International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(7s) 2026 pp. 957-974
https://doi.org/10.70917/ijcisim-2026-3167
Article

Air-Access Stress as an Entrepreneurial Risk Signal: Interpretable Machine Learning for Tourism Enterprise Resilience in Philippine Island Gateways

Karen Razelle M. Duyan [1,2,*], Juan Moshe M. Duyan [1,2], Valerie A. Abesamis Paloy [1], Eric A. Paloy [1], Melanie S. Manuel [1], Mark Joby Aguilar[1], Joe Robert G. Lucena [2,3], Luigi Carlo M. De Jesus [2,3]

[1]Kalinga State University, Tabuk City, Kalinga, Philippines; kmduyan@ksu.edu.ph; jmduyan@ksu.edu.ph; vaabesamis@ksu.edu.ph; eapaloy@ksu.edu.ph; msmanuel@ksu.edu.ph
[2]O.P. Jindal Global University, Sonipat, Haryana, India; karen.razelle_fs@jgu.edu.in; juan.moshe_fs@jgu.edu.in; luigi.carlo_fs@jgu.edu.in; joe.robert_fs@jgu.edu.in
[3]School of Engineering, Asia Pacific College, Makati City, Philippines; luigid@apc.edu.ph; joerobertpaull@apc.edu.ph
*Correspondence: kmduyan@ksu.edu.ph; karen.razelle_fs@jgu.edu.in

Abstract: Tourism entrepreneurship in island destinations depends on reliable transport access because disruptions can quickly affect customer flows, bookings, cash flow, staffing, procurement, supplier commitments, and service delivery. This study evaluates whether abnormal air-passenger decline can function as an upstream entrepreneurial risk signal for tourism enterprises and whether interpretable machine learning adds decision value beyond simple operational rules. Civil Aviation Authority of the Philippines records for 2018-2025 were transformed into 714 observed gateway-months across eight tourism gateways, with blank cells retained as unobserved and explicit no-flight records coded as zeros. Models trained on 2018-2021 data were tested out of time on 2022-2025 observations against persistence and seasonal-naive baselines. Severe access-stress exposure was strongly negatively associated with tourism direct gross value added share of GDP (Spearman rho = -0.952, $p < 0.001$), while passenger growth was positively associated with TDGVA growth (rho = 0.929, $p = 0.003$); these annual associations are exploratory and pandemic-dominated. Random forest ranked next-month stress risk well (AUC = 0.940), but persistence achieved the higher F1 score (0.708 versus 0.400). SHAP showed that passenger scale, volatility, and momentum drove predictions. The findings support a hybrid entrepreneurial decision-support approach: machine learning for prioritization and onset surveillance, persistence for ongoing episodes, and local booking, revenue, occupancy, and operating indicators for action. The study positions interpretable AI as a managerial tool for tourism-enterprise resilience rather than as an end in itself.

Keywords: tourism entrepreneurship; tourism enterprises; MSME resilience; entrepreneurial risk; island destinations; air access; business continuity; interpretable machine learning; SHAP; decision support

1. Introduction

Island and archipelagic destinations depend on transport access to convert tourism interest into actual visits, transactions, and service consumption. In the Philippines, air gateways connect source markets with hotels, restaurants, tour operators, transport providers, local attractions, and other tourism enterprises. A sudden decline in gateway traffic can therefore become an entrepreneurial and operating shock: rooms and tour capacity cannot be stored, staffing schedules are difficult to reverse at short notice, and procurement, supplier orders, pricing, promotions, working-capital needs, and local transport demand may all respond to changing visitor flows. Air access is thus not equivalent to tourism demand, but it is an enabling condition through which tourism entrepreneurs can reach customers and realize revenue [1, 2, 3].

This distinction matters for tourism-enterprise resilience. A high-volume gateway is not necessarily resilient, and a stable passenger series is not by itself evidence that entrepreneurs, workers, and destination businesses can absorb and recover from disruption. The concern is especially important for small tourism firms with limited staffing, inventory, liquidity, and market-diversification buffers. Recent tourism resilience research increasingly treats resilience as a dynamic, multidimensional process of resistance, adaptation, organizational learning, and recovery over time rather than as a fixed destination attribute [4, 5, 6]. This process is shaped by crisis-management capacity, structural conditions, policy support, and the ability of tourism enterprises to adjust strategies under uncertainty [7, 8, 6].

The second problem is methodological and managerial. Predictive models can appear strong when they are compared only with other machine-learning algorithms, yet entrepreneurs need tools that improve decisions relative to transparent, low-cost rules. Passenger movement is autocorrelated and shock episodes often persist across consecutive months, so a simple rule such as "next month will remain stressed if this month is stressed" can be difficult to beat. Without persistence and seasonal-naive baselines, high AUC or F1 values may overstate the practical contribution of machine learning. Recent tourism studies use interpretable machine learning while keeping tourism demand or performance as the substantive outcome [9, 10]. In the present study, model-attribution outputs are interpreted as predictive associations that support managerial judgment rather than as causal effects or automated business instructions.

The substantive contribution is to test whether abnormal air-access stress functions as an upstream tourism-enterprise risk signal rather than merely as an aviation anomaly. The empirical design asks whether such a signal can be detected one month ahead, whether interpretable machine learning adds value beyond simple operational rules, and whether the aggregate signal co-varies with official tourism-economic performance. To answer these questions, the study reconstructs a multi-year gateway-month panel from public aviation records, benchmarks three machine-learning models against persistence and seasonal-naive rules, uses SHAP for case-level explanations, evaluates cross-gateway transferability, and validates annual access-stress indicators against Philippine Tourism Satellite Account outcomes. This design places predictive analytics in the service of entrepreneurial market intelligence, operating continuity, resource allocation, and recovery planning. It also demonstrates how entrepreneurship research can convert public administrative data into decision support for tourism enterprises while maintaining clear boundaries between macro-level signals and firm-level outcomes.

Accordingly, this study addresses four research questions:

RQ1: How did abnormal air-access stress evolve across selected Philippine tourism gateways from 2018 to 2025?

RQ2: How much one-month-ahead predictive value do machine-learning models add over persistence and seasonal-naive baselines for enterprise-oriented risk monitoring?

RQ3: Which observable operating conditions drive random-forest predictions according to SHAP, and how transferable are the learned relationships across gateways?

RQ4: Do aggregate air-access stress patterns co-vary with tourism-specific economic outcomes relevant to the operating environment of tourism enterprises?

2. Literature Review and Conceptual Framework

2.1. Air Access, Tourism Enterprises, and Market Operations

The relationship between aviation and tourism is reciprocal and has direct implications for entrepreneurial market access. Papatheodorou [1] reviews how air services and tourism demand co-evolve through accessibility, market development, and policy. Mazzola et al. [2] show that air transport is closely linked to tourism flows in island settings, while Tang et al. [3] estimate positive tourism effects from international air-route expansion. For the Azores, Silveira et al. [11] demonstrate how air-transport liberalization altered travel and tourism patterns. In an archipelago, air transport may be the only time-efficient option for long-haul visitors and short stays, making gateway performance part of the market infrastructure on which tourism ventures depend.

For entrepreneurs and managers, the value of gateway information arises from timing. Passenger movement does not identify travel purpose and includes residents, workers, business travelers, visiting friends and relatives, and leisure tourists. Nevertheless, abrupt access changes can alter the pool of potential customers and affect booking expectations, cancellation exposure, staffing, inventory, supplier orders, transport coordination, pricing, promotional timing, and short-term cash-flow planning. The useful construct is therefore not "tourism demand measured by airport passengers," but "air-access stress as an upstream market-access signal for tourism-enterprise decisions."

2.2. Tourism-Enterprise Resilience as a Shock-Response-Recovery Process

The conceptual lineage of resilience distinguishes persistence from adaptive change. Holling [12] framed ecological resilience as the capacity of a system to persist under disturbance rather than simply return rapidly to equilibrium. Folke et al. [13] extended this perspective by emphasizing adaptability and transformability in social-ecological systems. Tourism scholarship has adopted this process view: Wang et al. [4] document the field's shift from recovery-centered definitions toward resistance, adaptation, and restructuring, while Liu et al. [5] develop a tourism and hospitality management framework organized around resilience and recovery processes. From an entrepreneurship perspective, resilience also concerns how firms protect essential operations, learn from shocks, reconfigure resources, and preserve market relationships. Duro et al. [14] show that tourism impacts and recovery are territorially heterogeneous; Bertella [15] emphasizes organizational learning paths during crisis; and Estiri et al. [7] examine policy responses supporting tourism-SME resilience. Hall et al. [16], Lee and Pennington-Gray [17], Smolčić Jurdana and Agbaba [8], and Qu et al. [6] further emphasize transitions, crisis-management capability, measurement, and actionable interventions.

