Hybrid ARIMA-LSTM Knowledge Management for Strategic Tourism Demand Forecasting in Morocco

Lamiae Benhayoun and Ikram Assaid
International University of Rabat, Morocco

lamiae.benhayoun@uir.ac.ma
ikram.assaid@uir.ac.ma

Abstract: Strategic foresight represents one of the most consequential yet undertheorized capabilities in the knowledge management literature. Whilst organizations increasingly invest in AI-driven forecasting tools, these systems are predominantly designed and evaluated as technical artefacts rather than as knowledge management (KM) instruments. In doing so, they overlook the governance decisions, knowledge architectures, and organizational learning principles that determine whether such tools genuinely enhance an organization's capacity to anticipate and respond to environmental complexity. Tourism management epitomizes this challenge acutely. The sector's inherent complexity, pronounced seasonality, and vulnerability to exogenous shocks, as the COVID-19 pandemic dramatically illustrated, require forecasting approaches that function as organizational knowledge systems rather than mere statistical exercises. This paper addresses this gap by developing and evaluating a hybrid ARIMA-LSTM model for predicting monthly international tourist arrivals in Morocco, framed explicitly within the KM literature on knowledge creation, codification, and strategic foresight. Drawing on a 20-year multivariate dataset (2005–2024) that integrates macroeconomic, climatic, behavioral, and event-driven knowledge assets, five models are systematically compared: ARIMA, SARIMA, XGBoost, LSTM, and the proposed hybrid. The hybrid model dramatically outperforms all alternatives (RMSE = 10.68; MAE = 9.52 vs. RMSE = 89.29 for the next-best approach), demonstrating that architecturally informed knowledge integration produces forecasting gains that no standalone model can achieve. Beyond algorithmic performance, the study advances several theoretical contributions. First, it operationalizes Nonaka and Takeuchi's (1995) combination process, demonstrating that dividing cognitive labor between an explicit linear knowledge layer (ARIMA) and a quasi-tacit nonlinear correction layer (LSTM) constitutes a genuine knowledge recombination strategy. Second, it theorizes data preprocessing as a knowledge governance decision, arguing that the choice of what historical knowledge to retain, discount, or reconstruct is as strategically consequential for organizational learning as the choice of the algorithm to deploy. Third, it reframes multivariate dataset construction as the deliberate assembly of a structured organizational knowledge portfolio distinguishing explicit, behavioral, and tacit-made-explicit knowledge assets. Fourth, this research introduces post-disruption selective reconstruction as a generalizable organizational learning principle applicable beyond tourism to any knowledge-intensive sector navigating structural discontinuities. Implications are drawn for tourism authorities, IS practitioners, and knowledge managers seeking to institutionalize AI-driven foresight within learning organization frameworks.

Keywords: Strategic forecasting, Hybrid ARIMA-LSTM, Tourism demand, AI and KM, Data governance

1. Introduction

Knowledge management has long recognized that the ability to anticipate the future is among the most strategically valuable capabilities an organization can possess (Nonaka & Takeuchi, 1995). In dynamic environments characterized by demand volatility, structural shocks, and behavioral complexity, this foresight capability depends not on intuition alone but on the disciplined conversion of heterogeneous data streams into actionable organizational knowledge. Tourism management represents one of the clearest contexts in which this challenge manifests: destinations must simultaneously interpret macroeconomic signals, climate dynamics, traveler behavior, and geopolitical disruptions to allocate resources, plan infrastructure, and sustain competitive positioning.

Morocco's tourism sector illustrates this tension acutely. Before the COVID-19 pandemic, tourism contributed approximately 19.5% of GDP and generated over two million direct and indirect jobs, underpinned by successive national strategies that positioned tourism as a pillar of the national development model (Didact & Bousstta, 2023; Kapera & Kapera, 2023). By 2019, Morocco welcomed 12.9 million international visitors. Yet the pandemic caused a 78% decline in arrivals in 2020, exposing a critical weakness: the absence of knowledge systems capable of anticipating, absorbing, and recovering from systemic shocks (Ouassou & Taya, 2022). The subsequent recovery, reaching 14.5 million visitors and 105 billion dirhams in revenues by 2023, has not resolved the underlying forecasting deficit. Mega-events, including the 2025 African Cup of Nations and the upcoming 2030 FIFA World Cup, further intensify planning demands (Meziani & Kirmi, 2024).

