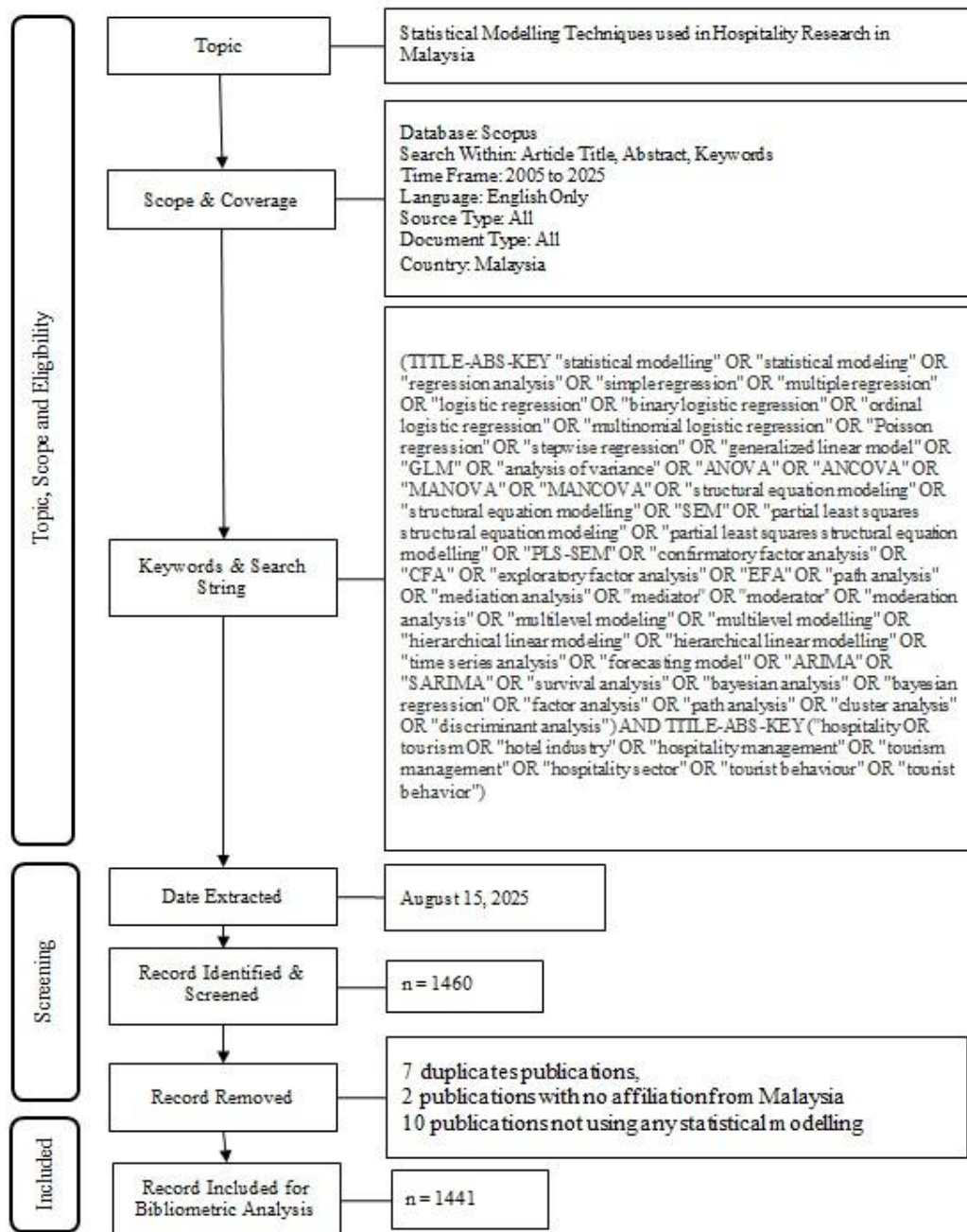


Figure 1. Flow Diagram of Search Strategies.

