tourism and hospitality

MDPI

Article

Open-Data Nowcasting of Ecuador's International Tourist Arrivals: Regularized Dynamic Regression with Wikipedia Attention and Copernicus Land Reanalysis Climate Signals

Julio Guerra [1,*], Sheyla Fernández [2], Danny Benavides [2], Víctor Caranquí [1] and Mónica Meneses [1]

[1] Facultad de Ingeniería en Ciencias Aplicadas, Universidad Técnica del Norte, Ibarra 100101, Ecuador; vmcaranqui@utn.edu.ec (V.C.); mamenesesb@utn.edu.ec (M.M.)
[2] Facultad de Ciencias de la Vida, Universidad Estatal Amazónica, Zamora 190150, Ecuador
* Correspondence: jeguerra@utn.edu.ec; Tel.: +593-98-8257720

Abstract

Timely monitoring of tourism demand is essential for destination management, yet official monthly arrival statistics are often released with delays and can be difficult to use for near-real-time decision-making, particularly under structural shocks such as coronavirus disease 2019 (COVID-19). This study develops a fully reproducible, open-data nowcasting pipeline for Ecuador's international tourist arrivals using a Python workflow. The framework integrates (i) the official monthly arrivals series published by Ecuador's Ministry of Tourism (MINTUR), (ii) open online attention proxies from Wikipedia pageviews retrieved via the Wikimedia REST application programming interface (API), and (iii) open climate covariates derived from the ERA5-Land land reanalysis. Multiple forecasting models are evaluated under a rolling-origin, one-step-ahead backtest, with a mandatory seasonal naïve benchmark and a regime-aware assessment that separates a stress-test window (2019–2021) from an operational post-COVID window (2022–2025). Forecast accuracy is summarized using root mean squared error (RMSE), mean absolute error (MAE), and symmetric mean absolute percentage error (sMAPE), and statistical significance of performance differences is assessed using the Diebold–Mariano (DM) test. Results show that a ridge-regularized autoregressive model (ridge_ar) achieves the best overall accuracy, reducing RMSE by approximately 79% relative to the seasonal naïve baseline over the full evaluation window. Windowed results confirm robust performance during the shock period and sustained improvements in the post-2022 operational regime, while the incremental benefit of broader exogenous signals is heterogeneous across windows, underscoring the importance of regularization and regime-aware reporting. The proposed approach provides a transparent, low-cost blueprint for reproducible tourism monitoring that is transferable to other destinations using open data and standard computational tools.

Keywords: tourism demand nowcasting; rolling-origin backtesting; regularized regression; Wikipedia pageviews

Check for updates

Academic Editor: Lewis Ting On Cheung

Received: 7 January 2026
Revised: 9 February 2026
Accepted: 12 February 2026
Published: 20 April 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Background, Problem Statement, and Study Aim

Tourism systems are exposed to rapid demand fluctuations driven by macroeconomic conditions, travel frictions, safety perceptions, and weather-sensitive seasonality. For destinations, monthly inbound-arrival statistics remain a central operational indicator for

capacity planning, marketing allocation, and short-run policy response. However, official tourism indicators are commonly published with non-negligible lags and may be revised, which limits their usefulness for near-real-time monitoring. The COVID-19 shock illustrated this limitation particularly clearly, because demand collapsed and recovered in a non-linear and highly time-varying manner that challenged standard forecasting assumptions and increased the value of timely monitoring signals (Collins-Kreiner & Ram, 2021).

Against this backdrop, tourism-demand forecasting research has increasingly incorporated internet-based and other high-frequency proxies as leading indicators, motivated by the idea that information-seeking behavior can precede realized travel. Evidence has accumulated that digital traces can add incremental predictive content for tourism series, particularly in short-horizon settings where "nowcasting" (i.e., producing an estimate for the current/next release using contemporaneous signals) is operationally relevant (Ting et al., 2025). Among these signals, Wikipedia pageviews have been proposed as an open, interpretable proxy of destination-related attention and information demand, with potential value for estimating tourism activity when traditional statistics are delayed. In addition, climate variability provides a physically grounded channel through which environmental conditions can affect travel comfort, accessibility, and perceived attractiveness, motivating the inclusion of reanalysis-based climate covariates such as ERA5-Land products (Hersbach et al., 2020).

Despite the growing use of digital traces in tourism forecasting, two methodological gaps remain in the tourism literature. First, many contributions rely on platform-dependent or partially opaque indicators and emphasize incremental predictive gains without a sufficiently explicit discussion of why such proxies should remain stable and behaviorally meaningful across time and shocks. Second, empirical implementations frequently remain close to classical regression-based designs (e.g., unpenalized dynamic regression and autoregressive integrated moving average with exogenous regressors (ARIMAX)/seasonal autoregressive integrated moving average with exogenous regressors (SARIMAX) specifications with exogenous regressors), which can become fragile when digital-trace predictors are numerous, strongly correlated, and subject to behavioral drift. Recent tourism studies using web or platform signals illustrate both the promise and the instability of such proxies across contexts and horizons, highlighting the need for designs that are reproducible, transparent, and explicitly evaluated under realistic rolling-origin protocols and regime shifts (Lee, 2025; Nie et al., 2024; Lei et al., 2025). This study addresses these gaps by (i) restricting inputs to fully open and reproducible signals, (ii) treating regularization as a principled extension of classical dynamic regression to stabilize estimation under collinearity, and (iii) adopting regime-aware, windowed evaluation (stress vs. post-COVID) to make clear when proxy-based nowcasting is operationally informative and when classical autoregressive dynamics dominate.

This study addresses the practical problem of producing reproducible, one-step-ahead monthly nowcasts of international tourist arrivals for Ecuador using fully open inputs. It develops a transparent feature-engineering pipeline that combines (i) the official monthly arrivals series, (ii) destination-related Wikipedia pageviews in Spanish and English, and (iii) ERA5-Land monthly climate variables. The modeling strategy emphasizes short-horizon operational performance under structural change conditions, explicitly acknowledging that multi-source exogenous indicators can be correlated and unstable around shocks, thereby motivating robust estimation and careful evaluation design.

Objective. The objective is to design and evaluate an open-data nowcasting framework that improves one-step-ahead monthly accuracy relative to a strong seasonal baseline, and that remains informative under a shock period and in a post-shock operational regime.

Research questions.

- RQ1. To what extent can open digital-trace and climate signals support accurate one-step-ahead monthly nowcasting of Ecuador's international tourist arrivals relative to a seasonal baseline?
- RQ2. How does model performance differ between a shock ("stress-test") window and a post-COVID operational window, and which modeling choices remain robust across both contexts?
- RQ3. Are the observed improvements over the baseline consistent enough to support an operational monitoring interpretation (i.e., not only visually plausible, but evaluable under rolling backtesting and windowed assessment)?

In short-horizon forecasting/nowcasting, exogenous signals can improve one-step-ahead accuracy only when they contain incremental information about the current/next outcome that is not already captured by the lag structure of the target series (i.e., they add predictive content beyond autoregressive persistence and seasonality). This is consistent with the general forecasting logic of dynamic regression with external predictors and with the notion that a predictor must add information not already embedded in past values of the series to be useful for prediction (Hyndman & Khandakar, 2008). In the tourism context, digital "attention" indicators are theoretically most relevant when they reflect pre-trip information-search and consideration processes that can precede realized travel, especially at short lags, whereas purely autoregressive models rely on stable persistence patterns that may weaken under shocks (Fodness & Murray, 1997; Xiang & Gretzel, 2010). Moreover, query/attention data have been shown in broader nowcasting settings to help "predict the present" precisely because they can respond contemporaneously to changing conditions, while official statistics are delayed (Choi & Varian, 2012).

Based on these arguments, we state three working expectations that guide interpretation (without treating them as strict hypotheses):

- (E1) Shock window advantage: During the COVID-related stress-test regime, attention-based exogenous signals (Wikipedia pageviews) are expected to yield larger incremental gains over purely autoregressive baselines because they can react quickly to abrupt changes in perceived risk, restrictions, and destination salience, while autoregressive persistence patterns are disrupted.
- (E2) Stabilized-regime attenuation: In the post-2022 operational regime, incremental gains from broad exogenous sets are expected to be smaller and potentially absent when the attention $\rightarrow$ travel mapping weakens (behavioral drift, media-driven attention, or representativeness shifts), in which case parsimonious autoregressive/state-space structures may dominate.
- (E3) Climate covariate nuance: Climate reanalysis variables provide a physically grounded channel (comfort/accessibility) but, at monthly resolution, their incremental predictive contribution is expected to be context-dependent and potentially modest relative to seasonality already captured by autoregressive components; therefore, their value is evaluated empirically rather than assumed.

The paper is organized to (i) describe data sources and reproducible feature construction, (ii) present the nowcasting models and evaluation protocol, and (iii) report full-sample and windowed performance evidence, including baseline comparisons, to support practical tourism monitoring use.

Conceptual Framing: Attention Signals, Open Data, and Regime Instability

From a behavioral tourism perspective, travel demand is preceded by an information-search phase in which prospective visitors form awareness, reduce uncertainty, compare

alternatives, and plan itineraries. The intensity and channels of this external search depend on perceived risk, involvement, and prior experience, and the resulting information-seeking behavior leaves observable digital traces (Fodness & Murray, 1997; Gursoy & McCleary, 2004; Xiang & Gretzel, 2010). In this study, we use the term attention to denote this revealed demand for destination-related information—a measurable manifestation of pre-trip cognition and planning activity rather than a direct measure of purchases or bookings. Within this interpretation, Wikipedia pageviews capture aggregate interest in destinations, attractions, and tourism-related topics, and can therefore provide timely signals that may lead realized arrivals with short lags (Alis et al., 2015; Khadivi & Ramakrishnan, 2016).

We also treat open data as a theoretical and operational design constraint. In the open-data literature, "open" implies not only accessibility but also legal and technical conditions enabling reuse and redistribution, which supports transparency, replication, and public value creation (Zuiderwijk & Janssen, 2014). Consistent with this principle, we restrict predictors to inputs that are openly accessible and reproducible (Wikipedia pageviews and ERA5-Land reanalysis), so that the full nowcasting workflow can be replicated and audited without proprietary platforms.

Finally, we use regime to describe periods in which the mapping between predictors and arrivals is approximately stable, while recognizing that shocks can induce structural breaks and parameter shifts that invalidate single-regime conclusions. In time-series terms, such breaks can be represented as regime changes or multiple structural changes (Hamilton, 1989; Bai & Perron, 2003). We therefore treat COVID-19 as a regime-breaking event and explicitly evaluate performance in separate windows (stress vs. post-shock operational), aligning the empirical protocol with the underlying instability that motivates near-real-time monitoring. The COVID-19 shock illustrated this limitation particularly clearly, because demand collapsed and recovered in a non-linear and highly time-varying manner that challenged standard forecasting assumptions and increased the value of timely monitoring signals. From an ex-ante perspective, fast-propagating travel-mediated shocks have also motivated anticipatory forecasting and risk-assessment research based on epidemic time-series signals and mobility/air-travel patterns, reinforcing why timely indicators matter when reporting and publication cycles lag behind real-world dynamics (Christidis & Christodoulou, 2020).

1.2. Literature Review

1.2.1. Tourism-Demand Forecasting and the Role of Short-Horizon Nowcasts

Tourism-demand forecasting has a long tradition in the tourism analytics literature, largely because inbound and domestic flows exhibit pronounced seasonality, trend shifts, and sensitivity to shocks; these properties create operational challenges for capacity planning, marketing timing, and public-sector monitoring. Reviews in Tourism Management emphasize that tourism demand is structurally different from many "standard" economic series because it mixes strong calendar effects with abrupt disruptions (e.g., crises) and destination-specific dynamics, which motivates both robust baselines and careful evaluation designs. A widely cited synthesis highlights that forecast performance depends strongly on horizon, model class, and the stability of the underlying demand-generating process (Song & Li, 2008).