Traditional approaches to tourism demand forecasting, primarily ARIMA and SARIMA time-series models, have served as the dominant knowledge codification tools in this domain. While interpretable and computationally efficient, they are inherently limited to linear, stationary data structures and cannot incorporate the multisource, heterogeneous knowledge assets that define contemporary tourism intelligence environments (Song et al., 2019; Rekiek et al., 2024). Deep learning models, particularly Long Short-Term Memory (LSTM) networks, address some of these limitations through their capacity to capture long-term temporal dependencies and nonlinear patterns. Yet their data hunger and interpretability constraints make standalone LSTM deployment problematic in disrupted data environments (Teixeira & Gunter, 2023).

This paper argues that a hybrid ARIMA-LSTM architecture, framed as a knowledge management system, offers a structurally superior approach: one that codifies linear organizational knowledge through ARIMA while using LSTM to capture the tacit, nonlinear residual patterns that escape statistical modelling. The study makes four contributions. First, it empirically demonstrates that the hybrid outperforms all standalone alternatives on Moroccan national tourism data (RMSE = 10.68 vs. 89.29 for ARIMA alone). Second, it theorizes forecasting pipeline design as a knowledge governance problem, showing that preprocessing decisions are as consequential as model selection. Third, it introduces a KM lens to interpret multivariate tourism data as organizational knowledge assets both explicit (economic indicators, climate) and tacit-made-explicit (event impact scores, search trends). Fourth, it provides actionable KM implications for tourism authorities, IS practitioners, and learning organizations seeking to institutionalize AI-driven strategic foresight.

The paper is structured as follows: Section 2 reviews the theoretical foundations at the intersection of KM and forecasting. Section 3 describes the data and methodology. Section 4 presents empirical results. We develop the theoretical and managerial contributions in Section 5 and conclude the paper in Section 6.

2. Literature Review

2.1 Forecasting as a Knowledge Management Process

The treatment of forecasting as a purely technical statistical exercise has been progressively challenged by the knowledge management literature. Nonaka and Takeuchi (1995) established that organizational knowledge takes two fundamental forms: explicit knowledge, which is codifiable, transferable, and storable (as in structured data and formal models), and tacit knowledge, which resides in experience, intuition, and contextual judgment. Effective organizational learning depends on the dynamic conversion between these forms, what Nonaka terms the SECI spiral (Socialization, Externalization, Combination, Internalization).

In forecasting contexts, this framework maps directly onto model design: explicit knowledge is captured in statistical structures (ARIMA parameters, feature engineering), while tacit knowledge, i.e. the nonlinear, context-dependent patterns that experienced analysts might intuit but cannot easily articulate, remains largely uncaptured by classical approaches. Machine learning and hybrid AI models represent, from this perspective, an attempt to externalize and codify tacit organizational knowledge embedded in historical patterns, behavioral signals, and systemic interactions.

This conceptualization aligns with Argote and Miron-Spektor (2011), who emphasize that learning organizations develop competitive advantage not merely by accumulating data, but by building routines and tools that convert experience into replicable, scalable knowledge. Forecasting pipelines, from data collection through model deployment to decision support, constitute precisely such routines. Their design, therefore, is a KM design problem, not merely a technical one.

2.2 AI and Knowledge Management in Learning Organizations

The integration of artificial intelligence into knowledge management has generated a substantial body of research exploring both opportunities and risks. AI systems excel at pattern recognition, knowledge extraction from large datasets, and the reduction of human cognitive biases in forecasting (Davenport & Ronanki, 2018). In tourism contexts, this has translated into the adoption of machine learning models capable of processing diverse knowledge inputs, search engine trends as proxies for planning behavior, climate data as seasonal knowledge signals, economic indicators as competitive positioning intelligence (Ouassou & Taya, 2022; Rekiek et al., 2024).

LSTM networks hold particular relevance for organizational knowledge management because of their gated memory architecture. Unlike standard recurrent neural networks, LSTMs selectively retain relevant historical knowledge, encoded in prior arrival patterns, seasonal cycles, and economic conditions, while discarding noise. This selective retention mirrors the KM concept of knowledge filtering: the ability to distinguish signal from noise in complex information environments (Teixeira & Gunter, 2023; Li & Cao, 2018).