The shock-response-recovery perspective adopted here is grounded in disaster, infrastructure, and economic resilience measurement. Bruneau et al. [18] formalize resilience as time-varying system performance after disruption, while Rose [19] develops complementary economic-resilience concepts centered on loss reduction and recovery. Building on this lineage, the study characterizes an access shock by the depth of performance loss, duration of impaired functioning, cumulative loss, and movement toward a recovery threshold. These quantities are applied narrowly to the access layer and interpreted as conditions in the external operating environment of tourism enterprises. Recovery months retrospectively describe observed passenger movement after a detected event; they are not predicted and are not interpreted as full destination recovery, firm recovery, or venture survival.

2.3. Interpretable Machine Learning and Entrepreneurial Decision Support

Tourism forecasting studies increasingly combine non-linear methods with data that reflect the timing and context of demand. Viverit et al. [20] use machine learning to cluster hotel booking curves before forecasting demand. Contessi et al. [21] propose an interpretable hotel-occupancy forecasting design that uses principal component analysis to compress booking-curve information. Huang and Zhang [22] evaluate an iTransformer for daily tourism demand, whereas Hu et al. [23] incorporate short-video information into forecasting. Lee [24] integrates web-search data into a SARIMAX model, and Vasenska [25] benchmarks machine-learning and deep-learning models using economic indicators. For entrepreneurship and industry management, the relevance of these methods lies not in algorithmic novelty alone, but in whether available information can improve market sensing, contingency planning, and resource decisions before the forecast horizon.

Interpretability is essential when analytical systems are intended for entrepreneurial and managerial use. Hou and Wang [26] propose a VSN-xLSTM model whose variable-selection mechanism exposes predictive drivers. Ghosh [9] uses XGBoost with SHAP and accumulated local effects to explain constraints on inbound tourism demand, while Oruç Erdoğan et al. [10] apply machine learning and SHAP to international tourism demand and environmental indicators. These precedents are methodologically relevant because the model serves a substantive management problem and attribution is treated as predictive explanation rather than causal effect. Interpretability is also distinct from ordinary feature importance: managers need to know whether a high-risk signal is associated with declining year-on-year traffic, rising volatility, a low passenger base, or another observed condition before translating the warning into staffing, inventory, supplier, pricing, or communication decisions.

Benchmarking is equally important for responsible entrepreneurial decision support. Persistence is a natural competitor when stress episodes span several months, while a seasonal-naive rule tests whether recurring calendar patterns are sufficient. A machine-learning model should therefore be judged not only by discrimination but also by whether its threshold-level precision, recall, and F1 improve on simple, transparent, and inexpensive rules. This study treats a failure to beat a naive baseline as a substantive result rather than hiding it, because an entrepreneur should not adopt a more complex system unless it adds practical value or provides explanations unavailable from the simpler alternative.

2.4. Entrepreneurship-Centered Conceptual Model

Figure 1 summarizes the entrepreneurship-centered conceptual model. The access layer generates market and operating signals from passenger scale, carrier structure, momentum, volatility, and seasonality. A one-month risk model converts these inputs into a predicted probability of severe access stress. SHAP and cross-gateway validation assess interpretation and transferability. The enterprise-decision layer shows how a warning may prompt review of bookings, staffing, inventory, supplier commitments, capacity, cash-flow exposure, and continuity arrangements. The tourism-economic layer then asks whether aggregate access stress co-varies with measured tourism outcomes. The boundary shown at the bottom of the figure is central: access monitoring can support entrepreneurial judgment but cannot substitute for tourist arrivals, bookings, occupancy, sales, expenditure, or firm-level performance data.

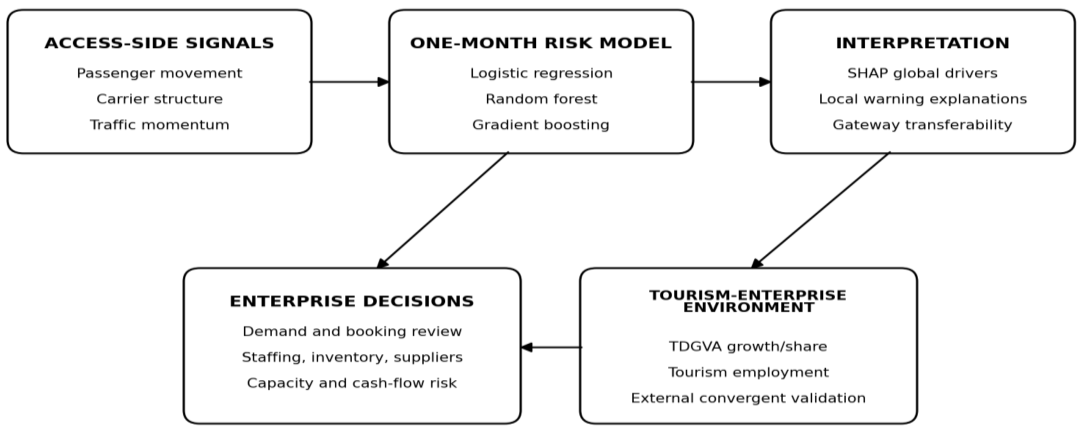

Figure 1. Entrepreneurship-centered conceptual framework linking air-access stress, one-month risk modeling, interpretable machine learning, tourism-enterprise decisions, and external tourism-economic validation.

3. Materials and Methods

3.1. Research Design

The study used a longitudinal, secondary-data design with the gateway-month as the primary unit of analysis. The empirical workflow contained five stages: (1) extract and harmonize monthly passenger and operator records; (2) identify observed versus unobserved months; (3) calculate access-stress events relative to gateway-specific same-month history; (4) evaluate one-month-ahead classification models against naive baselines; and (5) externally validate annual access indicators against PTSA outcomes. The design is predictive and diagnostic, not causal. Its intended management function is to develop an upstream risk-screening signal that can be combined with local enterprise data rather than to estimate firm-level demand, revenue, or survival.

3.2. Data Sources

Table 1 separates data actually used in this study from possible future extensions. The primary access panel came from annual aircraft, passenger, and cargo movement workbooks published by the Civil Aviation Authority of the Philippines [27]. Tourism-specific outcomes came from the Philippine Statistics Authority (PSA) Philippine Tourism Satellite Account statistical tables [28]. A short-window tropical-cyclone layer from NOAA IBTrACS [29] and airport coordinates from OurAirports [30] were retained only as a pilot extension and were excluded from headline model estimates because the available cyclone subset covered 2023-2025 rather than the full 2018-2025 panel.

Table 1. Public data sources used in the study.

Source	Study window	Variables	Role in analysis
CAAP annual movement workbooks	2018-2025	Monthly passenger movement by airport and operator	Primary gateway-month access panel
PSA Philippine Tourism Satellite Account	2018-2025 subset	TDGVA growth; TDGVA share of GDP; tourism employment share	External tourism-outcome validation
NOAA IBTrACS v04r01	2023-2025 pilot subset	Cyclone track position and intensity observations	Pilot extension; excluded from headline model
OurAirports open data	Current coordinate file used in pilot	Airport latitude and longitude	Spatial join for cyclone pilot

3.3. Gateway Selection and Missingness Protocol

The analysis covers eight tourism-relevant gateways that could be consistently matched across CAAP workbooks: Bacolod, Bohol/Panglao, Boracay/Caticlan, Davao, Iloilo, Kalibo, Laguindingan/Cagayan de Oro, and Siargao. The selection is a data-availability sample, not a census of Philippine airports. Airport sections with international and domestic subsections were combined at gateway level, and operator names were standardized before aggregation.

A conservative reporting protocol was applied. A gateway-month was treated as observed when at least one cell in the corresponding gateway section contained a numeric value or an explicit no-flight statement. Explicit text such as "No Commercial flight due to COVID-19 PANDEMIC" was coded as an observed zero because the source directly reported no passenger movement. Blank cells were retained as unobserved and did not generate zero-passenger records. Within an observed gateway-month, blank cells for individual operators were treated as no recorded activity for that operator. This distinction is important because the 2024 and 2025 workbooks contain uneven reporting windows across gateways.