Within this broader field, nowcasting (producing an estimate of the most recent/near-current period with minimal delay) is commonly treated as an operational short-horizon forecasting task, often implemented as one-step-ahead prediction for monthly series. In tourism contexts, one-step-ahead nowcasts are especially relevant because official arrival statistics can be released with lag, while decisions (pricing, staffing, promotion, congestion management) are typically monthly or intra-monthly. Methodologically, the literature increasingly argues that the key difficulty is not only "average" seasonality but also turning

points and regime changes—i.e., settings where methods that track gradual dynamics may fail when the process shifts abruptly (e.g., demand collapses or rebounds). Empirical work on forecasting turning points in tourism growth underscores the need for models and protocols that explicitly acknowledge non-stationarity and instability rather than assuming a single stable regime (Wan & Song, 2018; P. Wu et al., 2024).

1.2.2. Digital Trace Data as Leading Indicators for Tourism Demand

A central development in contemporary tourism forecasting is the use of digital traces (web search activity, online attention metrics, platform interactions) as leading indicators of demand. The conceptual motivation is straightforward: when potential travelers plan, compare destinations, or seek travel-related information, their activity leaves observable digital footprints that can precede realized arrivals. This logic aligns with "predicting the present" work in economics, showing that search intensity can contain timely signals about near-term outcomes (Choi & Varian, 2012).

Tourism-specific evidence has accumulated across destinations and data sources. A recent synthesis focused on internet-based tourism-demand forecasting documents consistent gains from incorporating web signals—especially at short horizons—while also emphasizing important caveats (platform changes, representativeness, language effects, and the risk of overfitting when many correlated signals are used). This line of research has expanded from early search-index approaches toward hybrid models and machine-learning pipelines, with a recurrent theme that the incremental value of web signals is often greatest in settings where official indicators are lagged or noisy and when forecasting is evaluated in realistic, rolling-origin setups.

Beyond search engines, Wikipedia pageviews can be situated more explicitly within tourism behavior theory as an information-search and attention signal. Destination choice and trip timing typically involve a pre-trip stage in which travelers seek factual knowledge, compare places, and reduce uncertainty; this external search is a well-established component of the tourist decision process (Fodness & Murray, 1997; Gursoy & McCleary, 2004). Because Wikipedia is a widely used, general-purpose knowledge base, pageview volumes for destination-relevant articles can be interpreted as an observable trace of aggregate destination-related attention—reflecting information acquisition consistent with planning, curiosity, and consideration rather than a direct purchase intention (Xiang & Gretzel, 2010). Empirically, prior work has shown that Wikipedia usage patterns can support tourism-demand analysis and forecasting, reinforcing the plausibility of pageviews as timely leading indicators when combined with appropriate controls and robust evaluation designs (Alis et al., 2015; Khadivi & Ramakrishnan, 2016). At the same time, the proxy is imperfect: pageviews may also reflect media coverage, educational interest, or other non-travel motivations. This motivates (i) regularization to stabilize inference under correlated multi-page/multi-language signals and (ii) windowed evaluation to test whether the attention–arrivals relationship remains informative across regimes and structural shocks.

From a critical tourism and "datafication" perspective, it is important to recognize that digital traces are not neutral sensors of demand but platform-mediated byproducts of how travel-related interest is expressed, captured, and surfaced online. In smart/digital tourism ecosystems, attention metrics emerge from socio-technical infrastructures (interfaces, content governance, referral pathways, and uneven digital access) that shape what becomes measurable and for whom, which can influence both explanatory interpretation and predictive stability (Gretzel et al., 2015; Kitchin, 2014). Consequently, the use of web-based proxies should be framed not only as an engineering choice but also as a measurement choice that embeds assumptions about how online information-seeking relates to offline travel behavior.

This implies several well-known failure modes that help explain why digital proxies may perform well in some contexts yet degrade in others. First, representativeness is structurally limited: platform users are not a random sample of travelers, and usage intensity can vary with socioeconomic status, connectivity, age, and language—factors that shift the "signal-to-demand" mapping across time and populations (van Deursen & van Dijk, 2014). Second, algorithmic and platform dynamics can introduce proxy instability: changes in search/referral mechanisms, user interfaces, content visibility, or platform practices can alter observed trace volumes without any underlying change in travel intentions, a central caution in the "big data" critique literature (Boyd & Crawford, 2012; Lazer et al., 2014). Third, behavioral drift is expected under structural shocks (e.g., COVID-19), when attention may reflect risk perception, news cycles, or constraint-driven curiosity rather than trip planning, leading to time-varying proxy validity. These considerations motivate the study's regime-aware evaluation (stress-test vs. post-2022 operational window) and the use of rolling-origin backtesting to assess whether any proxy-based gains persist under changing regimes rather than assuming a stable relationship.

1.2.3. Weather and Climate Reanalysis Data as Exogenous Predictors

Another stream relevant to tourism nowcasting is the inclusion of exogenous environmental variables (temperature, humidity, precipitation proxies) that can capture seasonal and anomalous conditions affecting travel comfort, mobility, and destination attractiveness. For empirical implementation, modern practice often relies on consistent reanalysis products that provide spatiotemporally complete coverage, enabling reproducible feature engineering at monthly resolution. The ERA5-Land reanalysis has become a common choice for land-surface applications because it offers a state-of-the-art, globally consistent dataset suitable for constructing temperature-/humidity-type covariates (Muñoz-Sabater et al., 2021).

In forecasting designs, environmental covariates are typically treated as complementary rather than standalone drivers: they can help explain seasonal structure, reduce omitted-variable bias when modeling demand dynamics, and improve robustness when combined with behavioral signals (web attention) and autoregressive structure.

1.2.4. Regularized Dynamic Regression Under Collinear High-Dimensional Signals

A practical challenge repeatedly noted in "open signals" tourism forecasting is that exogenous features are frequently high-dimensional and collinear (e.g., multiple related destination pageviews; climate features that move together seasonally). In such settings, unregularized regression can become unstable: coefficients can inflate, variance increases, and out-of-sample performance deteriorates—especially under rolling estimation, where sample sizes per training window are limited.

This is precisely the setting where regularization is theoretically motivated. Ridge regression was originally introduced to stabilize estimation under non-orthogonality/multicollinearity by shrinking coefficients (Hoerl & Kennard, 1970).

Elastic Net generalizes shrinkage by combining L1 and L2 penalties, allowing a compromise between stability and sparse selection when predictors are correlated (Zou & Hastie, 2005).

Tourism-forecasting work that blends internet signals with classical time-series components increasingly adopts regularized dynamic regression (autoregressive terms plus exogenous variables with shrinkage) as an engineering solution to the "many correlated predictors + short-horizon evaluation" problem, and reports that regularization can deliver more reliable gains than either purely autoregressive models or high-dimensional exogenous models without shrinkage (Bokelmann & Lessmann, 2019; Dergiades et al., 2018; Llewellyn et al., 2023).

Importantly, regularization primarily addresses estimation variance induced by multicollinearity and limited rolling-window samples, but it cannot correct measurement bias, representativeness limitations, or time-varying proxy validity in digital traces; these risks must be handled through careful evaluation design (rolling-origin protocols, windowed reporting under regime change) and transparent proxy interpretation (Boyd & Crawford, 2012; Lazer et al., 2014).

1.2.5. Evaluation Protocols and Statistical Comparison for Forecast Accuracy

Because tourism series are non-stationary and subject to regime shifts, the literature emphasizes evaluation designs that mimic real-time use: rolling-origin (walk-forward) backtesting is widely treated as the appropriate protocol for nowcasting and short-horizon forecasting. Performance is commonly summarized with scale-dependent errors (RMSE, MAE) and scale-free measures (e.g., sMAPE), but an important methodological point is that differences in average error are not sufficient when forecast errors are serially dependent.

Accordingly, forecast comparison often uses formal tests such as the Diebold–Mariano (DM) test, which evaluates whether the expected loss differential between two forecasting methods equals zero (Diebold & Mariano, 1995).

In tourism applications, DM-style comparisons are particularly useful when the goal is to claim incremental value over a strong benchmark, and when model performance is assessed across distinct regimes (e.g., "stress" vs. "post-shock" periods) rather than only in aggregate.

2. Materials and Methods

2.1. Study Design and Objective

This study develops a reproducible one-step-ahead monthly nowcasting pipeline for Ecuador's international tourist arrivals. Let y_t denote the number of international arrivals observed in calendar month t, for $t = 1, \ldots, T$, where T is the total number of months in the aligned sample. The operational target is the next-month nowcast $\hat{y}_{t+1|t}$, produced using information available up to the end of month t (i.e., an expanding-window, real-time-like evaluation).

The methodological contribution is a regularized dynamic regression framework (autoregressive structure + exogenous open signals) designed to remain stable under correlated predictors and post-shock regime changes, which are common in tourism time-series.

Notation and Indexing Conventions

To avoid ambiguity, we use the following indexing throughout the paper. The monthly time index is $t = 1, \ldots, T$. Rolling-origin evaluation uses τ to denote the forecast origin, with $\tau = \tau_0, \ldots, T-1$, where τ_0 is the first origin after an initial training period. Autoregressive lags are indexed by $i = 1, \ldots, p$. Exogenous covariates are indexed by $k = 1, \ldots, K$ and consist of K_P Wikipedia pageview series and K_C climate series ($K = K_P + K_C$); their distributed lags are indexed by $j = 0, \ldots, q$. Seasonality controls use month-of-year dummy variables $D_{m,t}$ with $m = 2, \ldots, 12$ (one category omitted for identification). We write the target series as $\mathrm{y} = (y_1, \ldots, y_T)^\top$ and the aligned exogenous covariate vector at month t as $\mathrm{x}_t = (x_{1,t}, \ldots, x_{K,t})^\top$.

2.2. Data Sources and Variables

2.2.1. Target Series: Monthly International Arrivals

The dependent variable y_t is the monthly international arrivals series published by Ecuador's tourism authority (MINTUR). The series is aggregated at the monthly frequency and used as the operational monitoring variable for the nowcasting task.

2.2.2. Open Exogenous Signals: Online Attention

To capture short-run demand signals not fully contained in autoregressive dynamics, the pipeline incorporates Wikipedia pageview counts for tourism-relevant pages (e.g., "Tourism in Ecuador", "Quito", "Galápagos Islands", in Spanish and English). These series serve as timely proxies of information demand/travel interest and are commonly used in nowcasting contexts when high-frequency digital traces are available. Conceptually, these pageviews are treated as aggregate pre-trip information-search attention signals (rather than bookings), consistent with established tourism information-search models (Fodness & Murray, 1997; Gursoy & McCleary, 2004). The complete set of Wikipedia page titles and languages used (index set S; with $| S | = K_P$) is provided in Supplementary Files to ensure full reproducibility.

2.2.3. Open Climate Covariates: ERA5-Land Reanalysis

Monthly climate covariates were derived from ERA5-Land, a global land reanalysis produced by ECMWF/Copernicus that provides physically consistent gridded variables suitable for downstream socioeconomic modeling (Muñoz-Sabater et al., 2021; Hersbach et al., 2020). ERA5-Land monthly means were retrieved via the Copernicus Climate Data Store (CDS), part of the Copernicus Climate Change Service (C3S), and downloaded in Network Common Data Form (NetCDF) format

The pipeline uses monthly aggregates for Ecuador (spatially aggregated over a fixed national bounding box/mask, depending on the implementation): near-surface air temperature (T_{2m}), dewpoint temperature ($T_{d,2m}$), and total precipitation (TP), with unit conversions: $T(°C) = T(K) - 273.15, TP(mm) = 1000 \cdot TP(m)$.