However, the KM literature also cautions against AI overdependence. Authentic organizational learning requires human sense-making, contextual judgment, and interpretive capacity alongside algorithmic intelligence (van Wyk, 2024). Hybrid models that combine the interpretability of statistical approaches with the pattern-capturing capacity of deep learning offer a more balanced knowledge architecture, one where human analysts can validate the ARIMA trend component while trusting the LSTM to correct for nonlinear residuals they cannot easily model themselves.

2.3 Data as Organizational Knowledge Assets

The conceptualization of data as a knowledge asset, rather than a raw material, has important implications for forecasting system design. Davenport (2014) distinguishes between data (raw observations), information (structured and contextualized data), and knowledge (actionable insight embedded in organizational routines). From this perspective, the multivariate dataset assembled in this study represents a knowledge portfolio: each data source contributes a distinct type of organizational intelligence, and the integration of these sources through feature engineering constitutes a knowledge combination process in Nonaka's terms.

The post-COVID disruption introduces a further KM challenge: the knowledge embedded in pre-pandemic historical patterns partially loses its validity, requiring organizations to decide which prior knowledge to retain, which to discount, and which to reconstruct. This is fundamentally a knowledge governance problem, one that this study addresses through deliberate preprocessing design, discussed in Section 3.

2.4 From Traditional Forecasting to Hybrid Knowledge Systems

ARIMA and SARIMA models have long served as the dominant knowledge codification tools in tourism forecasting (Song et al., 2019). Their interpretability, i.e. the ability to read off autoregressive and moving-average parameters as explicit statements about temporal structure, aligns well with the explicit knowledge tradition in KM. However, their inability to capture nonlinear relationships, incorporate exogenous variables without significant extension, or adapt to regime changes limits their value as comprehensive knowledge systems (Petrevska, 2017; Rekiek et al., 2024).

XGBoost and other ensemble methods offer richer knowledge extraction from structured features but lack intrinsic temporal memory, making them less suited to sequential knowledge dependencies. LSTM networks, by contrast, function as temporal knowledge repositories, but their black-box character and data requirements constrain their standalone utility in disrupted environments (Nguyen-Da et al., 2023).

Hybrid ARIMA-LSTM models synthesize these strengths into a two-phase knowledge architecture: the statistical layer captures and codifies explicit linear knowledge, while the neural layer extracts the residual tacit knowledge that eludes statistical representation. This design has been validated in tourism forecasting by Zhang et al. (2019), Ouassou and Taya (2022), and Mueller and Sobreira (2024), with consistent evidence of performance superiority over standalone approaches. The present study extends this validation to Morocco's national tourism context over a 20-year horizon, introducing a KM framing that enriches both the theoretical interpretation and the practical implications of these results. Following this literature review, Table 1 summarizes our foundational concepts and their applications in this research through a KM lens.

Table 1: The Hybrid ARIMA-LSTM Model as a KM Architecture

KM Dimension	Concept	Application in This Study
Knowledge Creation	Transforming raw data into actionable organizational intelligence	Multivariate dataset (economic, climate, behavioral) fused into a forecasting model
Knowledge Codification	Explicit capture of patterns for organizational reuse	Hybrid ARIMA-LSTM residual structure encodes linear trends + nonlinear corrections
Knowledge Sharing	Disseminating insights to decision-makers	Forecast outputs inform staffing, budget, infrastructure planning
Organizational Learning	Adapting to environmental disruptions over time	Post-COVID smoothing as a deliberate learning-from-disruption mechanism
Strategic Foresight	Using knowledge assets for proactive decision-making	Scenario-based forecasts enabling resilient tourism planning

3. Methodology

This study aims to demonstrate that a hybrid ARIMA-LSTM model constitutes a superior knowledge management system for strategic tourism demand forecasting, and that the design of such systems is as much a knowledge governance challenge as a technical modelling exercise. In this respect, the hybrid architecture theorized above is operationalized in three stages: assembling the knowledge portfolio, governing its quality, and designing the model architectures as detailed below.