Data Cleaning

Microsoft Excel and OpenRefine Software were two main tools used in the data cleaning process. Data cleaning process includes removing duplicates, correcting typing error and merge synonyms keyword such as structural equation modelling vs. SEM and filtering and screening publications that did not apply statistical modelling in the publications. If keywords mentioned by authors did not include any statistical modelling techniques used, then the titles and abstract were manually checked to identify relevant techniques. Publications with no affiliation from Malaysian institution were also excluded entirely from the dataset. Among 1460 publications retrieved, 7 duplicates

publications, 2 publications with no affiliation from Malaysia and 10 publications that not incorporate statistical modelling in the study were found and then removed. Remaining 1441 publications will be included and ready for further analysis.

Bibliometric analysis tools

The cleaned dataset was then analysed using VOSviewer software and Microsoft Excel. The VOSviewer software was used to visualise the network keyword co-occurrence while trend analyses were performed using Microsoft Excel to examine the evolution of statistical modelling techniques.

Results and Discussion

This section presents results of the bibliometric review. The results will be presented in three sections to cater the research questions which are: (1) Descriptive overview of the hospitality publications (2005-2025); (2) Publications with application of statistical modelling techniques; and (3) Evolution of statistical modelling techniques (2005-2025).

Descriptive overview of the hospitality publications (2005 – 2025)

This section presents general overview of hospitality research publications authored by Malaysian scholars over 20 years. It highlights the yearly publications trend, and productivity output in the field. The objective of this section is to provide readers a comprehensive background on how the hospitality research has evolve over time from 2005 until 2025.

Yearly publications

Figure 2 shows the publication trends on hospitality research made by Malaysia author from 2005 until 2025 (15th August 2025). The analysis showing a substantial growth on research in hospitality for the past 20 years with total of 6695 documents has been published in Scopus dataset alone. In the first 5 years, publication activity was very low with average less than 100 documents per year but began to steadily increase around 2009. However, on 2012, it can be seen a sudden increase on the number of publications in hospitality research. This sudden rise can be explained by several factors. During that year, Malaysia received about 25 million tourist arrivals which highlighted the importance of the hospitality industry to the economy. This achievement has led to the rising budget allocations to tourism and hospitality industry by the government compared to the previous years. This allocation has encouraged and provide a beneficial opportunity to the academicians and scholars to focus more on hospitality related issues leading to increase spending in research funds and number of active researchers in hospitality research. After the noticeable increase in hospitality research on 2012, publications by Malaysian authors declined in 2013 and remain stagnant through 2014. This slowdown is because to the shifted priorities on the governments budgets fundings for tourism and hospitality as compared to previous years due to global economic uncertainty on 2013 and industry related events such as aviation crisis on 2014. This has directed the research focus also shifted towards other fields such as risk, safety and crisis management. The trend on hospitality research then started to show a consistent increase every year starting from 2015 until 2024. However,

publications in 2025 showing a slight decline as compared to previous year. This decline happens since this bibliometrics data were retrieved on August 2025, which is in the 3rd quarter of 2025, hence the declining publications in 2025 not fully reflect the overall publications in 2025. These findings indicate that the steady increase in hospitality publications is consistent with the growth of the hospitality industry specifically in areas of advanced and more complex innovations. This trend not only highlights the growing academic output in this field but also be a sign of significance importance of Malaysian hospitality industry through scholarly attention and research engagement.