2.3. Temporal Alignment, Cleaning, and Monthly Indexing

All sources (arrivals, pageviews, climate) were aligned to a common monthly index. Each observation was mapped to the month start timestamp to ensure consistent merges and lag construction. Any model training and evaluation step only used information available up to the forecast origin (expanding-window), preventing look-ahead leakage.

2.4. Feature Engineering (Dynamic Regression Design Matrix)

2.4.1. Autoregressive Terms and Seasonality Controls

Tourism arrivals exhibit strong persistence and annual seasonality. Accordingly, the baseline dynamic structure includes p autoregressive lags y_{t-i} for $i = 1, \ldots, p$, and month-of-year seasonality controls via dummy variables $D_{m,t}$ for $m = 2, \ldots, 12$ (with one month omitted as the reference category).

2.4.2. Exogenous Lags (Online Attention + Climate)

Exogenous signals are incorporated with short distributed lags to represent delayed effects and reduce sensitivity to timing noise. For each exogenous covariate $x_{k,t}$ (Wikipedia pageviews or climate), we include lagged values $x_{k,t-j}$ for $j = 0, \ldots, q$. Because these exogenous features are often correlated within and across families (e.g., related pages/languages and climate variables), the modeling strategy emphasizes regularization for stability and generalization.

2.5. Forecasting Models

All equations reported in Section 2 are included for completeness and reproducibility and correspond to standard definitions in the forecasting and regularization literature (seasonal naïve benchmark; ridge and Elastic Net objective functions; standard accuracy metrics; Diebold–Mariano predictive accuracy test).

2.5.1. Mandatory Baseline: Seasonal Naïve

A seasonal naïve model was used as the required benchmark, as shown by Equation (1):

$$\hat{y}_t^{(SN)} = y_{t-12}. \quad (1)$$

This baseline is intentionally strong for seasonal tourism series and provides a conservative reference for incremental value.

2.5.2. Regularized Dynamic Regression

The main model class is dynamic regression with autoregressive and exogenous inputs:

$$y_t = \beta_0 + \sum_{i \in \mathcal{L}_y} \phi_i \, y_{t-i} + \sum_{j=1}^{p} \sum_{\ell=0}^{L} \beta_{j,\ell} \, x_{j,t-\ell} + \sum_{m=1}^{11} \gamma_m D_{m,t} + \varepsilon_t, \quad (2)$$

where $\mathcal{L}_y = \{1, 2, 12\}$, $D_{m,t}$ are month-of-year indicators (one omitted category), and $L = 3$ for exogenous lags.

Ridge (L2) regularization. Parameters are estimated by minimizing, Equation (3) ridge objective:

$$\min_{\theta} \sum_{t \in \mathcal{T}_{train}} (y_t - \hat{y}_t(\theta))^2 + \lambda \parallel \beta \parallel_2^2, \quad (3)$$

which is well-known to stabilize estimation under collinearity and improve out-of-sample performance (Hoerl & Kennard, 1970).

Elastic Net (L1 + L2) regularization. As an alternative that can shrink and partially select correlated predictors, Equation (4) Elastic Net objective:

$$\min_{\theta} \sum_{t \in \mathcal{T}_{train}} (y_t - \hat{y}_t(\theta))^2 + \lambda \left(\alpha \parallel \beta \parallel_1 + (1 - \alpha) \frac{1}{2} \parallel \beta \parallel_2^2 \right), \quad (4)$$

which is specifically motivated for settings with many correlated regressors.

The exogenous feature space in this application is intentionally "rich" and overlapping: multiple Wikipedia pageviews across related destination topics/languages and multiple climate covariates with distributed lags. This design naturally induces strong multicollinearity within and across signal families (e.g., lagged pageviews of related pages, temperature–dewpoint coupling), while the monthly sample size per rolling training window remains limited. In such settings, classical unpenalized regression and stepwise selection can become unstable and sensitive to small-sample perturbations.

Relative to other penalized linear alternatives, Lasso (pure L1) can yield sparse models, but it tends to select only one predictor from a correlated cluster and ignore the rest, which is undesirable when correlated proxies jointly represent a broader latent construct (e.g., "destination attention"). Ridge (pure L2) is highly effective for stabilization under multicollinearity, but it retains all predictors and therefore does not provide any degree of parsimony or partial selection when some exogenous signals are weak or redundant. Elastic Net combines L1 and L2 penalties to balance these trade-offs: it supports partial sparsity while also encouraging grouped retention of correlated predictors, improving robustness when predictors form correlated clusters (Zou & Hastie, 2005; Hong et al., 2023; Afreixo et al., 2024).

Finally, we prioritize Elastic Net over dimension-reduction approaches (e.g., principal component analysis (PCA), principal component regression (PCR), and partial least squares (PLS)) and more complex non-linear learners because the primary goal is operational nowcasting under repeated rolling re-estimation with a transparent, reproducible model structure. Elastic Net preserves a direct mapping between forecast performance and interpretable "families" of exogenous signals (attention vs. climate) while maintaining stability under correlated predictors—an advantage repeatedly noted in applied forecasting contexts with exogenous-variable selection problems (Sowinski, 2021).

In this application, regularization is not only a technical remedy for multicollinearity but it also encodes a conservative modeling stance about how open proxies should be used for tourism meaning-making under uncertainty. The exogenous design matrix intentionally includes overlapping and correlated indicators (multiple Wikipedia pageviews across related topics/languages and multiple climate variables with lags), which can be viewed as noisy, partially redundant measurements of broader constructs (e.g., destination-related attention and environmental comfort). Shrinkage therefore acts as a safeguard against over-interpreting any single proxy: it distributes predictive weight across correlated "families" of signals and stabilizes estimates in rolling re-estimation settings (Hastie et al., 2009; James et al., 2013).

This choice has direct interpretability consequences. Because shrinkage alters coefficient magnitudes (and, for Elastic Net, may partially select among correlated predictors), coefficients are not interpreted causally and are not treated as standalone measures of substantive effect sizes. Instead, interpretation is restricted to (i) whether exogenous families add incremental predictive value beyond autoregressive dynamics under a given regime, and (ii) whether the overall contribution of attention/climate signals is stable across evaluation windows. The primary claims of the study therefore remain predictive and operational (out-of-sample accuracy under rolling-origin evaluation), while conceptual interpretation is made at the level of robust, reproducible signal families rather than individual proxy coefficients.

2.5.3. Time-Series Benchmarks

To contextualize performance against established forecasting families, the pipeline also evaluates:

- SARIMAX (seasonal ARIMA with optional exogenous terms) as a classical statistical comparator.
- A state-space benchmark (structural/unobserved-components style) to represent robust smoothing-based dynamics under noise and shocks.

(Implementation details—orders, estimation routines, and any default constraints—should mirror what was executed in the provided scripts to maintain full reproducibility.)

2.6. Backtesting Protocol: Rolling-Origin One-Step-Ahead Evaluation

Performance was assessed using a rolling-origin (expanding-window) backtest that replicates operational nowcasting. For each forecast origin $\tau = \tau_0, \ldots, T-1$:

(i) Fit the model on $\{(y_t, x_t)\}_{t=1}^{\tau}$.
(ii) Produce the one-step-ahead nowcast $\hat{y}_{\tau+1|\tau}$.
(iii) Store the forecast error $e_{\tau+1} = \hat{y}_{\tau+1|\tau} - y_{\tau+1}$.

This design follows best practice for out-of-sample forecast evaluation in time-series, where re-estimation at each origin reflects real deployment conditions (Bergmeir et al., 2018; Fiorucci & Louzada, 2020).

2.7. Accuracy Metrics

For each model, the following metrics were computed over the rolling backtest:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}, MAE = \frac{1}{n}\sum_{t=1}^{n} | e_t |, \quad (5)$$

$$sMAPE = \frac{100}{n}\sum_{t=1}^{n} \frac{| y_t - \hat{y}_t |}{(| y_t | + | \hat{y}_t |)/2}. \quad (6)$$

Additionally, results were reported relative to the seasonal naïve baseline, using percentage improvement computed consistently per metric window.

2.8. Post-Shock Operational Validity: Evaluation Windows

To ensure the study remains policy-relevant and explicitly "post-shock robust", all accuracy metrics and statistical comparisons are reported not only for the full rolling evaluation span but also for two regime-specific windows that reflect substantively different operating conditions. Specifically, we report results for FULL, covering the entire rolling evaluation period as a global benchmark; STRESS (2019–2021), capturing the COVID-19 disruption and early recovery dynamics; and POST-2022, representing a post-reopening operational monitoring context characterized by attenuated irregular volatility and re-emerging seasonal persistence. Importantly, "operational" does not imply a complete return to pre-2020 levels or full "normalization"; rather, it denotes a later-phase monitoring environment in which tourism dynamics are less dominated by abrupt shock mechanics than during 2019–2021. This windowed reporting supports stronger inference by separating crisis-period performance from post-reopening performance, thereby avoiding conclusions driven by a single pooled regime.

2.9. Statistical Testing: Diebold–Mariano Against the Baseline

To test whether observed gains over the seasonal naïve baseline are statistically meaningful, the study applies the Diebold–Mariano (DM) test for equal predictive accuracy.

Let $L(\cdot)$ be a loss function (e.g., squared error). Define the loss differential, Equation (7) Diebold–Mariano:

$$d_t = L\left(e_t^{(model)}\right) - L\left(e_t^{(baseline)}\right). \quad (7)$$

The null hypothesis is $H_0 : \mathbb{E}[d_t] = 0$. The DM statistic is constructed from the sample mean $\overline{d}$ and an autocorrelation-consistent estimate of $\mathrm{Var}\left(\overline{d}\right)$ (for one-step-ahead forecasts, $h = 1$), producing a test statistic and p-value per evaluation window. A negative $\overline{d}$ (and typically negative DM statistic under conventional implementations) indicates that the model's loss is lower than the baseline.

2.10. Implementation and Reproducibility

All processing, modeling, and evaluation steps were implemented in Python 3.13.5 using a script-based pipeline (data ingestion, exogenous data collection, feature construction, rolling backtesting, figure/table export, and post-shock evaluation). All processing, modeling, and evaluation steps were implemented in Python using a script-based pipeline (data ingestion, exogenous data collection, feature construction, rolling backtesting, figure/table export, and post-shock evaluation). Reproducibility is ensured through deterministic monthly indexing and lag construction; an expanding-window backtesting design that prevents look-ahead leakage; persisted outputs (tables and figures) written to versioned output folders; and explicit environment management (e.g., a conda environment) with documented dependencies.

3. Results

3.1. Time-Series Structure and Shock Imprint (Robust STL Decomposition)

To characterize the intrinsic dynamics of Ecuador's monthly international arrivals and to motivate the subsequent model evaluation design, a robust STL (Seasonal–Trend decomposition using LOESS) was applied to the observed arrivals series. Figure 1 summarizes the resulting decomposition into observed, trend, seasonal, and residual components. STL is routinely used to separate persistent seasonality from slower-moving trends and irregular

shocks in monthly socioeconomic time-series, which is particularly relevant when tourism demand is exposed to abrupt disruptions.