3.1 Collection of Data as a Knowledge Portfolio

The study assembles a monthly multivariate dataset spanning January 2005 to December 2024, drawing on seven complementary sources. Framed as a knowledge portfolio, each source contributes a distinct type of organizational intelligence to the forecasting system. The target variable that is international tourist arrivals, was obtained from the Moroccan Ministry of Tourism and the Haut Commissariat au Plan (HCP). Macroeconomic knowledge assets include the Consumer Price Index (CPI from CEIC/World Bank) and exchange rates (MAD/USD and MAD/EUR from Bank Al-Maghrib), reflecting Morocco's price competitiveness and purchasing power signals for source markets. Climate knowledge is captured through monthly temperature and precipitation averages from Maroc Meteo. Behavioral intelligence is operationalized via a Google Trends composite search index for Morocco Travel queries, serving as a forward-looking proxy for travel intention. Tourism revenue (CEIC Data) proxies economic impact. Finally, a custom-built Event Impact Score encodes the tacit knowledge embedded in major cultural, religious, and sporting events as a deliberate attempt to externalize institutionalized but unstructured organizational memory. Table 2 reports our data collection from diverse sources of organizational knowledge assets.

Table 2: Data Sources as Organizational Knowledge Assets

Source	Variables	Period	Freq.	KM Type
Ministry of Tourism / HCP	Tourist Arrivals	2005-2024	Monthly	Explicit
Maroc Meteo	Temperature, Precipitation	2005-2024	Monthly	Explicit
CEIC / World Bank	Inflation Rate (CPI)	1976-2024	Monthly	Explicit
Bank Al-Maghrib	Exchange Rates (MAD/EUR; MAD/USD)	1999-2024	Monthly	Explicit
Google Trends	Search Volume Index	2004-2024	Monthly	Behavioral
CEIC Data	Tourism Revenue (USD millions)	1998-2024	Monthly	Explicit
Custom-built	Major Events Impact Score	2005-2024	Monthly	Tacit > Explicit

3.2 Knowledge Governance: Preprocessing as a Strategic Decision

Data preprocessing in this study is treated not as a routine technical step but as a deliberate knowledge governance decision, one that determines which historical patterns the organization chooses to learn from, and which it discounts as structurally invalid. The COVID-19 pandemic created a knowledge discontinuity: arrival data for 2020-2021 encodes an extreme disruption with no equivalent precedent in the 15-year pre-pandemic record. Training models directly on this period risks contaminating the knowledge base with patterns that reflect a unique crisis rather than generalizable demand dynamics.

To address this, post-COVID data (2021-2024) was smoothed using a 3% annual growth factor applied to pre-COVID seasonal averages (2017-2019). Stationarity was confirmed via the Augmented Dickey-Fuller test (ADF statistic = -3.949, p = 0.0017). Lagged variables (lag-1, lag-2, lag-3) and 3-month rolling averages were generated as feature engineering inputs. All features were normalized via MinMaxScaler for LSTM and XGBoost. This preprocessing design reflects an organizational learning choice: privileging continuity and generalizability over raw fidelity to a non-representative disruption period.

3.3 Model Architecture as a Knowledge System

Five forecasting models are developed and compared in this study, each conceptualized not merely as a statistical algorithm but as a distinct type of knowledge system, differing in the kind of organizational knowledge it can capture, retain, and deploy.

- ARIMA (1,0,0) — Explicit Knowledge Layer

ARIMA captures the most directly codifiable pattern in the data: the linear relationship between successive monthly arrival figures. It represents the most transparent form of organizational knowledge, structured, interpretable, and auditable by human analysts (best validation RMSE = 92.10). Its limitation is equally transparent: it cannot learn nonlinear patterns, incorporate external variables, or adapt to structural breaks.

- SARIMA (1,0,0)(1,1,1)12 — Seasonal Knowledge Extension

SARIMA extends ARIMA by capturing recurring seasonal patterns including summer peaks, shoulder seasons, and winter troughs. However, the post-COVID smoothing applied in Section 3.2 suppressed precisely the seasonal variability SARIMA depends on, causing it to misinterpret residual noise as cyclical signals (RMSE = 175.76). This illustrates a fundamental KM principle: a knowledge system optimized for one pattern becomes a liability when that pattern disappears from the data.