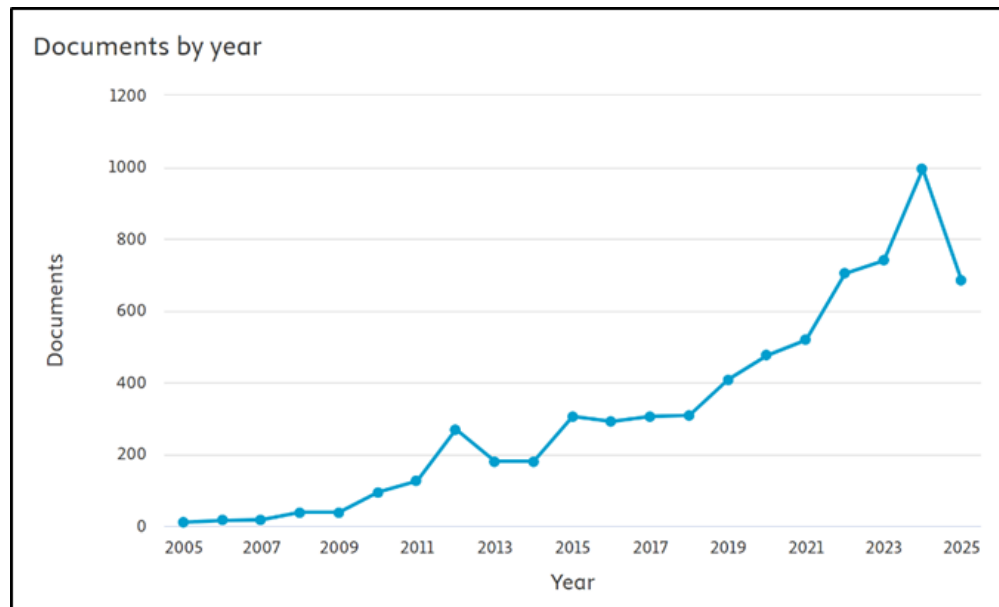


Figure 2. Trend of publications in hospitality research per year (2005-2025).

Document types

Figure 3 shows the breakdown of the documents type in hospitality publications. Majority of the publications were journal articles (69.4%), followed by conference paper (14.1%) and book chapter (10.1%). The remaining percentages were accumulated by other document type such as review, book and notes. This distribution highlights that journal articles have been the most preferred type of publications among Malaysian authors in the hospitality research for the past 20 years. These findings resulted from the fact that journal articles have higher academic values with wider reach and indexing following a thorough peer review process. It offers higher recognition, visibility and impacts to the authors.

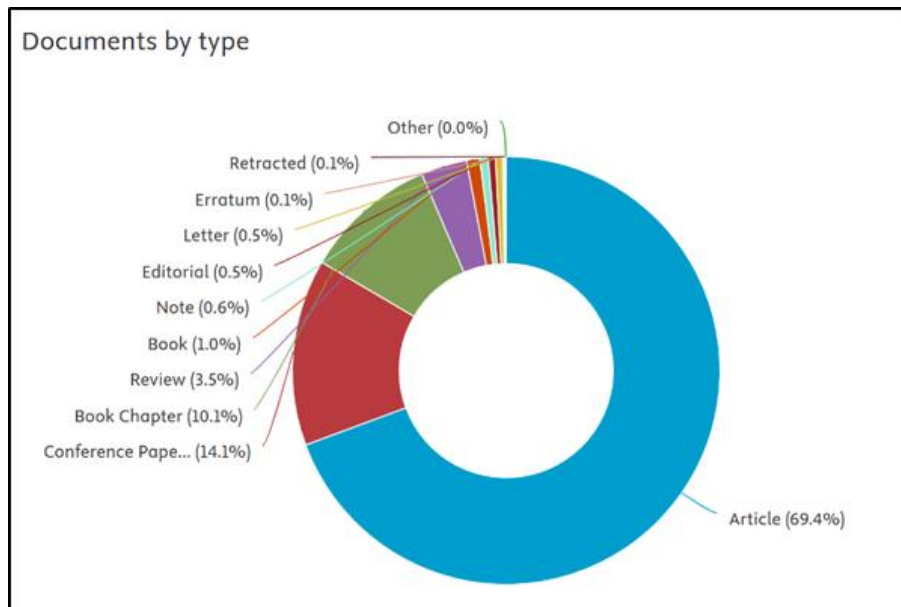


Figure 3. Overall document type.

Publication applying statistical modelling techniques in hospitality research

This section narrows the focus to publication that specifically applied statistical modelling techniques written by Malaysian authors in hospitality research. This aim for this section is to address the first research question regarding statistical modelling techniques are mostly applied in hospitality research among Malaysian authors. The analysis begins by exploring the yearly trends of publication that applied the statistical modelling techniques, and the document's types. Followed by keywords co-occurrence and frequency distributions of occurrence to further identifies the preferred statistical modelling for the past 20 years.