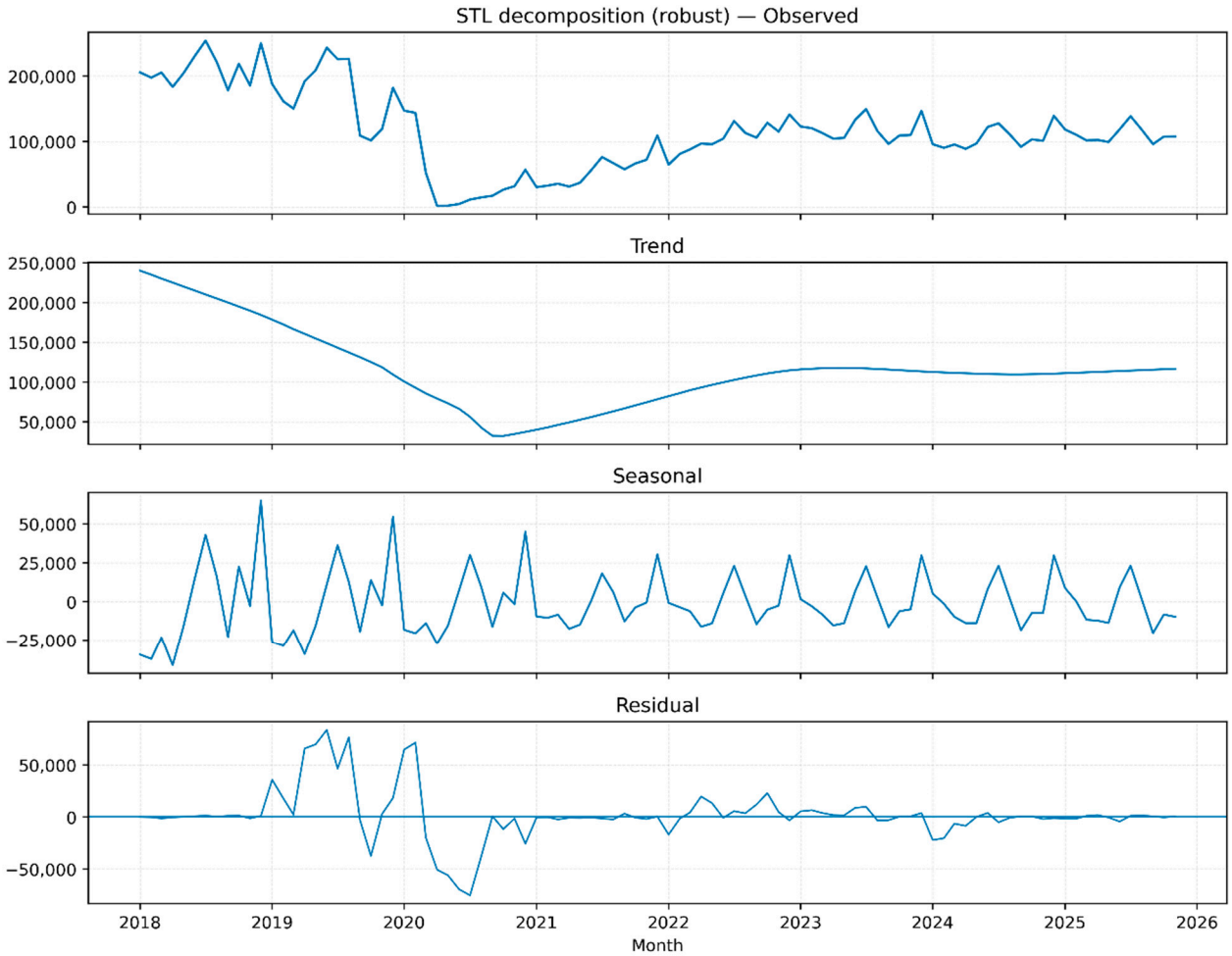


Figure 1. Robust STL decomposition of monthly international arrivals to Ecuador (MIN-TUR). The observed series is decomposed into trend, seasonal, and residual components. The trend exhibits a sharp discontinuity in early 2020 consistent with the COVID-19 shock, while annual seasonality persists across the sample, and the largest irregular residual excursions concentrate in the 2019–2021 stress window.

The decomposition evidences a clear structural break associated with the COVID-19 shock, visible as (i) a rapid collapse in the observed series and (ii) a pronounced discontinuity in the trend component around early 2020 (Figure 1). Quantitatively, the series mean declines from ~193k arrivals/month in 2018–2019 to ~49k arrivals/month in 2020–2021, with the trough at 1439 arrivals (April 2020)—a ~99.3% drop relative to the 2018–2019 mean. This pattern aligns with the well-documented susceptibility of tourism demand to mobility restrictions and pandemic control measures, which generate abrupt, non-regular demand contractions and slow recoveries.

Despite the shock, the seasonal component indicates that annual seasonality remains present across the sample (Figure 1), consistent with the known seasonality of international tourism flows. However, the post-shock regime exhibits a reduced scale: a simple month-of-year spread (mean of peak month minus trough month) decreases from ~96k (2018–2019) to ~45k (2022–2025), reflecting that even as seasonality persists, it operates around a lower post-shock level. The residual component concentrates much of its large-amplitude irregular variation in the 2019–2021 window, while the later period shows visibly smaller residual excursions (Figure 1), indicating a partial stabilization of arrivals after borders reopened

and tourism operations largely resumed, although the level and volatility structure remain distinct from the pre-2020 baseline.

These findings are consequential for model design and reporting. Because tourism-demand time-series frequently exhibit regime changes and structural breaks that degrade the validity of "single-regime" forecasting conclusions, the decomposition motivates (and justifies) a windowed evaluation strategy that separates stress and operational periods—rather than reporting only a single aggregate metric.

3.2. Exogenous Signals: Collinearity and Motivation for Regularized Models

The exogenous feature set was inspected for collinearity across Wikipedia pageviews (multiple articles and languages) and ERA5-Land climate covariates (temperature/dewpoint/precipitation/humidity, with multiple lags). Figure 2 summarizes a representative subset of the resulting correlation structure and shows a clear block pattern: (i) strong within-source collinearity across lagged versions of the same signal (expected in persistent attention and climate processes), and (ii) non-trivial cross-correlation among several climate variables (e.g., temperature–ddewpointcoupling) and their lags.

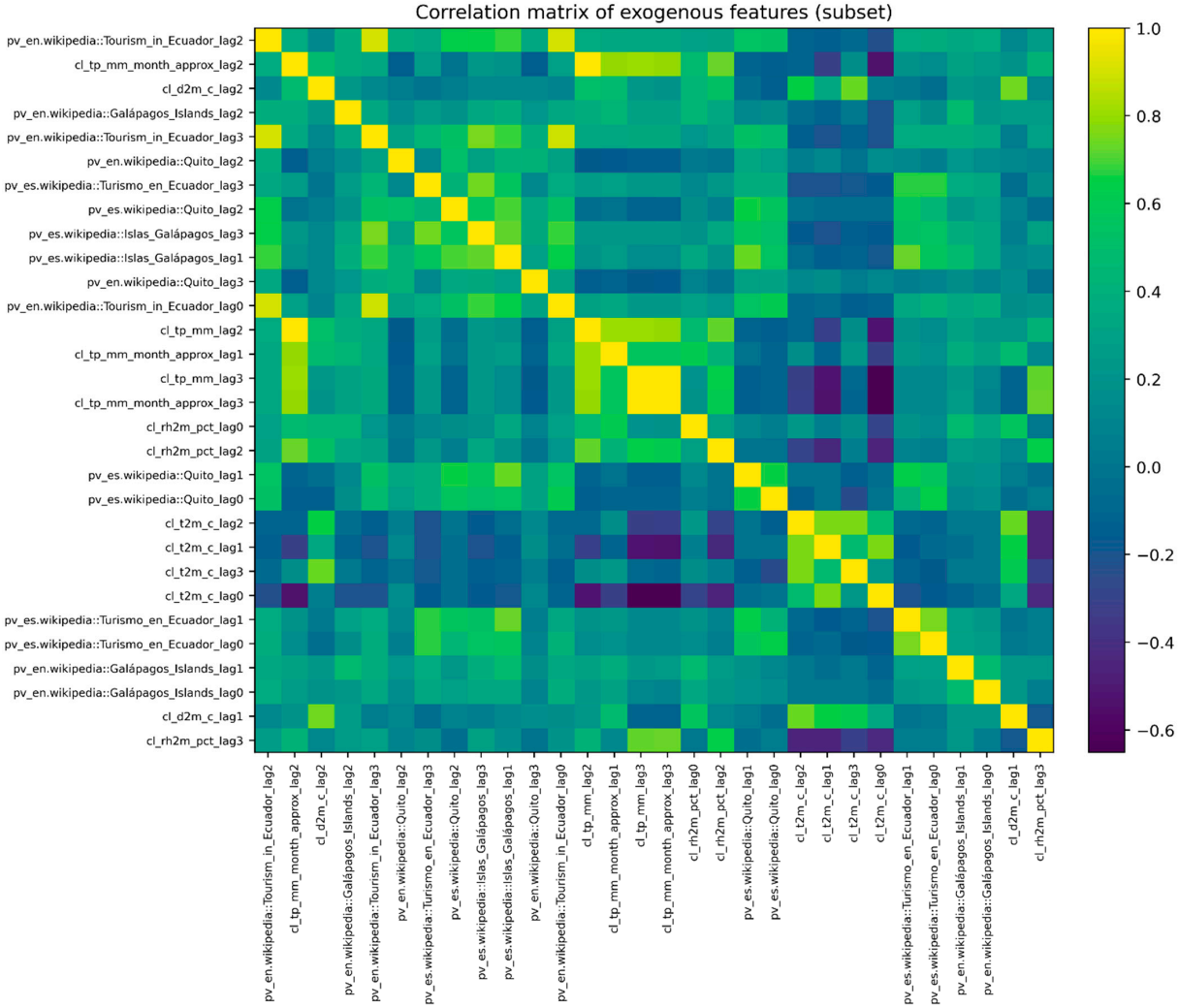


Figure 2. Correlation heatmap of a representative subset of exogenous features used for nowcasting. The plot highlights strong within-family correlations among lagged Wikipedia pageviews and among lagged climate variables, as well as cross-correlations between selected climate covariates. For readability, only a subset is shown; the full feature set exhibits the same block-collinearity pattern. This structure motivates the use of regularized dynamic regression (Ridge/Elastic Net) to stabilize coefficient estimates and improve out-of-sample nowcasting performance.

This correlation structure is consequential for modeling. In dynamic regression settings, highly correlated predictors inflate coefficient variance and make unregularized estimates unstable, particularly when the sample is limited and the feature space contains multiple lags and conceptually overlapping proxies of "tourist attention" (e.g., Wikipedia visits to related pages). In this context, Ridge regression provides a principled stabilization mechanism by shrinking coefficients and reducing variance under multicollinearity, often improving out-of-sample performance relative to unpenalized regression (Hoerl & Kennard, 1970). In parallel, Elastic Net combines shrinkage with sparse selection while handling correlated groups of predictors (a relevant property when attention and climate signals form correlated "families"), which makes it suitable when the aim is both predictive robustness and parsimonious signal use (Zou & Hastie, 2005).

The practical relevance of this design choice is consistent with prior tourism-forecasting evidence using online attention data. Search intensity and related digital traces can improve short-horizon tourism-demand forecasts, but they often exhibit strong correlation across languages/keywords/platforms and can contain artifacts that degrade forecast quality if modeled naïvely—precisely the type of risk that regularization is meant to mitigate (Bokelmann & Lessmann, 2019; Dergiades et al., 2018; Tibshirani, 1996).

Complementarily, Wikipedia pageviews have been proposed as a useful "big data" input for tourism statistics and demand estimation, reinforcing the motivation for including pageview-based signals while treating redundancy/collinearity as a first-order modeling constraint.

3.3. One-Step-Ahead Nowcasting Performance: Full Evaluation Window (Rolling Backtest)

To quantify operational nowcasting accuracy, the study uses a rolling-origin, one-step-ahead backtest that mimics real-time deployment by repeatedly re-estimating models and predicting the next month. This evaluation design is widely recommended for time-series forecasting because it preserves temporal ordering and reflects decision-making under information constraints (Hyndman & Koehler, 2006). Consistent with standard forecasting practice, performance is summarized using RMSE, MAE, and sMAPE, which jointly capture scale-dependent errors and relative (percentage) deviation.