After this protocol, the panel contained 714 observed gateway-months and 95,660,648 reported passengers. The data were unbalanced: Davao, Iloilo, and Bacolod each contributed 96 observed months, while other gateways had shorter series because of source availability. General aviation and military rows were excluded from commercial passenger totals and carrier concentration measures.

3.4. Access-Stress and Recovery Characterization

For gateway g in month t, expected passenger movement was defined as the median of up to five previous observations for the same gateway and calendar month, with a minimum of two historical observations required. This same-month baseline reduces ordinary seasonality without fitting a high-parameter time-series model to a small number of annual cycles. Shock depth was calculated as:

$$\mathbf{ShockDepth(g,t) = [Passengers(g,t) - Expected(g,t)] / Expected(g,t),} \quad \mathbf{(1)}$$

A gateway-month was classified as severe access stress when shock depth was less than or equal to -0.30. The threshold indicates at least a 30% shortfall relative to historical same-month expectation. The mathematical minimum of the ratio is -1.00 when recorded passengers equal zero and the expected baseline is positive; no clipping rule was applied. Consecutive severe-stress months formed an event only when they were adjacent calendar months. A gap in source reporting ended the event sequence.

Recovery was characterized descriptively as the first subsequent observed month in which passenger movement reached at least 90% of its same-month expected value. This variable is retrospective and is used only to describe event trajectories. The study does not claim to predict recovery time.

3.5. One-Month Prediction Task, Baselines, and Models

The prediction task estimated whether severe access stress would occur in the next observed calendar month. A target was created only when the exact next calendar month was available; row-wise shifting across reporting gaps was not permitted. Predictors available at warning time were log passenger movement, active carrier count, carrier HHI, year-on-year passenger growth, month-on-month passenger growth, rolling three-month volatility of log passenger movement, and sine and cosine terms for calendar seasonality. Carrier diversity was omitted because it is deterministically related to HHI.

The development period covered 2018-2021 (n = 360 usable records; 155 positive targets), and the out-of-time evaluation period covered 2022-2025 (n = 341; 20 positive targets). The split intentionally places the post-pandemic recovery and recent operating period entirely outside model training. A stricter 2024-2025 sensitivity holdout was also examined, but it contained only three positive targets and is therefore not used for headline claims.

Three models were compared: class-balanced logistic regression, a class-balanced random forest, and gradient boosting fitted with balanced sample weights. Missing predictor values were imputed using medians estimated from the training period only. No test-period values were used for imputation or model fitting. Model probabilities were converted to binary warnings at a 0.50 threshold.

Two naive baselines were evaluated on the same test observations. The persistence baseline predicted next-month's stress when the current month was already in stress. The seasonal-naive baseline predicted next-month stress when the same gateway had been in stress in the corresponding target month one year earlier. These rules test whether autocorrelation and recurring seasonality explain most of the apparent predictive skill.

3.6. Performance, Uncertainty, Transferability, and SHAP

Performance was reported using area under the receiver operating characteristic curve (AUC), average precision (AP), balanced accuracy, precision, recall, and F1. Gateway-cluster bootstrap confidence intervals were generated by resampling the eight gateways with replacement 500 times, preserving within-gateway temporal dependence. Because the number of clusters is small, the intervals are interpreted as descriptive uncertainty bands rather than exact frequentist coverage guarantees.

Spatial transferability was assessed through leave-one-gateway-out validation. For each gateway, the random forest was trained on all other gateways through 2021 and evaluated on the held-out gateway during 2022-2025. AUC was not reported when a held-out gateway contained no positive test events. This design directly tests whether the model can transfer beyond gateway-specific baselines.

SHAP values were computed for the random forest. Global importance was measured by mean absolute SHAP value across the out-of-time test set. Local explanations were then extracted for two correctly identified stress onsets and one missed onset, allowing the analysis to show not only which variables mattered overall, but why specific warnings succeeded or failed. Consistent with recent tourism applications of SHAP, the values are used as model-attribution measures for prediction explanation rather than as estimates of causal effects [9, 10].

3.7. External Tourism-Economic Validation

To assess whether the access-side signal is connected to the broader operating environment of tourism enterprises, annual access indicators were merged with official PTSA outcomes for 2018-2025 [28]. The access indicators were the share of observed gateway-months classified as severe stress, annual passenger growth, and mean shock depth. The tourism outcomes were PTSA Tourism Direct Gross Value Added (TDGVA) growth, TDGVA share of GDP, and tourism employment share. Spearman rank correlations were used because the validation layer contains only eight annual observations and includes the exceptional COVID-19 shock. Sample sizes vary across predictor-outcome pairs because of variable construction: stress-month share is available for all eight years, annual passenger growth is undefined for the first panel year and therefore uses seven observations, and mean shock depth requires at least two historical same-month observations and therefore uses six years. Four focal correlations are reported without multiplicity correction and are treated as exploratory associations. These associations are convergent validity checks at the tourism-economy level and should not be read as causal estimates, gateway-level relationships, or direct measures of enterprise performance.

3.8. Pilot Tropical-Cyclone Layer

A 2023-2025 NOAA IBTrACS subset [29] was spatially joined to gateway coordinates from OurAirports [30] using a 300 km radius to demonstrate how cyclone exposure could be added to the framework. Because the pilot window does not cover the full modeling period, cyclone variables were excluded from the primary model and from headline performance estimates. The pilot is retained only to document an extensible climate-risk pathway for future work with the complete historical IBTrACS record.

4. Results

4.1. Corrected Panel and Access-Stress Trajectories

The corrected panel contains 714 observed gateway-months, rather than a mechanically balanced 8 × 96 matrix. The selected gateways recorded 95.66 million passengers from 2018 to 2025. Davao accounted for the largest cumulative total (25.80 million), followed by Iloilo (16.44 million) and Laguindingan/Cagayan de Oro (12.08 million). Table 2 reports both scale and coverage so that large totals are not interpreted without considering unequal observation windows.

Table 2. Corrected gateway panel summary, 2018-2025.

Gateway	Observed months	Reported passengers	Mean monthly passengers	Mean active carriers	Mean HHI
Davao	96	25,802,111	268,772	7.35	0.334
Iloilo	96	16,440,480	171,255	4.48	0.375
Laguindingan/Cagayan de Oro	87	12,076,603	138,812	4.00	0.376
Boracay/Caticlan	77	11,612,126	150,807	5.22	0.399
Bacolod	96	11,403,071	118,782	3.75	0.384
Bohol/Panglao	76	7,908,112	104,054	5.33	0.327
Kalibo	94	7,591,168	80,757	6.54	0.299
Siargao	92	2,826,977	30,728	2.49	0.499

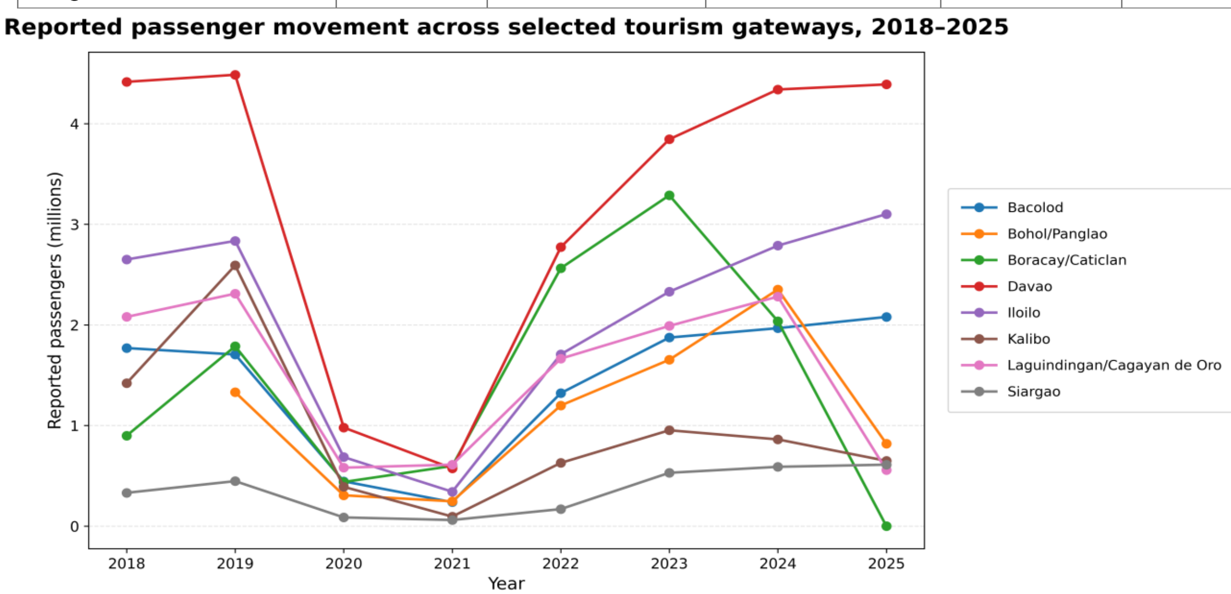

Figure 2. Reported passenger movement across selected Philippine tourism gateways, 2018-2025. Missing gateway-months are not plotted as zeros.