Yearly publications

Figure 4 shows the trend of hospitality publications for the past 20 years (2005 to 2025) that incorporated statistical modelling techniques in their studies. A total of 1441 publications that applied statistical modelling techniques extracted from the Scopus database, which accounted about 21.5% of overall publications on hospitality research by at least one Malaysian authors. In the early 5 years (2005 to 2010), the publications that incorporated statistical modelling techniques is very limited. However, the publication showing a consistent growth starting from 2011 until now. A sharp increase is observed in 2019 onwards with the highest output recorded on 2024 with 245 publications. There's a slight decrease on the publication this year since this bibliometrics data were retrieved on the 3rd quarter which not represent the whole year. These findings showing the significant growing interest on the statistical modelling among Malaysian authors. This rise is driven by the need to understand increasingly complex data-from bookings and guest reviews to smart technology usage-and to forecast critical business outcomes like occupancy and demand (Chuang, 2023). Other than that, this technique provides systematic and strong quantitative approach in identifying relationship, forecasting and prediction and at the same time provides

complex problem solving in which cannot be achieved through simple descriptive statistics.

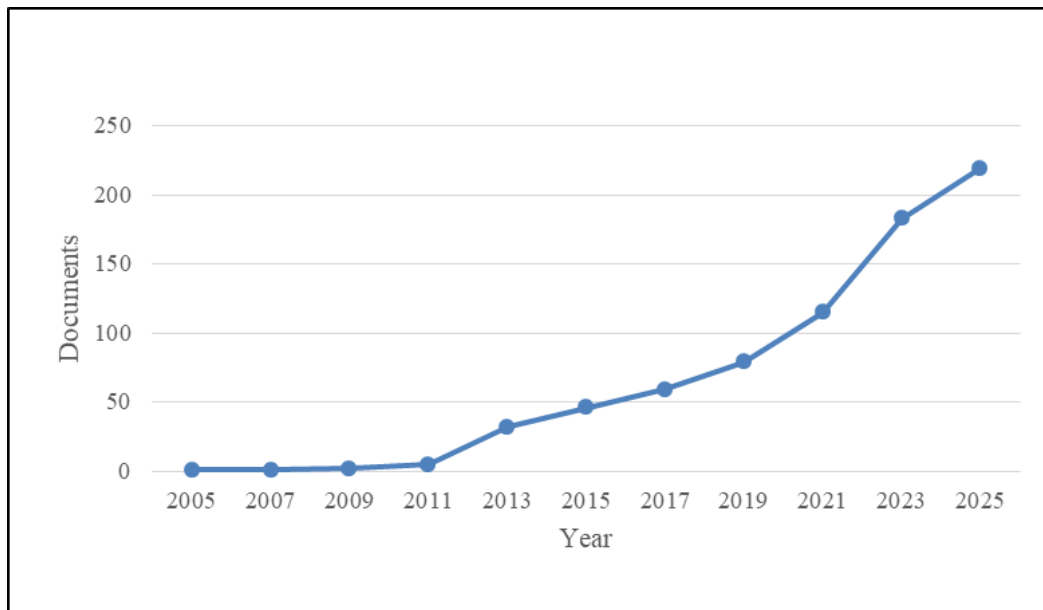


Figure 4. Publications on hospitality research with statistical modelling techniques (2005-2025).

Documents type

Table 1 shows the breakdown of the documents type in hospitality publications that applied statistical modelling techniques. Journal articles are the major contributor with 88.4%, followed by conference paper (6.7%) and book chapter (3.3%). The review papers contribute about 1.3% which is considered low. The remaining 0.3% are from other category such as notes and letters. These findings once again highlight the importance of recognition, visibility and credible approach in publishing hospitality research among Malaysian authors.

Table 1. Document type.

Category	Frequency	Percentage
Article	1274	88.4
Book chapter	47	3.3
Conference Paper	96	6.7
Review	19	1.3
Others	5	0.3
Grand Total	1441	100

Keywords co-occurrence

To further investigate and explore on the prevalent statistical modelling techniques in hospitality research, the author's keywords co-occurrence was analysed. The co-occurrence analysis presented in Figure 5 revealed several dominant groups of statistical modelling approach preferred by Malaysian authors in hospitality research from 2005 to 2025. The biggest node (yellow) situated at the centre of the map represents the most frequent occurrence keywords in publication. PLS-SEM emerges as the most dominant statistical modelling techniques used in the hospitality research. This