As reported in Table 1, the ridge-regularized autoregressive model is the top performer across the full rolling backtest window, achieving the lowest RMSE and MAE among all candidates. Relative to the required seasonal naïve baseline, ridge_ar reduces RMSE by ~79% (14,434 vs. 69,262), MAE by ~77% (10,826 vs. 46,169), and sMAPE by ~71% (15.77% vs. 54.80%). Importantly, the improvement in predictive accuracy versus the baseline is also supported by a Diebold–Mariano test on the full window ($p \approx 0.004$), indicating a statistically significant reduction in forecast loss. The Diebold–Mariano framework is a standard approach for comparing predictive accuracy across competing forecasts under an explicit loss function.

Based on the rolling-origin evaluation, the final operational nowcasting system is the ridge-regularized autoregressive model (*ridge_ar*). Its one-step-ahead nowcast is defined as

$$\hat{y}_{t \mid t-1} = \hat{\beta}_0 + \sum_{i=1}^{p} \hat{\phi}_i \, y_{t-i} + \sum_{m=2}^{12} \hat{\gamma}_m \, D_{m,t}, \quad (8)$$

where y_t denotes monthly international arrivals, y_{t-i} are autoregressive lags, and $D_{m,t}$ are month-of-year indicators (one omitted category). Parameters are estimated at each forecast origin under an expanding-window scheme by solving the ridge-penalized least squares problem (intercept unpenalized):

$$(\hat{\beta}_0, \hat{\phi}, \hat{\gamma}) = \arg\min_{\beta_0, \phi, \gamma} \sum_{\tau=1}^{T_t} (y_\tau - \beta_0 - \sum_{i=1}^{p} \phi_i y_{\tau-i} - \sum_{m=2}^{12} \gamma_m D_{m,\tau})^2 + \lambda(\| \phi \|_2^2 + \| \gamma \|_2^2), \quad (9)$$

with λ selected within the training window.

Corresponding result. Over the full rolling backtest window, *ridge_ar* achieves RMSE = 14,434, MAE = 10,826, and sMAPE = 15.77%, corresponding to a 79.2% RMSE reduction versus the seasonal naïve baseline.

Table 1. Overall one-step-ahead nowcasting accuracy from a rolling-origin (expanding-window) backtest over the full evaluation period. Models are ranked by RMSE (lower is better). MAE and sMAPE are reported to complement scale-dependent and percentage-based accuracy. The rightmost column reports percent RMSE reduction relative to the required seasonal_naive baseline (positive values indicate improvement).

Rank	Model	RMSE	MAE	sMAPE (%)	RMSE Reduction vs. Seasonal_Naive (%)
1	ridge_ar	14,434	10,826	15.77	79.2
2	state_space	22,383	17,108	21.39	67.7
3	ridge_exog	33,358	23,304	33.27	51.8
4	SARIMAX	41,140	29,859	41.97	40.6
5	elasticnet_exog	47,281	32,660	41.81	31.7
6	seasonal_naive	69,262	46,169	54.80	0.0

Moreover, the improvement versus the baseline is statistically supported by a Diebold–Mariano test (DM = −3.007, $p = 0.0039$).

Beyond aggregate metrics, Figure 3 provides a visual diagnostic of performance over time. The ridge_ar nowcast tracks the main recovery dynamics and recurring within-year fluctuations more closely than the baseline, while larger deviations concentrate around abrupt demand shifts, where tourism time-series are known to be harder to predict under purely seasonal heuristics.

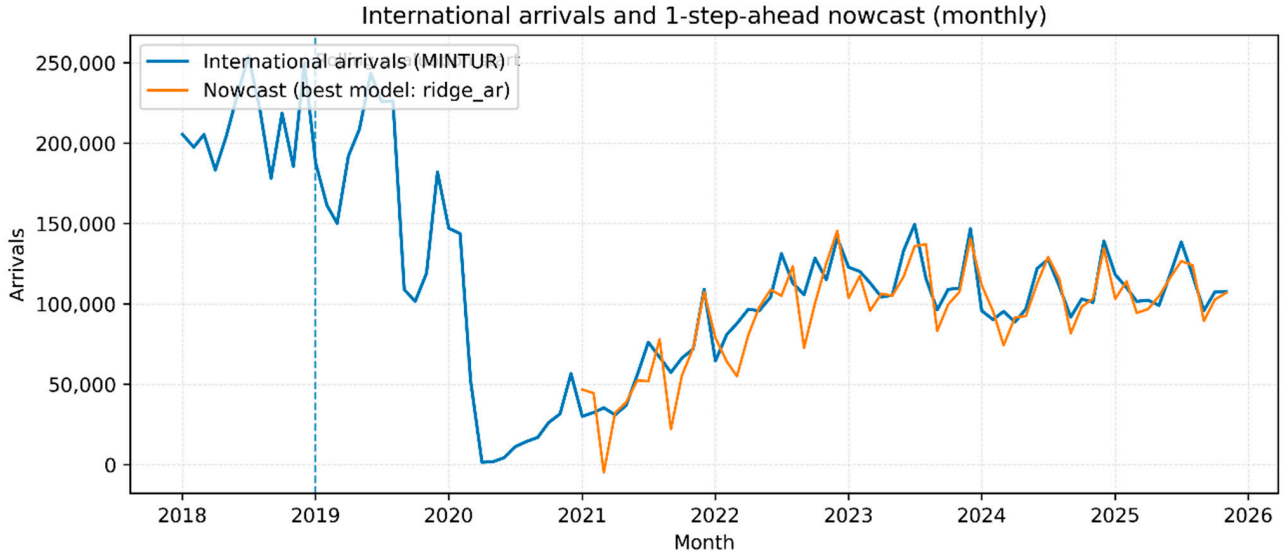


Figure 3. International arrivals (observed) and one-step-ahead nowcasts from the best-performing model (ridge_ar) produced under rolling-origin (expanding-window) evaluation. The vertical dashed line marks the start of the out-of-sample backtest period. Nowcasts are generated sequentially using information available up to the previous month, illustrating how the model tracks both seasonal dynamics and post-shock stabilization without look-ahead information.

3.4. Operational Validity: Windowed Evaluation (Stress vs. Post-Reopening Operational)

Because the study targets one-step-ahead monthly nowcasting, the core operational question is whether predictive accuracy remains stable when the data-generating process shifts under structural shocks. For that reason, nowcasting performance is assessed

separately for a COVID-era stress regime and a later post-reopening operational regime, rather than only on a single pooled window. This is an evaluation (not a horizon) distinction: all models still generate t + 1 nowcasts, but their accuracy is reported within windows that reflect substantively different conditions. This separation is consistent with the tourism-forecasting literature emphasizing crisis-period evaluation and robustness under COVID-era dynamics (Llewellyn et al., 2023).

Accordingly, Table 2 reports windowed rolling-origin, one-step-ahead accuracy for (i) the full sample, (ii) a stress-test window (2019–2021) capturing the shock and immediate disruption, and (iii) a post-COVID operational window (2022–2025) intended to reflect real-world nowcasting conditions after reopening and partial stabilization. Across all three windows, ridge_ar ranks first by RMSE, indicating that a regularized autoregressive specification provides the most stable operational performance. In the stress window, ridge_ar achieves a large RMSE reduction relative to the seasonal_naive baseline (≈81%), showing strong robustness under severe disruption (Table 2).

Table 2. Windowed one-step-ahead nowcasting accuracy under rolling-origin evaluation, reported separately for FULL, STRESS (2019–2021), and POST-2022 windows. For each window, *n* denotes the number of backtest months. "RMSE improvement vs. seasonal_naive" reports the percent RMSE change relative to the baseline within the same window (positive values indicate improvement; negative values indicate deterioration), enabling direct regime-specific assessment of operational usefulness.

Window	Model	*n*	MAE	RMSE	sMAPE (%)	RMSE Improvement vs. Seasonal_Naive (%)
FULL	ridge_ar	59	11,066.0	14,434.3	14.80	79.16
FULL	state_space	59	17,156.4	22,382.6	22.45	67.69
FULL	ridge_exog	59	25,039.9	33,357.8	30.43	51.85
FULL	elasticnet	59	27,009.4	36,929.9	34.13	46.69
FULL	ARIMA	59	31,802.5	45,625.4	39.11	34.14
FULL	seasonal_naive	83	50,822.8	69,262.5	56.49	0.00
POST (2022–2025)	ridge_ar	47	10,267.5	13,165.3	12.65	56.26
POST (2022–2025)	state_space	47	15,880.8	20,363.7	19.50	32.35
POST (2022–2025)	seasonal_naive	47	22,262.0	30,101.1	30.67	0.00
POST (2022–2025)	ARIMA	47	26,776.4	35,745.3	33.45	−18.75
POST (2022–2025)	elasticnet	47	30,904.2	41,626.1	39.23	−38.29
POST (2022–2025)	ridge_exog	47	26,789.9	33,357.8	35.13	−10.82
STRESS (2019–2021)	ridge_ar	12	15,246.4	18,588.2	49.44	81.30
STRESS (2019–2021)	state_space	12	23,299.2	28,965.3	67.64	70.85
STRESS (2019–2021)	ridge_exog	12	31,985.5	37,257.2	91.71	62.51
STRESS (2019–2021)	elasticnet	12	34,955.4	39,756.0	97.90	60.00
STRESS (2019–2021)	ARIMA	12	39,942.5	46,355.4	109.87	53.36
STRESS (2019–2021)	seasonal_naive	36	84,774.7	99,392.2	118.11	0.00

Interpretively, the post-2022 attenuation of the attention-based exogenous value is consistent with three non-exclusive mechanisms. First, once irregular volatility attenuates and seasonal persistence re-emerges after reopening, the arrivals series often exhibits strong seasonal persistence, so autoregressive/state-space structures can absorb most of the predictable variation; in that setting, broad attention proxies may offer limited incremental information beyond what is already encoded in lagged dynamics and month-of-year effects. Second, the proxy–behavior mapping may undergo concept drift: after the pandemic, online attention can reflect news cycles, risk salience, or generalized curiosity rather than pre-trip planning, weakening the short-lag link between information seeking and realized travel (Gama et al., 2014). Third, digital traces are subject to platform and representativeness

dynamics—changes in how users discover information (e.g., substitution toward social media/online travel agencies (OTAs) or shifts in referral patterns) can alter Wikipedia pageview volumes independently of true travel intention, a known failure mode in proxy-based nowcasting ("algorithm dynamics" and measurement instability) (Lazer et al., 2014). These considerations provide a plausible explanation for why exogenous-enriched models can be more helpful under disruption/recovery yet underperform in a later post-reopening operational regime, reinforcing the value of regime-aware evaluation rather than assuming proxy gains are time-invariant.

In the post-2022 operational window, ridge_ar remains best-in-class (≈56% RMSE reduction vs. seasonal_naive), and the state-space benchmark remains competitive (≈32% RMSE reduction). Notably, ridge_exog does not improve upon the seasonal baseline post-2022 (negative RMSE improvement), suggesting that the incremental value of the specific exogenous proxy set is time-varying and may be more beneficial under disruption/recovery than under the later, more stable regime (Table 2).

3.5. Statistical Significance vs. Baseline: Diebold–Mariano (DM) Tests

To ensure that performance differences are not only descriptive but also statistically defensible, the study applied Diebold–Mariano (DM) tests to compare each candidate model against the mandatory seasonal_naive baseline across the same evaluation windows used in the operational analysis. The DM framework is a standard approach for testing equal predictive accuracy using a loss-differential series computed from paired forecast errors.