The dominant system-wide stress occurred during 2020-2021, but the corrected event logic also detected later gateway-specific episodes. Twenty severe-stress events were identified. The longest continuous events were observed in Davao (25 months, March 2020-March 2022), Bacolod and Laguindingan/Cagayan de Oro (24 months each), and Kalibo (22 months, July 2020-April 2022). In later years, the principal severe episodes were shorter and more localized, including Kalibo in early 2023 and May-July 2025. Figure 3 shows that the share of observed months in severe stress fell sharply after 2022, while remaining non-zero for specific gateways.

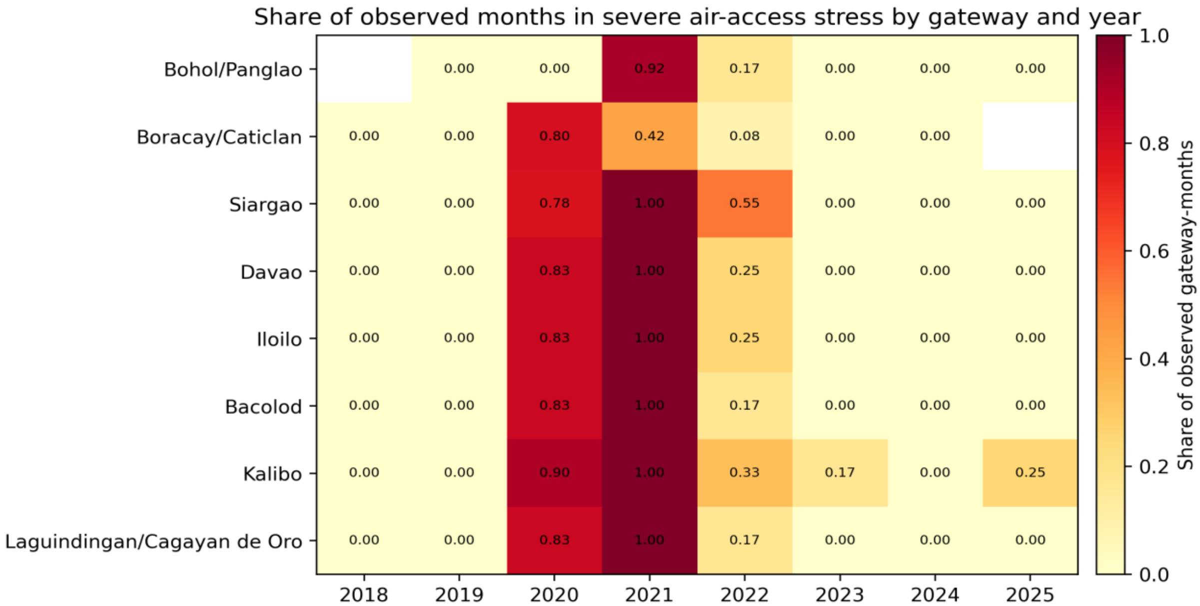

Figure 3. Share of observed months classified as severe air-access stress by gateway and year. Blank cells indicate that a gateway-year lacked sufficient historical baseline information or source coverage for the relevant calculation.

4.2. Out-of-Time Model and Baseline Performance

Table 3 answers RQ2. Random forest produced the strongest AUC (0.940) and highest AP (0.537) among the machine-learning models, with a gateway-cluster bootstrap AUC interval of [0.877, 0.990]. Gradient boosting achieved similar discrimination (AUC = 0.931), while logistic regression achieved AUC = 0.911. However, discrimination did not translate into superior threshold-level classification. The persistence baseline achieved precision = 0.607, recall = 0.850, and F1 = 0.708, substantially above the random-forest F1 of 0.400.

Table 3. Out-of-time one-month access-stress performance, 2022-2025.

Model	AUC (95% cluster bootstrap CI)	AP	Balanced accuracy	Precision	Recall	F1
Logistic regression	0.911 [0.833, 0.974]	0.418	0.745	0.255	0.600	0.358
Random forest	0.940 [0.877, 0.990]	0.537	0.794	0.280	0.700	0.400
Gradient boosting	0.931 [0.871, 0.987]	0.537	0.796	0.227	0.750	0.349
Persistence baseline	0.908 [0.868, 0.988]	0.525	0.908	0.607	0.850	0.708
Seasonal-naive baseline	0.789 [0.710, 0.878]	0.148	0.789	0.163	0.850	0.274

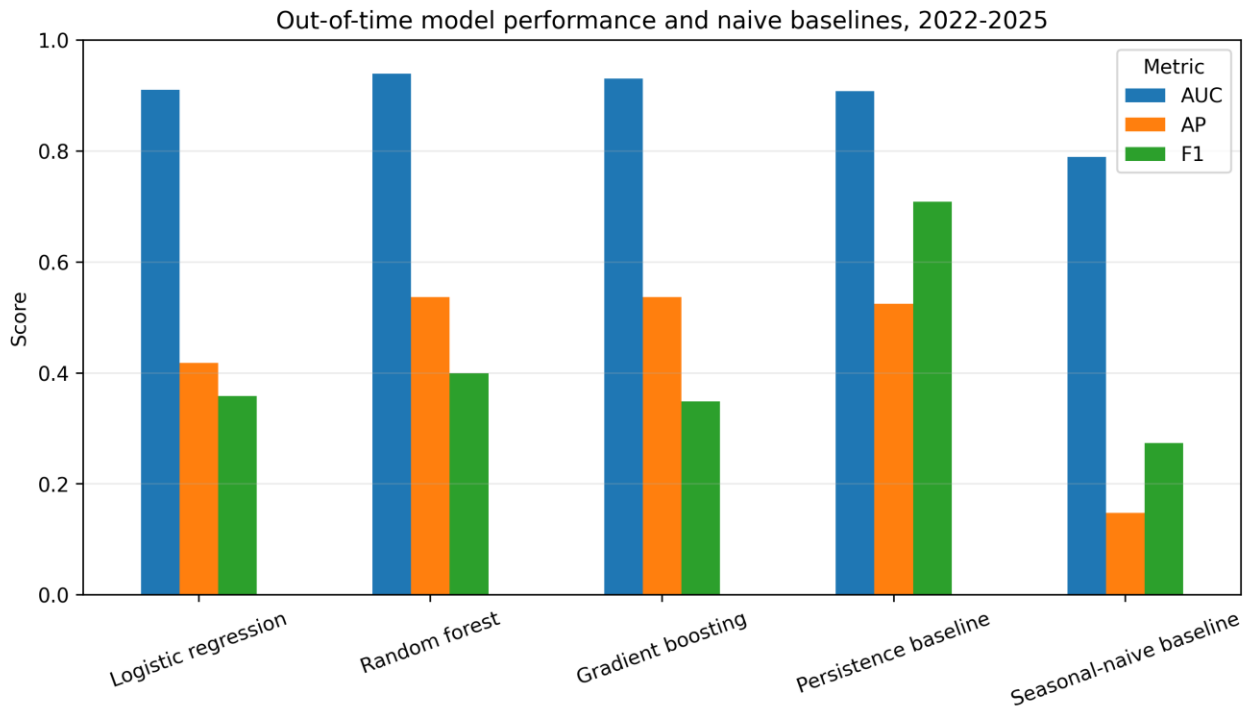

Figure 4. Out-of-time AUC, average precision, and F1 for the three machine-learning models and two naive baselines.

The baseline comparison changes the interpretation of the model. Random forest improved ranking discrimination over persistence (AUC 0.940 versus 0.908) and produced slightly higher average precision (0.537 versus 0.525), but persistence was substantially better at the fixed 0.50 operating point. The gateway-cluster bootstrap confidence intervals overlapped substantially (random forest [0.877, 0.990]; persistence [0.868, 0.988]), so the numerical AUC advantage should not be interpreted as a clearly separated performance gain. Seventeen of the 20 positive test targets (85%) were continuations of already active stress episodes, which explains why persistence was difficult to beat. The machine-learning contribution is therefore best described as incremental risk ranking and explanatory diagnostics, not operational superiority over a simple rule.