- XGBoost — Feature-Rich Knowledge Extraction

XGBoost draws on the full portfolio of external knowledge assets, inflation, exchange rates, climate, Google Trends, and event scores, to extract complex nonlinear interactions between variables. Its critical limitation is the absence of sequential memory: it treats each observation independently and cannot learn that what happened three months ago matters in a specific temporal order (RMSE = 260.41).

- LSTM — Temporal Knowledge Repository

The LSTM network is designed to retain and selectively deploy knowledge across long temporal sequences. Its gated memory architecture functions as an organizational memory system that learns which historical patterns predict future demand and which can be discarded as noise, externalizing the quasi-tacit temporal knowledge that experienced analysts might intuit but cannot easily formalize (RMSE = 120.39). Its weakness here is that the COVID disruption shortened its effective training window, limiting the depth of temporal knowledge it could accumulate.

- Hybrid ARIMA-LSTM — Integrated Knowledge Architecture

The hybrid synthesizes both approaches into a three-phase knowledge integration process. First, ARIMA generates the explicit linear baseline forecast. Second, the residuals, i.e. the gap between ARIMA's predictions and actual arrivals — are extracted as the uncaptured knowledge signal. Third, a lightweight LSTM (50 units, 6-step lookback) is trained exclusively on these residuals, learning to correct for the nonlinear patterns ARIMA could not capture. The final forecast combines both layers: ARIMA baseline + LSTM residual correction. This deliberate division of cognitive labor between complementary knowledge systems, one for explicit linear knowledge, the other for quasi-tacit nonlinear knowledge, mirrors Nonaka and Takeuchi's (1995) combination process and explains the hybrid's superior forecasting performance demonstrated in Section 4.

Models are evaluated on 2024 test data using Root Mean Squared Error (RMSE), penalizing large errors critical for capacity planning, and Mean Absolute Error (MAE), providing an interpretable measure of average forecasting deviation expressed in thousands of visitor arrivals.

4. Results

Table 3 reports and compares the performances of the assessed models according to the considered criteria. We explain these results in the following paragraphs.

4.1 Individual Model Performance

ARIMA (1,0,0) delivered a robust linear baseline (RMSE = 89.29, MAE = 68.09), performing most accurately in the second half of 2024 where the smoothed data closely matches its stationarity assumptions. This result confirms that explicit knowledge codification through ARIMA remains valuable when the knowledge environment has been stabilized through deliberate preprocessing.

SARIMA(1,0,0)(1,1,1)12 performed poorly (RMSE = 227.83, MAE = 205.91). The post-COVID smoothing suppressed the seasonal periodicity that SARIMA's knowledge architecture depends on, causing it to misinterpret residual noise as cyclical signals. This outcome illustrates a critical KM principle: a knowledge system optimized for one type of pattern becomes a liability when that pattern is no longer present in the data.

XGBoost, despite its rich feature-based knowledge extraction, achieved the highest errors (RMSE = 260.41, MAE = 191.06). It captured mid-year peaks but systematically overestimated peak summer arrivals, particularly in July and August. This reflects a sensitivity to feature variance in artificially stabilized data — the model extracted knowledge signals that were amplified by the smoothing process, leading to systematic overconfidence. This finding highlights the risk of knowledge extraction models that lack temporal sequential memory.

The standalone LSTM (RMSE = 120.39, MAE = 99.79) outperformed SARIMA and XGBoost but fell short of ARIMA. Its training window was effectively shortened by COVID-era data instability: only the period from 2021-2023 provided reliable sequential patterns, limiting the depth of temporal knowledge it could learn. This reinforces the insight that LSTM's knowledge repository function depends critically on the quality and continuity of the training knowledge base.

4.2 Hybrid Model: Integrated Knowledge Performance

The hybrid ARIMA-LSTM model substantially outperformed all standalone approaches (RMSE = 10.68, MAE = 9.52), an improvement factor of approximately 8x over ARIMA alone. Its two-phase knowledge architecture proved strategically effective: ARIMA absorbed the macro-level explicit trend knowledge, enabling the LSTM to focus exclusively on correcting the nonlinear residual knowledge that ARIMA could not capture. The model achieved near-perfect alignment with actual 2024 values across most months, particularly from June to December.

This result is not merely a technical superiority. It reflects the value of a knowledge integration strategy that deliberately divides cognitive labor between complementary knowledge systems, one optimized for explicit linear patterns, the other for tacit nonlinear corrections. The hybrid's superiority is therefore a function of architectural wisdom as much as algorithmic power.