In the implementation used here, the loss differential was defined so that negative DM statistics indicate lower forecast loss than seasonal_naive (i.e., better predictive accuracy), while positive DM statistics indicate worse performance than the baseline. Consequently, statistically significant improvements are identified by DM < 0 with $p < 0.05$ (two-sided). Because DM tests can be sensitive in finite samples—particularly in short windows—results for the stress-test window ($n = 12$ months) should be interpreted cautiously, and small-sample considerations discussed in the forecasting literature remain relevant.

Table 3 reports DM statistics, two-sided p-values, and the number of backtest months per window for each model-versus-baseline comparison. In tourism-demand forecasting studies, DM-type tests are widely used to support claims of forecast improvements beyond visual comparisons and point metrics (Ho et al., 2022).

3.6. Error Diagnostics and Reliability (Distribution and Time-Varying Error)

Because tourism-demand time-series can experience abrupt structural breaks (e.g., COVID-19 and its recovery phase), aggregate accuracy metrics alone can mask when a nowcasting system is operationally reliable versus when it is exposed to regime shifts. This motivates an explicit diagnostic view of the one-step-ahead errors over time and their distribution, which is standard practice when evaluating forecasting performance under interruptions and non-stationary conditions (Hyndman & Rostami-Tabar, 2025; Sun et al., 2017).

As shown in Figure 4a, the rolling one-step-ahead error series (defined as prediction − actual) exhibits clear time variation. The largest negative deviations cluster in the early recovery period (e.g., 2021–early 2022), indicating episodes of systematic under-prediction during rapid rebounds. After 2022, the error oscillations become visibly more contained around zero, consistent with a transition from "shock/recovery" dynamics to a more regular operational regime. Quantitatively, the dispersion decreases substantially from the earliest evaluation months to the post-2022 period (error standard deviation drops from ~18.4k to ~12.2k arrivals, and the 95th percentile absolute error reduces from ~37.6k to ~27.5k arrivals), supporting the visual evidence that the system becomes more stable once the post-shock regime is established. This pattern is consistent with tourism-forecasting

evidence showing that COVID-era regime changes amplify forecast risk and that post-shock evaluation is essential to judge operational usefulness (Hyndman & Koehler, 2006).

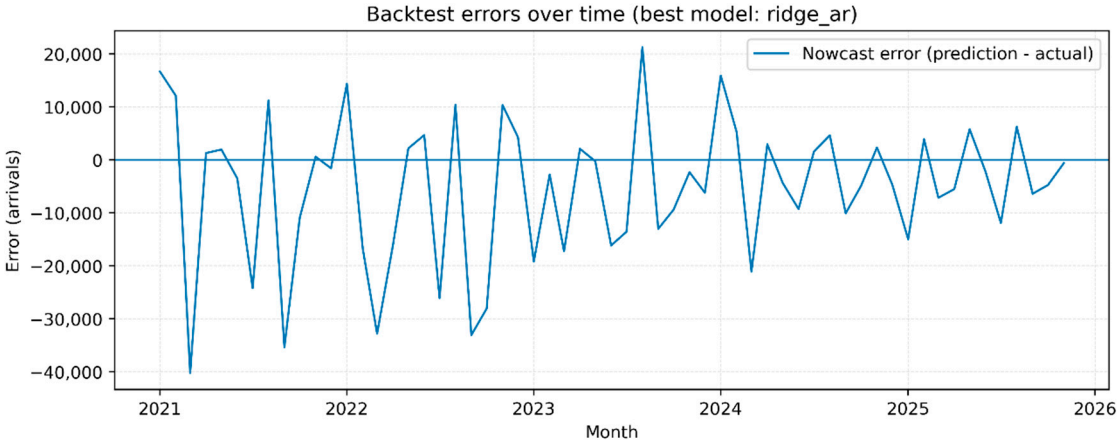
Table 3. Diebold–Mariano (DM) tests of predictive accuracy versus the seasonal_naive baseline, by evaluation window (two-sided *p*-values). DM statistics are computed from one-step-ahead loss differentials across rolling backtest months (*n*). Negative DM values indicate lower loss (better predictive accuracy) than the baseline under the chosen loss function; reported *p*-values assess whether improvements are statistically distinguishable from zero.

Window	Model (vs. Seasonal_Naive)	DM Stat	*p*-Value (Two-Sided)	*n*
FULL	ridge_ar	−3.007	0.0039	59
FULL	state_space	−2.509	0.0149	59
FULL	elasticnet_exog	1.055	0.2959	59
FULL	ridge_exog	−0.962	0.3401	59
FULL	SARIMAX	0.455	0.6506	59
POST_2022	ridge_ar	−2.621	0.0118	47
POST_2022	SARIMAX	2.063	0.0448	47
POST_2022	state_space	−1.764	0.0844	47
POST_2022	elasticnet_exog	1.648	0.1062	47
POST_2022	ridge_exog	0.428	0.6707	47
STRESS_2019_2021	ridge_ar	−2.480	0.0306	12
STRESS_2019_2021	state_space	−2.326	0.0402	12
STRESS_2019_2021	ridge_exog	−2.103	0.0593	12
STRESS_2019_2021	SARIMAX	−1.056	0.3136	12
STRESS_2019_2021	elasticnet_exog	−1.054	0.3144	12

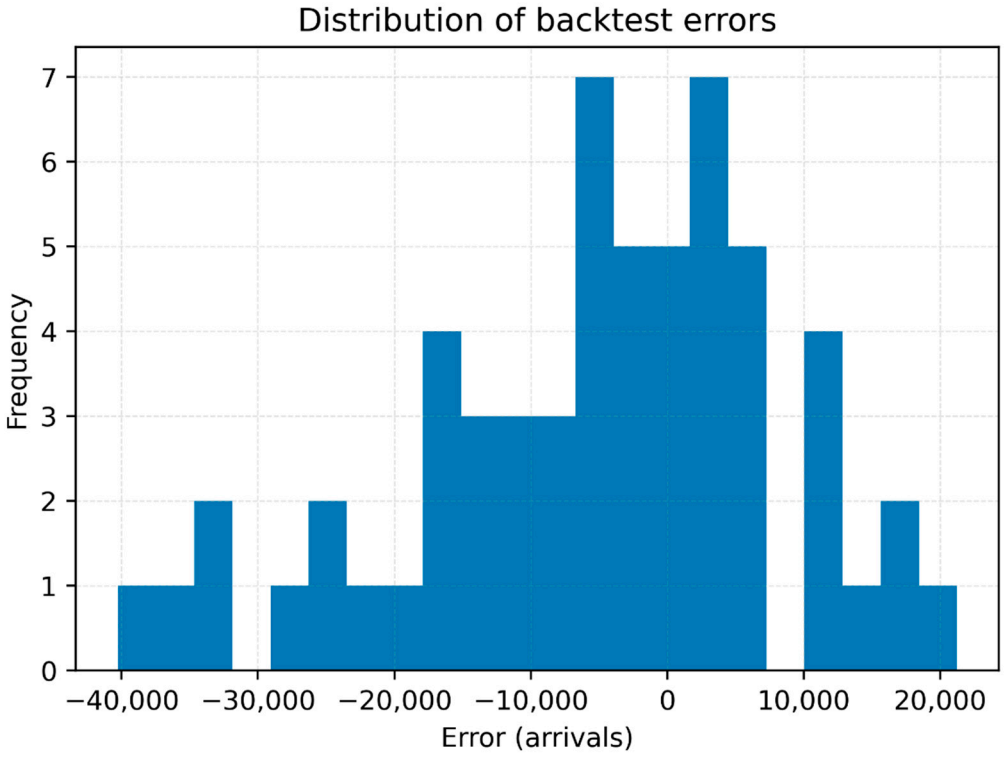
The distributional view in Figure 4b complements the temporal diagnostics. The histogram is broadly unimodal but shows a mild left-skew, reflecting a small but persistent tendency toward under-prediction (the mean error over the evaluation sample is approximately −5.4k arrivals). Importantly, the bulk of mass concentrates near zero in the post-2022 regime, while the heavier tails are driven by the earlier recovery months—consistent with the time-localized extremes observed in Figure 4a. In other words, the dominant source of "large misses" is not persistent model instability but a limited set of months associated with rapid, non-linear transitions that are known to challenge tourism-demand forecasting systems.

3.7. Practical Takeaway: What the Model Enables for Tourism Monitoring

The results indicate that the proposed workflow can be operationalized as a monthly tourism monitoring tool that produces reproducible one-step-ahead nowcasts as soon as the latest month's official arrivals are available and the corresponding exogenous signals have been updated. Concretely, the operational run requires only (i) the official monthly arrivals series (target), (ii) openly accessible online attention proxies (Wikipedia pageviews) that have been shown to carry timely information for tourism activity estimation, and (iii) openly accessible gridded climate reanalysis inputs (ERA5-Land monthly aggregates) that provide consistent country-scale climate covariates for the same monthly timeline. The use of online attention metrics as high-frequency leading indicators is consistent with prior evidence in tourism-demand monitoring and forecasting.



(**a**)



(**b**)

Figure 4. Error diagnostics for the best one-step-ahead nowcasting model under rolling backtest evaluation. (**a**) Time-varying nowcast error over the evaluation period, defined as prediction—actual international arrivals (MIN-TUR), highlighting periods of higher forecast risk during shock/recovery dynamics. (**b**) Histogram of one-step-ahead errors across evaluation months, summarizing bias and tail behavior. Negative values denote under-prediction, while tighter concentration around zero indicates greater operational stability.

4. Discussion

4.1. Interpretation of the Findings in an Ecuadorian, Latin American, and Global Context

The empirical pattern in this study is consistent with a central regularity in tourism-demand time-series: at a monthly frequency, a substantial fraction of the predictable variation is often carried by the series' own dynamics (seasonality, persistence, and gradual reversion), such that well-specified autoregressive structures can remain highly competitive even when additional "big-data" signals are available. This aligns with the long-standing

tourism-demand forecasting literature emphasizing that model performance is frequently driven by how well the approach captures serial dependence and seasonality rather than by adding many external regressors indiscriminately (D. C. Wu et al., 2025).

In operational terms, this matters for Ecuador and many Latin American contexts because "fast, low-cost, reproducible" monitoring solutions are often constrained by data availability and dissemination lags. A model class that can deliver reliable monthly nowcasts primarily from the official series and a small, robust set of features reduces dependence on proprietary data and on complex data integration pipelines that may not be sustainable for public-sector analytics teams. At the same time, the windowed evaluation highlights an important nuance: the value of open exogenous signals (digital attention and climate reanalysis) is not uniform across regimes. The evidence that exogenous-enriched models can be more helpful in shock-like periods but not necessarily in stabilized post-shock periods mirrors broader findings that the predictive contribution of big-data proxies is context-dependent, sensitive to behavioral shifts, and prone to degradation when the underlying "interest → travel" mapping changes (Mikulić & Baumgärtner, 2025; D. C. Wu et al., 2025).

From a global perspective, the post-2020 tourism environment is widely recognized as a structural-break setting: demand dynamics changed abruptly with mobility restrictions and then evolved during reopening. Forecasting in such environments benefits from evaluation designs that explicitly separate "stress" windows from "operational" windows (i.e., periods where nowcasting is intended to be used routinely). The two-window framing used here—stress-test vs. post-2022—directly targets this methodological requirement and supports interpretable guidance for deployment: a model may look strong on average while failing precisely when decision-makers need robustness most (Bergmeir et al., 2018; Zheng & Nikolopoulos, 2025).