The strict 2024-2025 sensitivity period contained 150 usable observations but only three positive targets. Random forest achieved AUC = 0.950 but F1 = 0.167. The very small number of positive events makes this recent-period estimate unstable; it is reported only as a sensitivity check and is not used as the headline result.

4.3. Test-Event Composition and Cross-Gateway Transferability

Appendix A lists every positive target month in the 2022-2025 test period, including whether stress was already active at the warning month. This disclosure shows that most positives were multi-month continuations rather than new onsets. The three true onsets were Kalibo in January 2023, Siargao in July 2022, and Kalibo in May 2025. The random forest correctly warned the latter two but missed the Kalibo January 2023 onset at the 0.50 threshold. The transferability was mixed. AUC exceeded 0.95 for Bacolod, Bohol/Panglao, Davao, Iloilo, and Laguindingan/Cagayan de Oro when both classes were present. Kalibo and Siargao were more difficult, with AUC values of 0.756 and 0.771, respectively. Boracay/Caticlan had no positive test targets, so AUC was undefined. Threshold behavior was also uneven: the Siargao model recovered all five positives but generated many false alarms, whereas the Davao model failed to classify positive at the 0.50 threshold. These results caution against assuming that a pooled model automatically transfers across destinations.

Table 4. Leave-one-gateway-out random-forest validation.

Held-out gateway	Test n	Positive targets	AUC	Precision	Recall	F1
Bacolod	47	1	0.957	0.250	1.000	0.400
Bohol/Panglao	39	1	0.974	0.167	1.000	0.286
Boracay/Caticlan	30	0	n/a	0.000	0.000	0.000
Davao	47	2	0.956	0.000	0.000	0.000
Iloilo	47	2	1.000	0.500	1.000	0.667
Kalibo	47	8	0.756	0.227	0.625	0.333
Laguindingan/Cagayan de Oro	38	1	1.000	0.500	1.000	0.667
Siargao	46	5	0.771	0.135	1.000	0.238

4.4. SHAP Explanations

Figure 5 addresses RQ3. Mean absolute SHAP values show that log passenger movement was the leading global driver, followed by recent three-month volatility and year-on-year growth. Month-on-month growth, active carriers, and carrier HHI contributed less, while calendar seasonality terms were comparatively small. The ranking indicates that the model primarily learned scale, instability, and momentum patterns rather than a simple seasonal calendar rule.

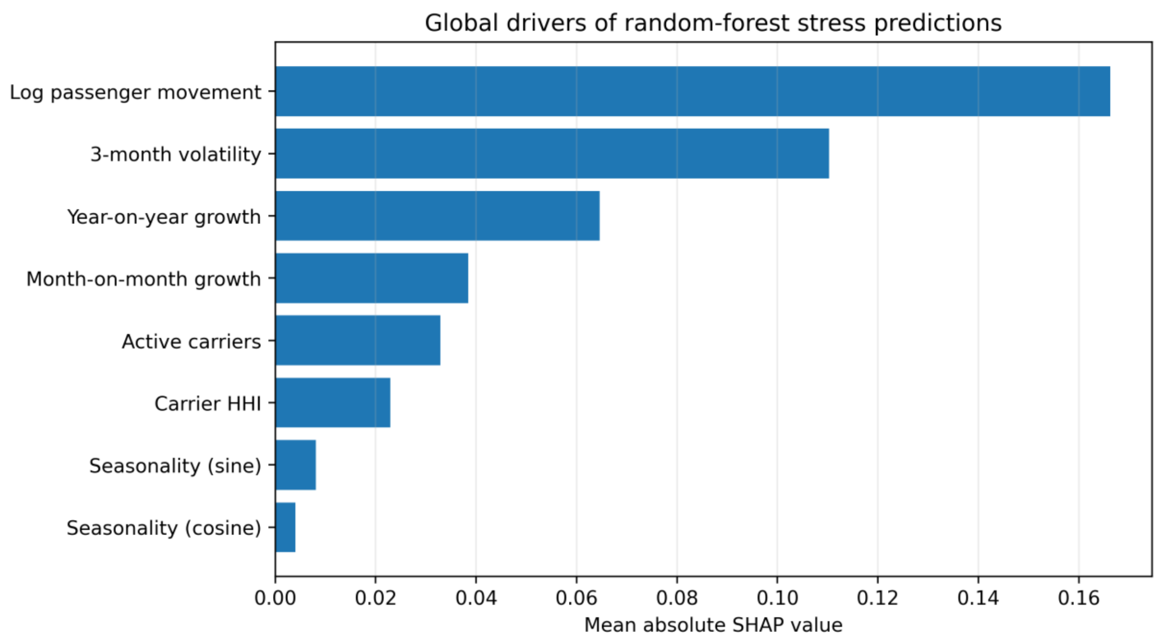

Figure 5. Global random-forest feature importance measured by mean absolute SHAP value on the 2022-2025 out-of-time test set.

The local explanations illustrate why case-level interpretation matters. For the Siargao June 2022 warning, low passenger scale and strong year-on-year change pushed the July 2022 stress probability upward to 0.774. For Kalibo April 2025, recent volatility and negative year-on-year growth were the strongest positive contributors, producing a risk of 0.787. The missed Kalibo December 2022 onset had a predicted probability of 0.454: low recent volatility reduced the score even though passenger scale and year-on-year growth pushed it upward. SHAP therefore helps diagnose both successful and failed warnings, but it does not establish causal mechanisms.

Table 5. Local SHAP explanations for two detected onsets and one missed onset.

Gateway	Warning month	Target month	Predicted risk	Class	Rank	Local driver	SHAP contribution
Siargao	2022-06	2022-07	0.774	1	1	Log passengers	+0.213
Siargao	2022-06	2022-07	0.774	1	2	YoY growth	+0.047
Siargao	2022-06	2022-07	0.774	1	3	Active carriers	+0.039
Siargao	2022-06	2022-07	0.774	1	4	Carrier HHI	-0.029
Kalibo	2025-04	2025-05	0.787	1	1	3-month volatility	+0.141
Kalibo	2025-04	2025-05	0.787	1	2	YoY growth	+0.097
Kalibo	2025-04	2025-05	0.787	1	3	Log passengers	+0.069
Kalibo	2025-04	2025-05	0.787	1	4	MoM growth	-0.031
Kalibo	2022-12	2023-01	0.454	0	1	3-month volatility	-0.138
Kalibo	2022-12	2023-01	0.454	0	2	Log passengers	+0.119
Kalibo	2022-12	2023-01	0.454	0	3	Active carriers	-0.031
Kalibo	2022-12	2023-01	0.454	0	4	YoY growth	+0.027

4.5. External Tourism-Economic Validation: Access Stress as an Enterprise-Environment Risk Signal

RQ4 asks whether the access-side signal is related to tourism-specific economic conditions relevant to the operating environment of tourism enterprises. Across the eight annual observations, the share of gateway-months in severe stress was strongly and negatively associated with TDGVA share of GDP (Spearman rho = -0.952, $p < 0.001$) and negatively associated with tourism employment share (rho = -0.708, p = 0.050). Annual selected-gateway passenger growth was strongly and positively associated with TDGVA growth (rho = 0.929, p = 0.003). Mean shock depth was also positively associated with TDGVA growth (rho = 0.886, p = 0.019). The reported sample sizes differ because stress-month share is available for 2018-2025 (n = 8), annual passenger growth is undefined for 2018 (n = 7), and mean shock depth is unavailable for 2018-2019 because the same-month baseline requires sufficient prior-year observations (n = 6). Because four focal correlations were examined without adjustment for multiple comparisons, the results are interpreted as exploratory rather than confirmatory.

Table 6. Selected external tourism-outcome validation associations.

Access indicator	Tourism outcome	n	Spearman ρ	p-value
Stress-month share	TDGVA share of GDP	8	-0.952	<0.001
Stress-month share	Tourism employment share	8	-0.708	0.050
Annual passenger growth	TDGVA growth	7	0.929	0.003
Mean shock depth	TDGVA growth	6	0.886	0.019

The validation results support the interpretation of air-access stress as a tourism-enterprise-environment risk signal, but the evidence is deliberately bounded. The eight annual observations are dominated by the 2020-2021 pandemic collapse and operate at national PTSA scale, whereas the access panel is built from eight gateways. The correlations therefore provide convergent rather than causal, gateway-specific, or firm-level validation. They justify linking the aviation signal to the economic environment in which tourism enterprises operate, but not treating passenger movement as a complete measure of customer demand, sales, profitability, cash flow, or entrepreneurial resilience.