Table 3: Model Performance — Knowledge System Comparison

Model	RMSE	MAE	KM Fit	Rank
ARIMA	89.29	68.09	Linear trends only	2nd
SARIMA	227.83	205.91	Seasonal patterns — disrupted	4th
XGBoost	260.41	191.06	Feature-rich, no temporal memory	5th
LSTM	120.39	99.79	Temporal dependencies — data-hungry	3rd
ARIMA-LSTM Hybrid	**10.68**	**9.52**	**Full KM pipeline: linear + nonlinear knowledge**	**1st**

Note: Lower RMSE and MAE indicate superior forecasting accuracy. KM Fit column interprets each model's knowledge architecture

5. Discussion

5.1 Theoretical Contributions

- Forecasting as a KM Knowledge Combination Process

The hybrid ARIMA-LSTM architecture operationalizes Nonaka and Takeuchi's (1995) combination process, the synthesis of multiple explicit knowledge sources into a new, more powerful knowledge system. ARIMA encodes the explicit, codifiable knowledge embedded in linear temporal structure. The residual LSTM encodes the harder-to-articulate, quasi-tacit knowledge expressed in nonlinear deviations from that structure. Their combination produces a knowledge system whose forecasting accuracy (RMSE = 10.68) dramatically exceeds what either component could achieve independently, consistent with the KM principle that knowledge recombination generates disproportionate organizational value (Argote & Miron-Spektor, 2011).

This framing extends Zhang et al. (2019) and Ouassou and Taya (2022), who demonstrated hybrid superiority without theorizing it within a KM framework. By doing so, this study positions hybrid forecasting models as KM artefacts — tools through which organizations externalize, combine, and institutionalize knowledge that would otherwise remain dispersed across data silos and individual expertise.

- Preprocessing as Knowledge Governance

A central and undertheorized finding of this study is that the preprocessing design, specifically the decision to smooth post-COVID arrival data using pre-pandemic seasonal baselines, was as consequential for model performance as the choice of algorithm. ARIMA's surprisingly strong baseline performance is explained not by the algorithm's intrinsic power in volatile environments (the literature consistently finds the opposite) but by the knowledge governance decision that stabilized the data before training.

This finding introduces a theoretically important extension to the KM literature on *data-as-knowledge-asset*: the quality of an organization's knowledge assets is not fixed but is actively shaped by the governance mechanisms applied to them. In disrupted environments, knowledge governance, i.e. the deliberate decisions about what to retain, what to discount, and how to reconstruct, becomes as strategically important as the knowledge management tools themselves. This aligns with and extends Davenport's (2014) distinction between data and knowledge, arguing that the conversion of raw data into organizational knowledge is itself a strategic act requiring managerial judgment, not merely technical execution.

Conversely, this same smoothing decision impaired SARIMA and LSTM, whose knowledge architectures depend on precisely the seasonal variability and data richness that smoothing suppresses. The model-data interaction effect identified here, where the same preprocessing design simultaneously enables one knowledge system and constrains another, represents a novel contribution to both the forecasting and KM literatures, with direct implications for how organizations design their data intelligence pipelines.

- Multivariate Tourism Data as a Structured Knowledge Portfolio

This study reframes the dataset construction process as the assembly of a structured knowledge portfolio. Drawing on Nonaka and Takeuchi's (1995) distinction between explicit and tacit knowledge, the seven data sources can be classified along a continuum: economic and climate data represent codified explicit knowledge; Google Trends search indices represent behavioral knowledge partially externalized from collective tacit intent; and the Event Impact Score represents the most ambitious conversion, the transformation of institutionally held, experiential knowledge about event impacts into a quantifiable explicit feature.

This portfolio framing has theoretical implications for how IS researchers conceptualize multivariate forecasting datasets. Rather than treating variable selection as a purely statistical exercise (maximizing predictive power), it foregrounds the organizational knowledge question: which types of knowledge does the organization possess, and how can they be integrated into a coherent intelligence system? The finding that tourism revenue and arrival data are strongly correlated while CPI and exchange rates relate nonlinearly reinforces this: different knowledge assets contribute differently to organizational foresight, requiring knowledge architectures that can accommodate both linear and nonlinear knowledge relationships simultaneously.