Table 2 below represents frequency distributions of statistical modelling techniques used by Malaysian authors on hospitality research for the past 20 years. These frequencies represent the number of keyword occurrence identified in hospitality publication. The distribution of statistical techniques can be classified as core statistical modelling and supporting statistical modelling. Four category of core statistical modelling techniques were identified being used in hospitality research which are Structural Equation Modelling (SEM) family, Regression Techniques, Time Series/ Forecasting and Machine Learning approaches; followed by one category of supporting statistical modelling. SEM family dominate the techniques applied in hospitality research among Malaysian authors where it accounted about 70.9% (n = 1144) of the modelling techniques applied. Among the statistical modelling applied under SEM, partial least squares (PLS-SEM) were the most preferable techniques (n = 646), followed by CB-SEM (n=330) and mediation analysis (n=95). Structural Equation Modelling (SEM) is widely adopted in social sciences because it allows researchers to test complex theoretical frameworks involving multiple latent variables, direct and indirect effects and mediation or moderation simultaneously. Unlike traditional regression, SEM accounts for measurement error and provides a comprehensive approach to theory testing and model validation. Regression techniques are the second leading most popular techniques after SEM, accounted 14.8% of the statistical techniques applied. Two regression techniques that commonly used by Malaysian authors in hospitality research are Multiple Regression (n=64) and Multiple Linear Regression (n = 59). Multiple regression and multiple linear regression are well-established tools that allow researchers to test hypotheses with clear assumptions and widely understood interpretation frameworks. They serve as foundational methods for more advanced modelling approaches, while also being sufficient for many research objectives that focus on prediction or explanation. Time series and machine learning only accounted the remaining with 2.9% and 1.9% of the techniques applied. SARIMA (n=10) and ARIMA (9) are common techniques applied in hospitality research. Although less prominent among Malaysian authors, time series has its own fans because its ability to capture temporal patterns, seasonality and long-term forecasting (Song and Li, 2008; Witt and Witt, 1995) (Figure 5).

Table 2. Frequency of statistical modelling techniques.

Category	Core modelling	Frequency	Percentage
SEM Family	PLS-SEM (646), CB-SEM (342), Mediation (95), Moderation (30), Path Analysis (7), CFA (13),	1,133	73.2%
Regression Techniques	Multiple Regression (64), Multiple Linear Regression (59), Logistic Regression (21), Hierarchical Regression (22), OLS (12), Quantile (12), Simple Linear (7), Panel Data (6), LASSO (2), Bayesian (1), Tobit (1), Canonical (1)	208	13.4%
Time Series / Forecasting	ARDL (5), SARIMA (10), ARIMA (9), GA-SARIMA (1), QARDL (1), VAR (1), Time Series Analysis (13)	40	2.6%
Machine Learning Approaches	ANN (10), RNN (4), CNN (1), Machine Learning (18), Boosted Trees (1)	34	2.2%
Factor & Multivariate Analysis	Factor Analysis (85), EFA (25), ANOVA (12), MANOVA (1), ANCOVA (2), SANCOVA (2), Cluster/K-means (4), Discriminant (1)	132	8.5%
Total		1547	100.0%

Evolution of statistical modelling techniques (2005 – 2025)

This section aims to look into the evolution of the statistical modelling techniques applied from 2005 to 2025 in order to address second research question. Table 3 and Figure 6 represents the summary of table and line chart presents the evolution of

statistical modelling techniques applied in hospitality research across four periods: 2005 – 2009, 2010 – 2014, 2015 – 2019 and 2020 – 2025. In the first period (2005 – 2009), regression analysis, factor analysis and SEM were techniques applied in the early period. During these years, regression analysis techniques dominate in hospitality research compared to SEM techniques. These findings revealed that most of Malaysian authors relies on traditional quantitative modelling approach in hospitality research and at the same time showing interest on the more complex relationships among variables. On the second period (2010 – 2014), SEM, regression analysis and factor analysis remain as the preferred techniques and show a significant rise in the number of applications in hospitality research. During this second periods, SEM and regression application occupied in most of the publications, and time series analysis also started to gain popularity in addition to the current techniques reflecting the interest in the applications of forecasting in hospitality research. For the third period (2015 – 2019), the application of SEM, regression analysis, factor analysis and time series continue to growth in hospitality research. During this years, SEM, regression analysis and factor analysis remain as popular techniques and time series analysis only showing a slow evolution. At this period, attention of the Malaysian authors also started to shift to machine learning techniques, showing the demand and evolution of hospitality research to more holistic and deeper insights. Lastly, during the fourth period (2020 – 2025), all the stated statistical modelling techniques mentioned continues to grow and getting higher popularity among Malaysian authors. However, SEM remains as the most popular techniques preferred as compared to other methods. Since its introduction in the early period, SEM shows tremendous application in every year confirming its positions as the most preferable techniques applied in hospitality research by Malaysian authors. The interest in machine learning in early 2015 shows the transition of hospitality research towards big data analytics. At the same time, regression analysis remains as relevance statistical techniques modelling used for the past 20 years among Malaysian authors because it offers simplicity, robustness and interpretability when examining relationships between variables and the factor analysis as preferable supporting analysis that complement the core modelling. because it offers simplicity, robustness and interpretability when examining relationships between variables. This finding suggest that the hospitality research is also evolving towards forecasting, predictive analytics and data driven decision making.