4.2. Why Regularization Was Methodologically Justified, and Why It Did Not Guarantee Gains in Every Window

A key contribution of the work is the explicit treatment of multicollinearity among exogenous signals. Correlated proxies—such as multiple Wikipedia pageview series for related topics and climate variables that co-move seasonally—can inflate variance in non-regularized regressions and lead to unstable coefficient estimates, especially in small-to-moderate samples typical of monthly tourism series. Ridge regression is a canonical remedy in precisely this setting because it shrinks coefficients continuously and stabilizes estimation under collinearity; Elastic Net extends this by combining ridge-like stabilization with sparsity-inducing behavior when some predictors are redundant (Alis et al., 2015; J. Li et al., 2018; Song & Li, 2008).

However, regularization addresses estimation stability, not predictive relevance. The post-2022 window results—where purely autoregressive or state-space structures outperform models that incorporate the broader exogenous set—suggest that in the stabilized regime, the open proxies may be (i) weakly aligned with inbound-arrival variation at monthly resolution, (ii) too noisy relative to the signal in the target, or (iii) structurally "shifted" in meaning after COVID-era behavioral changes (e.g., information seeking not translating to travel). This is consistent with recent tourism big-data forecasting reviews that emphasize heterogeneity of results and the need for careful feature selection, stability checks, and regime-aware evaluation, rather than assuming that more data sources automatically improve accuracy (H. Li et al., 2020; D. C. Wu et al., 2025).

More broadly, this pattern aligns with the cautionary literature on proxy-based nowcasting: performance gains from digital traces can degrade when the measurement process changes (platform design, ranking/referrals, user composition) even if the underlying phenomenon does not, and when "big data" signals are treated as stable sensors rather than evolving socio-technical traces (Lazer et al., 2014). In forecasting terms, this is a form

of concept drift, where the statistical relationship between predictors (attention) and the target (arrivals) changes over time; drift is especially plausible across a structural break such as COVID-19 and during the subsequent normalization of travel behavior (Gama et al., 2014). The two-window evidence in this study therefore supports a practical implication: open attention signals should be treated as conditionally useful features whose value must be monitored and revalidated across regimes, rather than assumed to provide persistent improvements in a stabilized operational setting.

4.3. Positioning Against Comparable Literature on Big-Data-Augmented Tourism Nowcasting

The approach is aligned with the "nowcasting with digital traces" paradigm popularized in economics and extended in tourism, where online activity is treated as a contemporaneous proxy for demand. Foundational work demonstrated that search intensity can track near-term economic behavior ("predicting the present"), motivating analogous applications to tourism demand (D. C. Wu et al., 2025).

Within tourism specifically, a substantial body of research has examined Google Trends and related search indices as demand predictors, often reporting improvements but also documenting instability across destinations, horizons, and structural breaks (Liu et al., 2025; D. C. Wu et al., 2025). More recent studies show that big-data augmentation can be particularly useful under disruptive conditions when conventional patterns break down—for example, during the pandemic—because online attention and information seeking can respond faster than official statistics (Fiorucci & Louzada, 2020; H. Li et al., 2020; Zheng & Nikolopoulos, 2025).

The present work contributes to this line in three differentiating ways that are particularly relevant:

1. Open-by-design signal set: the study uses Wikipedia pageviews (open access) and ERA5-Land (open reanalysis) alongside the official arrivals series, avoiding reliance on proprietary booking, payments, or platform data. The ERA5-Land dataset is specifically designed as a globally consistent land reanalysis suitable for downstream feature engineering at scale, which improves reproducibility and cross-country portability (Bergmeir et al., 2018).
2. Engineering-motivated model class: the emphasis on regularized dynamic regression directly addresses multicollinearity and small-sample instability, which are common in monthly tourism applications with many correlated proxies. This differs from many tourism studies that either (i) add external regressors to SARIMAX-like specifications without strong shrinkage, or (ii) deploy complex ML pipelines without explicitly diagnosing collinearity and coefficient stability (J. Li et al., 2018; Mikulić & Baumgärtner, 2025).
3. Regime-aware evaluation with statistical testing: separating stress vs. post-2022 operational windows and pairing the comparison with Diebold–Mariano tests strengthens the inferential claim beyond "metric deltas," while also making it clear where statistical evidence is weaker due to short samples (e.g., the 12-month stress window) (Bergmeir et al., 2018).

This positioning also clarifies why the findings are not contradictory to prior "big data helps" results: the study's evidence supports a more precise claim—open digital traces can add value, but the benefit is conditional on regime and on the stability of the proxy–demand linkage—which is increasingly emphasized in the most recent tourism-forecasting assessments (Mikulić & Baumgärtner, 2025; D. C. Wu et al., 2025).

A spatial lens further clarifies why the incremental value of digital attention proxies can be regime- and scale-dependent. Prior research shows that the predictive contribution of online search/attention signals varies with the unit of analysis: models estimated at

finer geographic resolution (cities, destinations, or attractions) often capture a more direct and localized "attention → visitation" mapping, whereas national aggregates can attenuate (or even offset) such signals through compositional effects across multiple destinations and substitution patterns. For example, evidence comparing country- versus city-level forecasting finds that the magnitude of forecast gains from search data depends on spatial scale, underscoring that proxy usefulness is not invariant to aggregation. Likewise, studies focusing on high-frequency tourism demand at destination/attraction scales emphasize heterogeneity and location-specific dynamics, which motivates spatiotemporal or high-capacity models tailored to localized demand processes (Huang & Zhang, 2024; K. Li et al., 2020). Complementarily, spatial analyses of "network attention" around tourism resources document strong geographic clustering of digital attention, implying that the signal is inherently spatially structured rather than spatially homogeneous. Interpreted in this context, our post-2022 underperformance of exogenous attention models is consistent with the idea that a single national arrivals series blends multiple destination-level attention–travel relationships; once aggregate demand re-stabilizes, autoregressive seasonality dominates and the remaining attention signal may be diluted (Zeng et al., 2025). This supports the practical implication that open-proxy nowcasting should be evaluated—and where feasible, operationalized—at destination level with spatially matched covariates, rather than relying exclusively on national aggregation.

4.4. What This Work Adds to Tourism Analytics and Monitoring Practice

The primary practical contribution is a reproducible, low-cost monthly nowcasting workflow that can be deployed with only a laptop and open data, producing consistent performance monitoring and interpretable model comparisons. This is a meaningful operational advance for destinations where timely official releases, disaggregated demand indicators, or commercial data partnerships may be limited or delayed.

Conceptually, the study reinforces an actionable methodological principle for tourism monitoring: in stable operating regimes, robust time-series structure may dominate; in disruption regimes, carefully regularized exogenous augmentation may help—but only if the proxies remain behaviorally aligned with travel outcomes. This principle is compatible with broader evidence that tourism demand is shaped by both persistent seasonal mechanisms and abrupt externalities (health crises, policy shifts, media discourse), and that model evaluation should be matched to the decision context rather than relying on a single aggregate period (Park et al., 2021).

Implications for tourism governance: platform dependence, surveillance logics, and algorithmic decision-making.

Beyond methodological validation, the regime-dependent findings have governance implications for how destinations operationalize "data-driven" tourism policy. A reproducible open-data nowcasting pipeline can strengthen public-sector accountability by reducing reliance on opaque proprietary datasets and vendor-specific analytics, enabling transparent documentation, replication, and audit of monitoring outputs. However, the results also underscore a conceptual tension: even "open" attention indicators remain platform-mediated traces whose measurement process is external to tourism agencies and may change over time. In that sense, proxy-based monitoring should be treated as a conditional evidence stream rather than an objective sensor of demand.

This matters for tourism governance because platform dependence and digital surveillance assumptions can quietly enter policy design. The critical literature on big-data systems emphasizes that digital traces are shaped by socio-technical infrastructures and can reproduce representativeness biases, algorithmic dynamics, and shifting meanings (Boyd & Crawford, 2012; Lazer et al., 2014). Related critiques highlight how data extraction

and behavioral prediction can consolidate power in ways that outpace democratic oversight ("surveillance capitalism") and normalize unequal data relations ("data colonialism") (Zuboff, 2019; Couldry & Mejias, 2019). In tourism settings, these concerns translate into practical cautions: attention signals can be informative for situational awareness, but their governance use should avoid equating "more monitoring" with better policy, especially when monitoring implies surveillance practices or restrictive digital controls that may have broader social and economic consequences (Gozgor et al., 2024).

Operationally, we recommend treating nowcasts as a decision support rather than automated decision-making. Good governance practice in tourism monitoring would therefore include (i) full transparency about data sources, model versions, and limitations; (ii) periodic revalidation across regimes and explicit drift monitoring; (iii) triangulation with non-platform indicators when available; and (iv) human oversight for decisions that affect livelihoods and mobility (e.g., resource allocation, risk communication). The central lesson of the two-window evidence is that "digital" does not guarantee stable informational value; robust governance requires continuously testing whether platform-derived proxies remain behaviorally aligned with the policy constructs they are assumed to measure.

While the timing of future shocks cannot be forecast as a calendar event, the operational purpose of nowcasting is to make shocks *visible* as early as the available data allow. In a monthly one-step-ahead system like ours, a new disruption becomes detectable through two complementary channels: (i) proxy-layer anomalies, where high-frequency attention signals (e.g., Wikipedia pageviews) exhibit abrupt shifts that may precede or coincide with changes in observed arrivals; and (ii) model-diagnostic anomalies, where one-step-ahead errors and residual behavior deviate persistently from their pre-shock distribution, indicating that the data-generating process or the proxy–outcome mapping has changed. In practice, the earliest operational alert can arise in the first affected month (through proxy shifts and sudden forecast-error inflation), whereas more defensible confirmation typically requires multiple consecutive monthly observations showing sustained abnormal errors or evidence consistent with a structural break/concept drift (Bai & Perron, 2003; Gama et al., 2014).

Accordingly, rather than asserting when the next shock will occur, we recommend an explicit drift/break monitoring layer in deployment. Concretely, this can be implemented by tracking rolling RMSE/MAE and standardized one-step-ahead residuals, and flagging "instability" when diagnostics exceed predefined thresholds for more than one month (e.g., repeated large standardized errors, a sustained rise in rolling error relative to recent history, or break-consistent behavior). When flagged, the system should switch to a "stress-mode" reporting stance (greater uncertainty, emphasis on robust baselines, and revalidation of proxy usefulness) until diagnostics stabilize. This operationalizes the regime-aware philosophy of the paper and clarifies when the next disruption would become visible to tourism stakeholders.

4.5. Limitations

Limitations should be made explicit because they shape interpretation and generalization:

- Spatial aggregation of climate features: using national-scale spatial averages can dilute the climate–tourism relationship, particularly for a country where tourism demand can be concentrated in climatically distinct zones and in destination-specific ecosystems. Climate–tourism relationships are well-established but are often destination- and activity-specific, so finer spatial matching (or tourism climate indices) would likely be more informative than country-level means.
- Proxy validity and behavioral drift: Wikipedia pageviews are an "attention" signal and can reflect information seeking by residents, students, journalists, or prospective

travelers; the mapping from attention to arrivals can change over time, particularly around crises. Such drift can explain why exogenous proxies appear more useful in some windows than others, consistent with recent critical assessments of tourism big-data forecasting.

- Short stress window and inference power: the stress-test window (2019–2021) contains relatively few monthly observations, which reduces the power of statistical comparison tests and makes p-values more sensitive to outliers. Therefore, statistical conclusions in the stress window should be interpreted cautiously and framed as indicative rather than definitive.

4.6. Future Research Directions

To strengthen both scientific contribution and operational value, the next steps that most directly address the above limitations are

- Destination-level nowcasting (spatial disaggregation) using province/city arrivals (if available) and spatially matched climate features, potentially incorporating tourism climate indices and extreme-event descriptors. This would align with established quantitative approaches to evaluating climate impacts on tourism and would likely increase the explanatory relevance of weather/climate features.
- Regime-switching or time-varying parameter models to explicitly represent structural breaks (COVID-like shocks, reopening phases) rather than treating them only through evaluation windows. This would directly connect to the tourism literature on post-crisis recovery forecasting.
- Probabilistic nowcasts and decision thresholds, producing prediction intervals and risk flags (e.g., "high uncertainty month") to better support policy and operational decisions, especially under disruption.