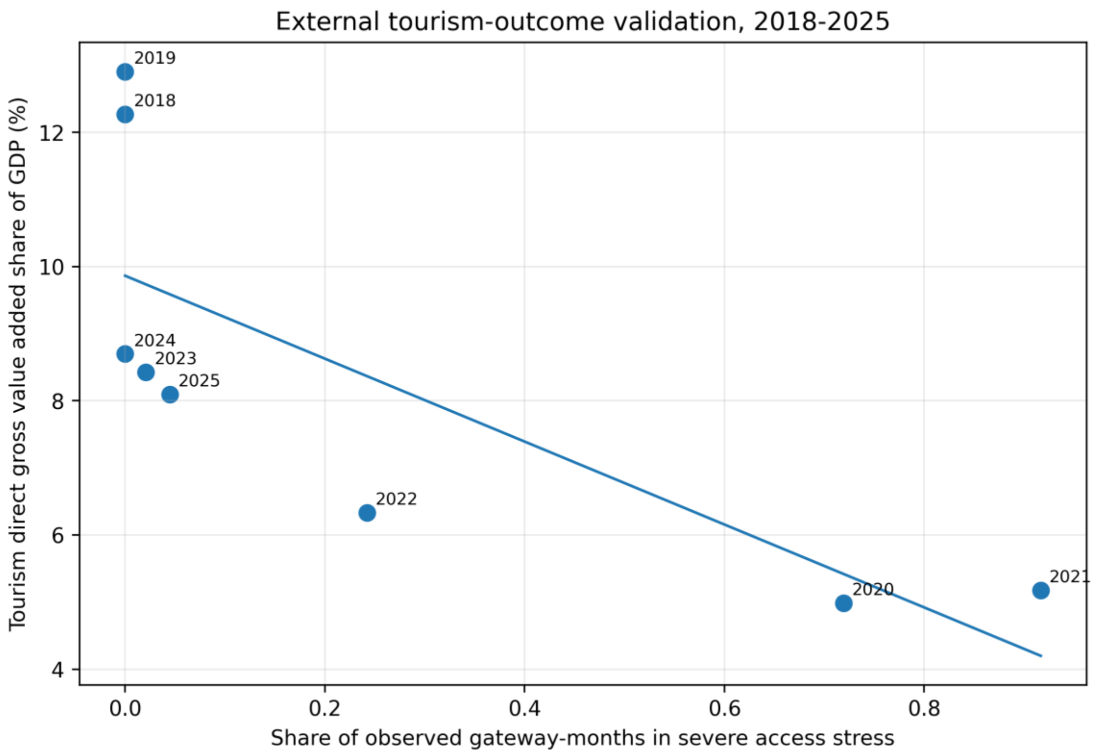

Figure 6. External validation of annual severe access-stress share against the tourism direct gross value added share of GDP, 2018-2025. The fitted line is descriptive; the association is not a causal estimate.

5. Discussion

5.1. Theoretical Contribution: Access Stress as an Upstream Tourism-Enterprise Risk Signal

The central theoretical contribution is a boundary clarification. Air access is neither full destination resilience nor direct enterprise demand; it is an upstream market and operating condition that can transmit disruption into tourism ventures and local entrepreneurial ecosystems. Papatheodorou [1] and Mazzola et al. [2] position accessibility as a key enabling condition for tourism, particularly where island systems depend on air links. Hall et al. [16] and Qu et al. [6] emphasize transitions, recovery trajectories, and actionable responses following disturbance. The shock-response-recovery framing is therefore more useful than a single composite gateway score because it observes when performance departs from a gateway-specific baseline, how long stress persists, and whether access recovers. Conceptually, this positions access monitoring within entrepreneurial risk management, market sensing, service continuity, and crisis and recovery planning rather than within transport performance alone.

The PTSA validation layer makes the management contribution empirically testable at the tourism-economy level. The access indicator is checked against tourism-specific economic outcomes rather than linked to enterprise performance only through hypothetical downstream effects. The strong annual associations do not support causal inference and are dominated by the pandemic period, but they are consistent with the proposition that broad access disruption and the environment for tourism enterprise activity are connected. The resulting construct is therefore an upstream tourism-enterprise risk signal: useful for early attention and operational verification, but not a substitute for direct measures of bookings, occupancy, sales, expenditure, employment, cash flow, profitability, business closure, or entrepreneurial adaptation.

5.2. What Machine Learning Adds to Entrepreneurial Decision-Making

The benchmark results provide a deliberately restrained interpretation of machine learning. Random forest and gradient boosting ranked risk well, and SHAP exposed the importance of scale, volatility, and momentum. However, persistence produced much higher F1 because most positive targets were continuations of existing stress. An entrepreneur or manager who only needs a binary “remain on alert” rule may therefore obtain more practical value from persistence than from the machine-learning model. This finding reinforces a central principle of responsible business analytics: complexity should be justified by incremental decision value.

The value of machine learning lies elsewhere. First, probability rankings can prioritize monitoring when multiple destinations or enterprise locations require attention. Second, SHAP can explain why a case is scored as high or low risk, making it easier to connect the warning to specific operating reviews. Third, the model can sometimes detect true onsets that persistence necessarily misses, as observed for Siargao in July 2022 and Kalibo in May 2025. These benefits are incremental rather than universal, and their entrepreneurial value depends on false-alarm costs, local data availability, working-capital constraints, lead times, and the decision threshold used by managers.

Taken together, the design is best understood as an interpretable-ML entrepreneurial risk-signal study with naive-baseline benchmarking: the machine-learning component serves a tourism-enterprise management problem rather than constituting the substantive outcome itself. This positioning is consistent with recent tourism applications in which interpretable machine learning is used to investigate substantive outcomes rather than treated as an end [9, 10]. It also aligns information-systems research with entrepreneurial problem solving by requiring the analytical output to be understandable, comparable with simple rules, and connected to specific managerial decisions.

5.3. Implications for Tourism Entrepreneurs, MSMEs, and Public Policy

For destination management organizations, local tourism offices, enterprise-support institutions, and industry associations, the model can function as a screening layer within a broader tourism-enterprise risk dashboard. A rising probability of access stress should trigger verification rather than automatic intervention. Decision makers can check airline schedule changes, cancellation patterns, hotel booking pace, visitor arrivals, weather advisories, local transport conditions, and event calendars before adjusting campaigns, market communication, enterprise assistance, events, or contingency plans.

For tourism entrepreneurs and hospitality firms, the warning signal is most useful when linked to operating lead times and business continuity. A high-risk score may justify closer review of flexible staffing rosters, perishable inventory, supplier orders, transport coordination, group bookings, cancellation exposure, pricing, promotional commitments, cash-flow needs, and customer communication. The persistence result also suggests that once a destination enters a verified stress episode, maintaining an alert state can be more reliable than repeatedly relying on a complex model. Machine learning is therefore most useful for prioritization and onset surveillance, while persistence remains a strong rule for ongoing episodes.

For public agencies, the findings support integrated tourism and enterprise intelligence rather than a new standalone airport index. CAAP access signals, PSA tourism outcomes, visitor statistics, accommodation indicators, local business conditions, and hazard data should be connected at compatible spatial and temporal scales through cross-sector data partnerships. Such integration would allow future systems to predict enterprise-relevant outcomes directly - including bookings, occupancy, sales, employment, service disruption, and business continuity - while retaining air-access stress as an upstream warning layer.

5.4. Implications for Entrepreneurship Education and Enterprise Development

For BS Entrepreneurship education, the study provides a practical example of how market research, business analytics, strategic management, operations, and risk management can be integrated around a real industry problem. Students can learn to transform public administrative data into market intelligence, compare advanced models with simple decision rules, interpret model explanations, and translate risk signals into choices involving staffing, inventory, suppliers, pricing, promotion, capacity, and continuity. The pedagogical emphasis is not on training students to become computer scientists, but on developing entrepreneurs who can ask sound business questions, evaluate analytical evidence, and use technology responsibly.

For entrepreneurship faculty, incubators, and extension programs, the framework offers a basis for enterprise-risk dashboards, scenario exercises, consultancy projects, and local industry partnerships. It demonstrates a legitimate interdisciplinary role for entrepreneurship scholars: defining the business problem, identifying decision-relevant indicators, setting practical benchmarks, interpreting results within enterprise constraints, and ensuring that technological tools remain subordinate to entrepreneurial judgment and stakeholder needs.