- Post-Disruption Organizational Learning as a Forecasting Design Principle

The COVID-19 pandemic created what Argote and Miron-Spektor (2011) would term a knowledge discontinuity, a rupture in the organization's accumulated experience base that renders prior learning partially obsolete. How organizations respond to such discontinuities is a central concern of the organizational learning literature: do they attempt to preserve prior knowledge (conservatism), discard it entirely (over-adaptation), or selectively retain and reconstruct it (adaptive learning)?

This study demonstrates that the selective reconstruction strategy, retaining pre-pandemic seasonal structure while acknowledging the structural break through smoothing and growth-factor adjustment, yields superior forecasting outcomes. This is not merely a technical finding; it is an organizational learning design principle applicable beyond tourism forecasting to any knowledge-intensive sector experiencing post-disruption recovery: financial institutions managing post-crisis data, healthcare systems adapting post-pandemic planning models, or supply chains rebuilding after geopolitical disruptions. The hybrid model's resilience under these conditions validates the adaptive learning approach as both a KM strategy and a forecasting architecture.

5.2 Managerial Implications

- Building a Tourism Knowledge Intelligence Platform

The most fundamental implication of this study is that tourism forecasting must be reconceptualized as a knowledge management infrastructure investment. Moroccan tourism authorities currently operate with fragmented, delayed, and regionally inconsistent data, a structural weakness that no algorithm alone can overcome. The solution is a national tourism knowledge intelligence platform that standardizes data collection, integrates real-time economic, behavioral, and event signals, and creates a shared organizational memory accessible to planners, marketers, and crisis managers alike, with each data source explicitly designated as an organizational knowledge asset with defined collection protocols and update frequencies.

- Embedding AI-Driven Foresight in Organizational Planning Cycles

For hotel managers, national strategists, and destination management organizations, the hybrid model offers a practically deployable forecasting tool that can be embedded directly into quarterly and annual planning cycles. Its two-phase architecture is deliberately interpretable: managers can inspect the ARIMA component to understand baseline demand trajectories and monitor LSTM residual corrections to identify where nonlinear factors, weather anomalies, exchange rate shocks, event-driven surges, are expected to deviate from trend. This interpretability directly addresses the black-box barrier to AI adoption in knowledge-intensive contexts (Davenport & Ronanki, 2018), enabling human sense-making alongside algorithmic intelligence.

- Scenario-Based Forecasting as Organizational Resilience

Because Morocco's tourism demand operates in an environment shaped by geopolitical shocks, pandemic risks, and exchange rate volatility, single-point forecasts systematically underestimate uncertainty. The hybrid architecture enables scenario-based forecasting, adjusting ARIMA baselines and LSTM residual weights to generate probabilistic ranges, communicating uncertainty as an organizational input rather than a modelling failure, with direct application to budget planning, crisis reserve allocation, and infrastructure investment sequencing.

- Data Governance as a Strategic KM Priority

This study elevates data governance from a technical afterthought to a strategic KM priority. Preprocessing decisions are as consequential as algorithmic choices, requiring executive-level attention and cross-functional protocols: explicit preprocessing documentation standards, formal structural break encoding procedures, periodic knowledge continuity audits, and feedback loops that enable systematic organizational learning from forecasting errors.

6. Conclusion and Future Directions

This paper has developed, evaluated, and theorized a hybrid ARIMA-LSTM model for strategic tourism demand forecasting in Morocco, positioning it explicitly as a knowledge management system rather than a purely statistical tool. The empirical results are unambiguous: the hybrid model (RMSE = 10.68, MAE = 9.52) dramatically outperforms all standalone alternatives across a 20-year multivariate dataset spanning one of the most disruptive periods in modern tourism history.

The study's theoretical contributions extend beyond forecasting performance. By framing the hybrid architecture as a knowledge combination process (Nonaka & Takeuchi, 1995), preprocessing as a knowledge governance decision, and multivariate data as a structured knowledge portfolio, this paper argues that AI-driven forecasting systems are fundamentally KM artefacts. The finding that preprocessing quality is as consequential as algorithmic choice introduces a new dimension to both the KM and IS literatures: data governance as a strategic organizational capability with direct implications for practitioners navigating post-disruption recovery.