Table 3. Summary of evolution of statistical modelling applied in four periods.

Year	No. of Publication	Statistical modelling techniques applied (frequency)
2005-2009	11	Regression Analysis (9), Factor Analysis (4) and SEM family (2)
2010-2014	91	SEM family (50), Regression Analysis (34), Factor Analysis (23) and Time Series (6)
2015-2019	300	SEM family (215), Regression Analysis (75), Factor Analysis (27) and Time Series Analysis (9) and Machine Learning (1)
2020-2025	1039	SEM family (870), Regression Analysis (111), Factor Analysis (78), Time Series Analysis (25) and Machine Learning (33)
Total	1441	

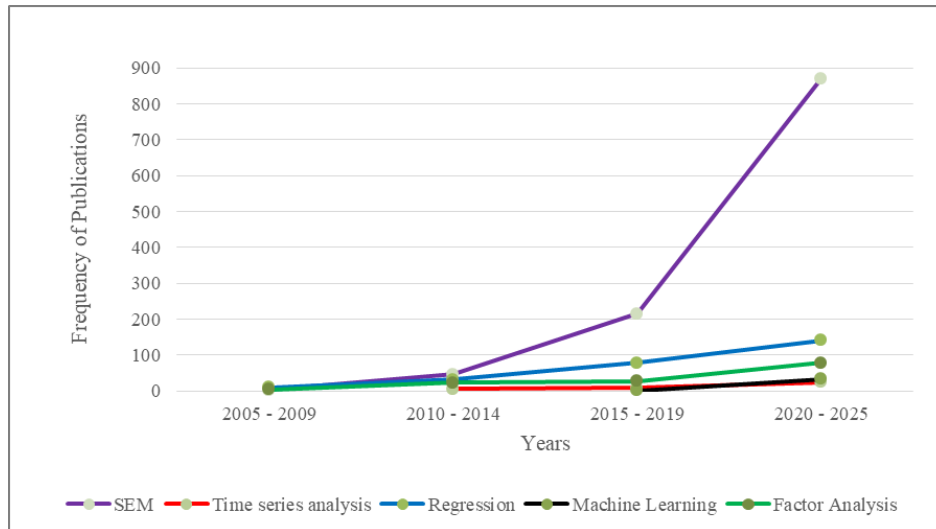


Figure 6. Evolution of statistical modelling techniques applied in hospitality research (2005-2025).

Conclusion

A bibliometric review was used in this study to get an overview of the statistical modelling techniques in hospitality research from Malaysian authors perspectives for 2005 to 2025. Hospitality research can be seen showing a significant increase every year for the past 20 years with a total of 6695 publications incorporating varieties of methodological approaches published mostly in article journal. About 21.5% of the overall publications applied statistical modelling techniques, also showing a rapid growth in applications every year. To address the most applied statistical modelling techniques in hospitality research among Malaysian authors (RQ1), an analysis of keywords co-occurrence with frequency distribution were presented. The findings highlight that structural equation modelling (PLS-SEM/CB-SEM), regression analysis and factor analysis (EFA/CFA) were most widely applied statistical modelling techniques. These approaches are foundational to theory testing, measurement validation and the exploration of causal relationships within hospitality contexts, reflecting the methodological alignment of Malaysian research practices (Hair Jr et al., 2014; Nunkoo et al., 2013).