5.5. Interpretation Cautions

Four cautions follow from the entrepreneurship-centered interpretation. First, SHAP explains the behavior of the fitted model, not the causal effect of a feature on tourism, aviation, or firm performance. Second, strong annual access-PTSA correlations are partly shaped by the pandemic and should not be generalized as stable structural coefficients. Third, the model does not identify why a passenger decline occurred: airline capacity decisions, storms, public-health restrictions, economic conditions, destination closures, and reporting changes can produce similar patterns. Fourth, the study does not observe individual enterprises, entrepreneurs, sales, cash flow, or adaptation decisions. Operational use therefore requires contextual verification and local firm-level data.

6. Limitations and Future Research

The study has seven main limitations. First, the CAAP passenger records do not distinguish tourists from residents, workers, business travelers, or visiting friends and relatives. Second, the panel contains only eight selected gateways and is unbalanced because source reporting coverage varies by year and airport. Third, the same-month historical median is a transparent baseline rather than a structural counterfactual model; unusually low or high prior years can affect shock depth. Fourth, the out-of-time test period contains relatively few new stress onsets, and threshold-level performance is weaker than the persistence baseline. Fifth, PTSA validation uses only eight national annual observations, so the strong associations are descriptive and pandemic-dominated. Sixth, the IBTrACS layer is only a 2023-2025 pilot and is excluded from headline model estimates. Seventh, the study does not contain firm-level measures of bookings, revenue, cash flow, staffing responses, inventory losses, business closure, innovation, or entrepreneurial adaptation; consequently, it identifies an external risk signal rather than estimating enterprise outcomes.

Future research should prioritize tourism-enterprise outcome data at the same spatial and temporal scale as the access panel. Monthly bookings, visitor arrivals, hotel occupancy, average daily rate, sales, cancellations, payroll, supplier commitments, inventory spoilage, cash-flow pressure, business closures, and adaptation strategies would allow the model to test whether access stress precedes measurable enterprise losses or strategic responses. A complete historical IBTrACS layer, route-level capacity and frequency data, aircraft movements, fares, and local weather could then be added as explanatory inputs. Methodologically, future work should compare calibrated probability models, decision-cost thresholds, onset-specific models, survival analysis for recovery duration, and prospective evaluation on newly arriving data. Cross-destination and cross-enterprise validation should remain mandatory. Field testing with tourism entrepreneurs, business incubators, local governments, and entrepreneurship students could further assess whether the warnings are understandable, timely, and actionable.

7. Conclusions

Across 2018-2025, abnormal air-access stress co-varied with measured Philippine tourism performance: years with greater severe-stress exposure had lower tourism direct gross value added share of GDP, while passenger growth moved with TDGVA growth. These national annual associations are exploratory and pandemic-dominated, so they do not establish causality or firm-level effects. Their contribution is narrower but important: they support air-access stress as an upstream signal of the external market and operating environment faced by tourism enterprises, complementing direct measures of bookings, occupancy, sales, expenditure, employment, and entrepreneurial adaptation.

The operational finding is that no single predictive tool dominates. Random forest provided strong risk ranking and interpretable SHAP diagnostics, but persistence achieved the higher F1 score because most positive targets were continuations of existing stress. A hybrid entrepreneurial decision-support approach is therefore most defensible: using machine learning to prioritize locations and watch for onset, persistence rules to maintain alerts during verified episodes, and local tourism-enterprise indicators to confirm action. This combination aligns predictive analytics with market intelligence, resource planning, business continuity, and recovery rather than treating passenger movement or an AI score as a standalone measure of enterprise performance.

From an entrepreneurship perspective, the broader contribution is a replicable approach for converting public-sector data into responsible and interpretable market intelligence. The study shows how entrepreneurship researchers and educators can work with information-systems and engineering specialists while retaining leadership over problem definition, enterprise relevance, managerial interpretation, and translation into MSME resilience and industry-development actions.

Abbreviations

The following abbreviations are used in this manuscript:

AP	**Average Precision**
AUC	Area Under the Receiver Operating Characteristic Curve
CAAP	Civil Aviation Authority of the Philippines
CI	Confidence Interval
F1	F1 Score
HHI	Herfindahl–Hirschman Index
IBTrACS	International Best Track Archive for Climate Stewardship
LOGO	Leave-One-Gateway-Out
ML	Machine Learning
PSA	Philippine Statistics Authority
PTSA	Philippine Tourism Satellite Account
RF	Random Forest
SHAP	SHapley Additive exPlanations
TC	Tropical Cyclone

Acknowledgments: The authors acknowledge the Civil Aviation Authority of the Philippines for publishing annual movement workbooks, the Philippine Statistics Authority for the Philippine Tourism Satellite Account, NOAA NCEI for IBTrACS, and OurAirports for open airport-coordinate data. During the preparation of this manuscript, ChatGPT (GPT-5.5) was used for language editing, code organization support, and document-formatting assistance. The authors reviewed and verified the source extraction logic, data processing, analyses, results, and interpretations and take full responsibility for the content of the manuscript.

Funding: This research received no external funding.

Author Contributions: Conceptualization, K.R.M.D., J.M.M.D. and L.C.M.D.J.; methodology, K.R.M.D., J.M.M.D., L.C.M.D.J. and J.R.G.L.; software, L.C.M.D.J. and J.R.G.L.; validation, K.R.M.D., V.A.A.P., E.A.P., M.S.M. and J.R.G.L.; formal analysis, L.C.M.D.J. and J.R.G.L.; investigation, K.R.M.D., J.M.M.D., V.A.A.P., E.A.P. and M.S.M.; resources, K.R.M.D., J.M.M.D., V.A.A.P., E.A.P. and M.S.M.; data curation, J.M.M.D., L.C.M.D.J. and J.R.G.L.; writing-original draft preparation, K.R.M.D., J.M.M.D. and L.C.M.D.J.; writing-review and editing, K.R.M.D., V.A.A.P., E.A.P., M.S.M., L.C.M.D.J. and J.R.G.L.; visualization, L.C.M.D.J. and J.R.G.L.; supervision, K.R.M.D. and M.S.M.; project administration, K.R.M.D.; funding acquisition, K.R.M.D. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Data Availability Statement: The raw aviation workbooks are publicly available from the Civil Aviation Authority of the Philippines Aircraft, Passenger, and Cargo Movements page. The Philippine Tourism Satellite Account statistical tables are publicly available from the Philippine Statistics Authority. The pilot tropical-cyclone data are available from NOAA NCEI IBTrACS, and airport coordinate data are available from OurAirports. The corrected processed panel, event tables, benchmark results, uncertainty estimates, SHAP outputs, and analysis code are included in the supplementary materials prepared for submission. Source URLs and access dates are provided in the References.

Institutional Review Board Statement: Not applicable. The study used publicly available administrative, economic, geospatial, and climate datasets and did not involve human participants or animals.

Informed Consent Statement: Not applicable.

Appendix A

Appendix A.1 Positive Out-of-Time Stress Targets

Table A1. Positive next-month stress targets in the 2022-2025 out-of-time evaluation period.

Gateway	Warning month	Stress target month	Stress already active?	Persistence correct?
Bacolod	2022-01	2022-02	Yes	Yes
Bohol/Panglao	2022-01	2022-02	Yes	Yes
Davao	2022-01	2022-02	Yes	Yes
Davao	2022-02	2022-03	Yes	Yes
Iloilo	2022-01	2022-02	Yes	Yes
Iloilo	2022-02	2022-03	Yes	Yes
Kalibo	2022-01	2022-02	Yes	Yes
Kalibo	2022-02	2022-03	Yes	Yes
Kalibo	2022-03	2022-04	Yes	Yes
Kalibo	2022-12	2023-01	No	No
Kalibo	2023-01	2023-02	Yes	Yes
Kalibo	2025-04	2025-05	No	No
Kalibo	2025-05	2025-06	Yes	Yes
Kalibo	2025-06	2025-07	Yes	Yes
Laguindingan/Cagayan de Oro	2022-01	2022-02	Yes	Yes
Siargao	2022-02	2022-03	Yes	Yes
Siargao	2022-03	2022-04	Yes	Yes
Siargao	2022-06	2022-07	No	No
Siargao	2022-07	2022-08	Yes	Yes
Siargao	2022-08	2022-09	Yes	Yes