Several limitations warrant acknowledgment. The post-COVID smoothing may partially inflate hybrid model performance metrics, and future research should test the model on unsmoothed data with explicit structural break indicators to disentangle the contributions of knowledge governance design and algorithmic choice. Transfer learning approaches, pre-training LSTM on comparable tourism markets before fine-tuning on Moroccan data, offer a promising path for expanding the knowledge base, while qualitative fieldwork with Moroccan tourism authorities could validate the KM architecture interpretation empirically.

As organizations increasingly face environments characterized by structural discontinuities and data heterogeneity, integrating KM principles into AI system design will become a practical necessity rather than a theoretical refinement. This study contributes one step toward that integration.

AI Declaration: The authors used the AI tool ChatGPT (OpenAI, version GPT-5.5) strictly for language refinement.

Ethics Declaration: This study did not require ethical approval because it exclusively relied on aggregated secondary data obtained from publicly accessible institutional and digital sources, including governmental statistics, macroeconomic indicators, climate records, and online search trend data. No human participants, personal identifiable information, or sensitive individual-level data were involved in the research process.

References

Argote, L., & Miron-Spektor, E. (2011). Organizational learning: From experience to knowledge. Organization Science, 22(5), 1123-1137.

Davenport, T. H. (2014). Big data at work: Dispelling the myths, uncovering the opportunities. Harvard Business Review Press.

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. Harvard Business Review, 96(1), 108-116.

Didast, F., & Bousstta, O. (2023). The influence of tourism on the Moroccan economy. African Scientific Journal.

Intarapak, S., Supapakorn, T., & Vuthipongse, W. (2022). Classical forecasting of international tourist arrivals to Thailand. Journal of Statistical Theory and Applications.

Kapera, I., & Kapera, A. (2023). Determinants, directions, and challenges of tourism development in Morocco. Geography and Tourism, 11(1).

Kapera, I., & Kapera, A. (2024). Sustainable tourism and environmental challenges in Morocco. Tourism Management.

Khatri, B. P. (2024). Trend analysis of tourist arrivals in Nepal using ARIMA model. Journal of Tourism and Adventure.

Laaroussi, H., Guerouate, F., & Sbihi, M. (2022). A novel hybrid deep learning approach for tourism demand forecasting. International Journal of Electrical and Computer Engineering.

Li, Y., & Cao, H. (2018). Prediction for tourism flow based on LSTM neural network. Procedia Computer Science.

Meziani, B., & Kirmi, B. (2024). Development of Morocco's tourism sector: Economic impact and preparations for international sporting events. International Journal of Applied Management and Economics.

Mueller, R., & Sobreira, N. (2024). Tourism forecasts after COVID-19: Evidence of Portugal. Annals of Tourism Research Empirical Insights.

Nguyen-Da, T., Li, Y.-M., Peng, C.-L., Cho, M.-Y., & Nguyen-Thanh, P. (2023). Tourism demand prediction after COVID-19 with deep learning hybrid CNN-LSTM. Sustainability (MDPI).

Nonaka, I., & Takeuchi, H. (1995). The knowledge-creating company. Oxford University Press.

Ouassou, E. H., & Taya, S. (2022). Forecasting regional tourism demand in Morocco from traditional and AI-based methods to ensemble modeling. Forecasting.

Petrevska, B. (2017). Predicting tourism demand by ARIMA models. Economic Research - Ekonomska Istrazivanja.

Rekiek, S., Jebari, H., & Reklaoui, K. (2024). Prediction of booking trends and customer demand in the tourism and hospitality sector using AI-based models. IJACSA.

Song, H., Qiu, R., & Park, J. (2019). A review of research on tourism demand forecasting. Annals of Tourism Research.

Teixeira, J., & Gunter, U. (2023). Editorial for special issue: Tourism forecasting: Time-series analysis of world and regional data. Forecasting (MDPI).

Tovmasyan, G. (2021). Forecasting the number of incoming tourists using ARIMA model: Case study from Armenia. Marketing and Management of Innovations.

van Wyk, B. (2024). Agentic AI and knowledge management in learning organizations. University of Pretoria Working Paper.

Zhang, B., et al. (2019). The comparison of forecasting analysis based on the ARIMA-LSTM hybrid models. Sustainability.