In relation to RQ2, the analysis reveals a clear evolution in statistical modelling techniques applied over time. Earlier hospitality research relied heavily on basic regression techniques, while more recent works demonstrate a marked shift toward SEM-based approaches, particularly PLS-SEM. This transition suggests an increasing emphasis on testing complex theoretical models, structural relationships and latent constructs, paralleling broader global trends in hospitality and tourism research (Sarstedt et al., 2020; Ali et al., 2018). Despite the positive growth, advanced modelling approaches such as time series or machine learning remains limited in the Malaysia context. Based on these findings, several recommendations can be proposed. First, as the present study relies solely on bibliometric data retrieved from Scopus dataset, future research should incorporate additional reputable sources such as Web of Science (WoS) to enable cross-database comparisons and improve the robustness of the results. Second, since the analysis was limited to publications authored by Malaysian researchers, it is

recommended that future reviews extend to hospitality research conducted in other counties as well. Such comparative analyses would provide a broader perspective on global trends and mythological practices in the hospitality fields.

Acknowledgement

This research is self-funded.

Conflict of interest

The authors confirm that there is no conflict of interest involve with any parties in this research study.

REFERENCES

- [1] Ali, F., Rasoolimanesh, S.M., Sarstedt, M., Ringle, C.M., Ryu, K. (2018): An assessment of the use of partial least squares structural equation modeling (PLS-SEM) in hospitality research. – *International Journal of Contemporary Hospitality Management* 30(1): 514-538.
- [2] Assaf, A.G., Tsonas, M.G. (2019): Quantitative research in tourism and hospitality: an agenda for best-practice recommendations. – *International Journal of Contemporary Hospitality Management* 31(7): 2776-2787.
- [3] Balasubramanian, K., Hanafiah, M.H. (2022): What innovations would enable the tourism and hospitality industry in Malaysia to rebuild? – *Worldwide Hospitality and Tourism Themes* 14(6): 596-603.
- [4] Chuang, C.M. (2023): The conceptualization of smart tourism service platforms on tourist value co-creation behaviours: an integrative perspective of smart tourism services. – *Humanities and Social Sciences Communications* 10(1): 1-16.
- [5] Cohen, J., Cohen, P., West, S.G., Aiken, L.S. (2013): *Applied multiple regression/correlation analysis for the behavioral sciences*. – Routledge 536p.
- [6] Dutta, I. (2024): Adapting to change: Exploring contemporary issues in the tourism and hospitality industry. – *International Journal of Advanced Academic Studies* 6(9): 1-6.
- [7] Hair, J.F., Risher, J.J., Sarstedt, M., Ringle, C.M. (2019): When to use and how to report the results of PLS-SEM. – *European Business Review* 31(1): 2-24.
- [8] Hair Jr, F.J., Sarstedt, M., Hopkins, L., Kuppelwieser, V.G. (2014): Partial least squares structural equation modeling (PLS-SEM) An emerging tool in business research. – *European Business Review* 26(2): 106-121.
- [9] Kadir, N., Karim, M.Z.A. (2012): Tourism and economic growth in Malaysia: Evidence from tourist arrivals from ASEAN-S countries. – *Economic Research-Ekonomska Istraživanja* 25(4): 1089-1100.
- [10] Low, M.P., Siegel, D. (2020): A bibliometric analysis of employee-centred corporate social responsibility research in the 2000s. – *Social Responsibility Journal* 16(5): 691-717.
- [11] Nunkoo, R., Ramkissoon, H., Gursoy, D. (2013): Use of structural equation modeling in tourism research: Past, present, and future. – *Journal of Travel Research* 52(6): 759-771.
- [12] Sarstedt, M., Hair Jr, J.F., Nitzl, C., Ringle, C.M., Howard, M.C. (2020): Beyond a tandem analysis of SEM and PROCESS: Use of PLS-SEM for mediation analyses. – *International Journal of Market Research* 62(3): 288-299.
- [13] Song, H., Li, G. (2008): Tourism demand modelling and forecasting-A review of recent research. – *Tourism Management* 29(2): 203-220.

- [14] Van Leeuwen, T. (2006): The application of bibliometric analyses in the evaluation of social science research. Who benefits from it, and why it is still feasible. – *Scientometrics* 66(1): 133-154.
- [15] Witt, S.F., Witt, C.A. (1995): Forecasting tourism demand: A review of empirical research. – *International Journal of Forecasting* 11(3): 447-475.