REFERENCES

1. Papatheodorou, A. (2021). A review of research into air transport and tourism: Launching the Annals of Tourism Research Curated Collection on Air Transport and Tourism. Annals of Tourism Research, 87, 103151. https://doi.org/10.1016/j.annals.2021.103151
2. Mazzola, F., Cirà, A., Ruggieri, G., & Butler, R. (2022). Air transport and tourism flows to islands: A panel analysis for southern European countries. International Journal of Tourism Research, 24(5), 639–652. https://doi.org/10.1002/jtr.2527
3. Tang, J., Wu, H., Ramos, V., & Sriboonchitta, S. (2023). The impact of international air routes on tourism. Annals of Tourism Research, 101, 103583. https://doi.org/10.1016/j.annals.2023.103583
4. Wang, T., Yang, Z., Chen, X., & Han, F. (2022). Bibliometric analysis and literature review of tourism destination resilience research. International Journal of Environmental Research and Public Health, 19(9), 5562. https://doi.org/10.3390/ijerph19095562

5. Liu, L.-W., Pahrudin, P., Tsai, C.-Y., & Hao, L. (2024). Disaster, risk and crises in tourism and hospitality field: A pathway toward tourism and hospitality management framework for resilience and recovery process. Natural Hazards Research, 4(4), 653–668. https://doi.org/10.1016/j.nhres.2024.06.001
6. Qu, M., Fu, Y., Chen, D., Zhang, Y., & Bao, J. (2026). From impact to action: Enhancing international tourism resilience through counterfactual explanations. Tourism Management, 114, 105356. https://doi.org/10.1016/j.tourman.2025.105356
7. Estiri, M., Heidary Dahooie, J., & Skare, M. (2022). COVID-19 crisis and resilience of tourism SME's: A focus on policy responses. Economic Research-Ekonomska Istraživanja, 35(1), 5556–5580. https://doi.org/10.1080/1331677X.2022.2032245
8. Smolčić Jurdana, D., & Agbaba, R. (2025). Management of crisis situations towards tourism destination sustainability: Key factors and measures. Sustainability, 17(23), 10871. https://doi.org/10.3390/su172310871
9. Ghosh, B. K. (2026). Infrastructure, governance, and price stability as binding constraints on inbound tourism to India. Tourism and Hospitality, 7(2), 57. https://doi.org/10.3390/tourhosp7020057
10. Oruç Erdoğan, E., Özdemir, O., Erdoğan, M., Durmuş Özdemir, E., & Özdemir, Ş. (2026). Environmental sustainability indicators and international tourism demand: Evidence from machine learning and SHAP analysis. Tourism and Hospitality, 7(6), 170. https://doi.org/10.3390/tourhosp7060170
11. Silveira, L., Santos, N., & Moreira, C. O. (2023). The liberalisation of air transport and the impacts on travel and tourism: The case of the Azores Archipelago. Sustainability, 15(5), 4488. https://doi.org/10.3390/su15054488
12. Holling, C. S. (1973). Resilience and stability of ecological systems. Annual Review of Ecology and Systematics, 4, 1–23. https://doi.org/10.1146/annurev.es.04.110173.000245
13. Folke, C., Carpenter, S. R., Walker, B., Scheffer, M., Chapin, T., & Rockström, J. (2010). Resilience thinking: Integrating resilience, adaptability and transformability. Ecology and Society, 15(4), 20. https://doi.org/10.5751/ES-03610-150420
14. Duro, J. A., Perez-Laborda, A., & Fernandez, M. (2022). Territorial tourism resilience in the COVID-19 summer. Annals of Tourism Research Empirical Insights, 3(1), 100039. https://doi.org/10.1016/j.annale.2022.100039
15. Bertella, G. (2022). Discussing tourism during a crisis: Resilient reactions and learning paths towards sustainable futures. Scandinavian Journal of Hospitality and Tourism, 22(2), 144–160. https://doi.org/10.1080/15022250.2022.2034527
16. Hall, C. M., Prayag, G., & Fang, S. (2024). Destination transitions and resilience following trigger events and transformative moments. Scandinavian Journal of Hospitality and Tourism, 24(4–5), 390–411. https://doi.org/10.1080/15022250.2024.2344605
17. Lee, S., & Pennington-Gray, L. (2025). Measuring resilience of the tourism sector: Reflective Resilience Index (REFLEX) approach. Annals of Tourism Research, 114, 103983. https://doi.org/10.1016/j.annals.2025.103983
18. Bruneau, M., Chang, S. E., Eguchi, R. T., Lee, G. C., O'Rourke, T. D., Reinhorn, A. M., Shinozuka, M., Tierney, K., Wallace, W. A., & von Winterfeldt, D. (2003). A framework to quantitatively assess and enhance the seismic resilience of communities. Earthquake Spectra, 19(4), 733–752. https://doi.org/10.1193/1.1623497
19. Rose, A. (2004). Defining and measuring economic resilience to disasters. Disaster Prevention and Management: An International Journal, 13(4), 307–314. https://doi.org/10.1108/09653560410556528
20. Viverit, L., Heo, C. Y., Pereira, L. N., & Tiana, G. (2023). Application of machine learning to cluster hotel booking curves for hotel demand forecasting. International Journal of Hospitality Management, 111, 103455. https://doi.org/10.1016/j.ijhm.2023.103455
21. Contessi, D., Viverit, L., Pereira, L. N., & Heo, C. Y. (2024). Decoding the future: Proposing an interpretable machine learning model for hotel occupancy forecasting using principal component analysis. International Journal of Hospitality Management, 121, 103802. https://doi.org/10.1016/j.ijhm.2024.103802
22. Huang, J., & Zhang, C. (2024). Daily tourism demand forecasting with the iTransformer model. Sustainability, 16(23), 10678. https://doi.org/10.3390/su162310678
23. Hu, M., Dong, N., & Hu, F. (2024). Tourism demand forecasting using short video information. Annals of Tourism Research, 109, 103838. https://doi.org/10.1016/j.annals.2024.103838
24. Lee, G.-C. (2025). A data-driven approach to tourism demand forecasting: Integrating web search data into a SARIMAX model. Data, 10(5), 73. https://doi.org/10.3390/data10050073
25. Vasenska, I. (2025). Comparative analysis of machine learning and deep learning models for tourism demand forecasting with economic indicators. FinTech, 4(3), 46. https://doi.org/10.3390/fintech4030046
26. Hou, H., & Wang, H. (2026). Research on interpretable tourism demand forecasting based on VSN-xLSTM model. Systems, 14(2), 146. https://doi.org/10.3390/systems14020146
27. Civil Aviation Authority of the Philippines. (2026). Aircraft, passenger, and cargo movements. Available at: https://www.caap.gov.ph/aircraft-passenger-and-cargo-movements/ (Accessed on: 6 July 2026).
28. Philippine Statistics Authority. (2026). Philippine Tourism Satellite Account statistical tables, 2000–2025. Available at: https://psa.gov.ph/statistics/tourism (Accessed on: 6 July 2026).
29. Gahtan, J., Knapp, K. R., Schreck, C. J., III, Diamond, H. J., Kossin, J. P., & Kruk, M. C. (2024). International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4.01 [Western Pacific subset, 2023–2025]. NOAA National Centers for Environmental Information. https://doi.org/10.25921/82ty-9e16 (Accessed on: 6 July 2026).

30. OurAirports. (2026). Open data downloads. Available at: https://ourairports.com/data/ (Accessed on: 6 July 2026).

Biographical Sketches

Karen Razelle M. Duyan is affiliated with Kalinga State University and O.P. Jindal Global University. Her research interests include entrepreneurship, tourism-enterprise resilience, innovation management, market research, and data-informed decision support.

Juan Moshe M. Duyan is affiliated with Kalinga State University and O.P. Jindal Global University. His research interests include tourism, public administration, indigenous governance, and evidence-informed policy and management.

Valerie A. Abesamis Paloy is affiliated with Kalinga State University and contributed to the investigation and validation of the study.

Eric A. Paloy is affiliated with Kalinga State University and contributed to the investigation and validation of the study.

Melanie S. Manuel is affiliated with Kalinga State University and contributed to investigation, validation, and research supervision.

Mark Joby T. Aguilar is affiliated with Kalinga State University and contributed to the investigation and validation of the study.

Joe Robert G. Lucena is affiliated with O.P. Jindal Global University and the School of Engineering, Asia Pacific College. His contributions include methodology, software, formal analysis, validation, and visualization.

Luigi Carlo M. De Jesus is affiliated with O.P. Jindal Global University and the School of Engineering, Asia Pacific College. His contributions include conceptualization, methodology, software, formal analysis, writing, and visualization.