5. Conclusions

This study developed and validated a fully reproducible, open-data pipeline for monthly one-step-ahead nowcasting of Ecuador's international tourist arrivals using only a laptop-based Python 3.13.5 workflow. The results demonstrate that a regularized autoregressive specification (ridge_ar) delivers the strongest and most reliable performance relative to a mandatory seasonal naïve benchmark across the full evaluation period, and its improvement is supported by formal forecast-comparison testing.

A key contribution is the regime-aware evaluation that separates a COVID-era stress window (2019–2021) from a post-2022 post-reopening operational window, showing that model robustness under disruption and performance under later-phase monitoring conditions can differ materially. The analysis further indicates that while open exogenous signals (Wikipedia pageviews and ERA5-Land climate covariates) are theoretically motivated and technically feasible at scale, their incremental benefit is not uniform across windows; in particular, attention-based augmentation can be more informative during disruption/recovery, yet attenuate or underperform in the later post-reopening operational context, reinforcing the need for regularization, disciplined feature design, and windowed reporting in tourism monitoring applications.

Accordingly, the work offers a practical blueprint for transparent, low-cost tourism monitoring in data-constrained contexts, while emphasizing an operational standard for responsible deployment: open inputs, rolling-origin backtesting, windowed inference, and ongoing drift monitoring. Future extensions—including destination-level disaggregation, explicit regime-switching/time-varying parameter models, and probabilistic nowcasts with decision thresholds—can build on this foundation to better align open-data signals with tourism decision contexts and governance needs.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/tourhosp7040113/s1.

Author Contributions: Conceptualization, J.G. and M.M.; methodology, J.G.; software, J.G.; validation, J.G., S.F., D.B., V.C. and M.M.; formal analysis, J.G.; investigation, J.G., S.F. and D.B.; resources, S.F., D.B. and V.C.; data curation, J.G.; writing—original draft preparation, J.G.; writing—review and editing, J.G., S.F., D.B., V.C. and M.M.; visualization, J.G.; supervision, M.M.; project administration, J.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the Supplementary Material. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors acknowledge the Ministry of Tourism of Ecuador (MINTUR) for making aggregate tourism indicators available via its public dashboard; the Wikimedia Foundation for providing open access to Wikipedia pageviews data through a public API; and the Copernicus Climate Change Service/ECMWF for providing ERA5-Land reanalysis products through the Copernicus Climate Data Store. During the preparation of this manuscript, the authors used ChatGPT—5.2 (OpenAI) to support English-language editing. The authors reviewed and take full responsibility for the content of this publication.

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

Abbreviations

The following abbreviations are used in this manuscript:

AR	Autoregressive
CDS	Copernicus Climate Data Store
C3S	Copernicus Climate Change Service
DM	Diebold–Mariano (test)
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA5-Land	ECMWF fifth-generation reanalysis over land (Copernicus)
L1/L2	Lasso/Ridge regularization penalties
LOESS	Locally Estimated Scatterplot Smoothing
MAE	Mean Absolute Error
MINTUR	Ministry of Tourism of Ecuador (Ministerio de Turismo del Ecuador)
NetCDF	Network Common Data Form (climate data container)
RMSE	Root Mean Squared Error
sMAPE	Symmetric Mean Absolute Percentage Error
SARIMAX	Seasonal ARIMA with exogenous regressors
STL	Seasonal-Trend decomposition using LOESS
API	Application Programming Interface
ARIMA	Autoregressive Integrated Moving Average
PCA	Principal Component Analysis
PCR	Principal Component Regression
PLS	Partial Least Squares
ARIMAX	Autoregressive Integrated Moving Average with Exogenous Regressors

References

Afreixo, V., Tavares, A. H., Enes, V., Pinheiro, M., Rodrigues, L., & Moura, G. (2024). Stable variable selection method with shrinkage regression applied to the selection of genetic variants associated with Alzheimer's disease. *Applied Sciences*, *14*, 2572. [CrossRef]

Alis, C. M., Letchford, A., Moat, H. S., & Preis, T. (2015, June 28–July 1). *Estimating tourism statistics with Wikipedia page views* [Conference paper]. ACM Web Science Conference, Oxford, UK. [CrossRef]

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*, 1–22. [CrossRef]

Bergmeir, C., Hyndman, R. J., & Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis*, *120*, 70–83. [CrossRef]

Bokelmann, B., & Lessmann, S. (2019). Spurious patterns in Google Trends data—An analysis of the effects on tourism demand forecasting in Germany. *Tourism Management*, *75*, 1–12. [CrossRef]

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, *15*, 662–679. [CrossRef]

Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record*, *88*, 2–9. [CrossRef]

Christidis, P., & Christodoulou, A. (2020). The predictive capacity of air travel patterns during the global spread of the COVID-19 pandemic: Risk, uncertainty and randomness. *International Journal of Environmental Research and Public Health*, *17*(10), 3356. [CrossRef]

Collins-Kreiner, N., & Ram, Y. (2021). National tourism strategies during the COVID-19 pandemic. *Annals of Tourism Research*, *89*, 103076. [CrossRef]

Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media*, *20*, 336–349. [CrossRef]

Dergiades, T., Mavragani, E., & Pan, B. (2018). Google Trends and tourists' arrivals: Emerging biases and proposed corrections. *Tourism Management*, *66*, 108–120. [CrossRef]

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*, 253. [CrossRef]

Fiorucci, J. A., & Louzada, F. (2020). GROEC: Combination method via generalized rolling origin evaluation. *International Journal of Forecasting*, *36*, 105–109. [CrossRef]

Fodness, D., & Murray, B. (1997). Tourist information search. *Annals of Tourism Research*, *24*, 503–523. [CrossRef]

Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys*, *46*, 44:1–44:37. [CrossRef]

Gozgor, G., Lau, C. K., Lin, Z., & Zeng, Y. (2024). The impact of digital governance on tourism development. *Journal of Digital Economy*, *3*, 1–13. [CrossRef]

Gretzel, U., Sigala, M., Xiang, Z., & Koo, C. (2015). Smart tourism: Foundations and developments. *Electronic Markets*, *25*, 179–188. [CrossRef]

Gursoy, D., & McCleary, K. W. (2004). Travelers' prior knowledge and its impact on their information search behavior. *Annals of Tourism Research*, *31*, 929–951. [CrossRef]

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*, 357–384. [CrossRef]

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer Series in Statistics.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*, 1999–2049. [CrossRef]

Ho, J. S., Choo, W. C., Zhangyu, R., Yee, C. L., & Lau, W. T. (2022). Forecasting the high-frequency exchange rate volatility with smooth transition exponential smoothing. *Asian Academy of Management Journal of Accounting and Finance*, *18*. [CrossRef]

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*, 55–67. [CrossRef]

Hong, F., Tian, L., & Devanarayan, V. (2023). Improving the robustness of variable selection and predictive performance of regularized generalized linear models and Cox proportional hazard models. *Mathematics*, *11*, 557. [CrossRef]

Huang, J., & Zhang, C. (2024). Daily tourism demand forecasting with the iTransformer model. *Sustainability*, *16*, 10678. [CrossRef]

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, *27*. [CrossRef]

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*, 679–688. [CrossRef]

Hyndman, R. J., & Rostami-Tabar, B. (2025). Forecasting interrupted time series. *Journal of the Operational Research Society*, *76*, 790–803. [CrossRef]

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R*. Springer.

Khadivi, P., & Ramakrishnan, N. (2016, May 17–20). *Wikipedia in the tourism industry: Forecasting demand and modeling usage behavior* [Conference paper]. 10th International AAAI Conference on Web and Social Media (ICWSM), Cologne, Germany.

Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*, 1–12. [CrossRef]

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google Flu: Traps in big data analysis. *Science*, *343*, 1203–1205. [CrossRef] [PubMed]

Lee, G.-C. (2025). A data-driven approach to tourism demand forecasting: Integrating web search data into a SARIMAX model. *Data*, *10*, 73. [CrossRef]

Lei, C.-F., Chen, F., & Chu, C.-W. (2025). Optimizing SARIMAX model with big data to predict gaming tourism destination demand. *Mathematics*, *13*, 3276. [CrossRef]

Li, H., Hu, M., & Li, G. (2020). Forecasting tourism demand with multisource big data. *Annals of Tourism Research*, *83*, 102912. [CrossRef]

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, *68*, 301–323. [CrossRef]

Li, K., Liang, C., Lu, W., Li, C., Zhao, S., & Wang, B. (2020). Forecasting of short-term daily tourist flow based on seasonal clustering method and PSO-LSSVM. *ISPRS International Journal of Geo-Information*, *9*, 676. [CrossRef]

Liu, X., Liu, A., Chen, J. L., Li, G., & Song, H. (2025). Tourism demand forecasting in normal and crisis times: Combining bootstrap-aggregating and Bayesian approaches. *Journal of Hospitality & Tourism Research*. [CrossRef]

Llewellyn, M., Ross, G., & Ryan-Saha, J. (2023). COVID-era forecasting: Google Trends and window and model averaging. *Annals of Tourism Research*, *103*, 103660. [CrossRef]

Mikulić, J., & Baumgärtner, R. M. (2025). Google Trends and Baidu index data in tourism demand forecasting: A critical assessment of recent applications. *Tourism Management*, *110*, 105164. [CrossRef]

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J.-N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, *13*, 4349–4383. [CrossRef]

Nie, X., Leung, X. Y., Qian, J., Jiang, J., & Xue, L. (2024). Tourism crisis management in the digital age: Leveraging search engine data for interval forecasting. *Sustainability*, *16*, 6682. [CrossRef]

Park, E., Park, J., & Hu, M. (2021). Tourism demand forecasting with online news data mining. *Annals of Tourism Research*, *90*, 103273. [CrossRef]

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

Sowinski, J. (2021). The impact of the selection of exogenous variables in the ANFIS model on the results of the daily load forecast in the power company. *Energies*, *14*, 345. [CrossRef]

Sun, R., Yuan, H., & Liu, X. (2017). Effect of heteroscedasticity treatment in residual error models on model calibration and prediction uncertainty estimation. *Journal of Hydrology*, *554*, 680–692. [CrossRef]

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *58*, 267–288. [CrossRef]

Ting, C.-J., Hsiao, Y.-L., & Shen, J.-J. (2025). Tourist sites and visitor numbers in Taiwan: An online buzz analysis. *Bulletin of Applied Economics*, 135–147. [CrossRef]

Van Deursen, A. J. A. M., & van Dijk, J. A. G. M. (2014). The digital divide shifts to differences in usage. *New Media & Society*, *16*, 507–526. [CrossRef]

Wan, S. K., & Song, H. (2018). Forecasting turning points in tourism growth. *Annals of Tourism Research*, *72*, 156–167. [CrossRef]

Wu, D. C., Zhong, S., Wu, J., & Song, H. (2025). Tourism and hospitality forecasting with big data: A systematic review of the literature. *Journal of Hospitality & Tourism Research*, *49*, 615–634. [CrossRef]

Wu, P., Li, G., Wen, L., & Liu, H. (2024). Tourism demand forecasting with multiple mixed-frequency data: A reverse mixed-data sampling method. *Journal of Travel Research*, *63*, 2023–2041. [CrossRef]

Xiang, Z., & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management*, *31*, 179–188. [CrossRef]

Zeng, H., Tang, C., Zhou, C., & Zhou, P. (2025). Spatial analysis of network attention on tourism resources for sustainable tourism development in Western Hunan, China: A multi-source data approach. *Sustainability*, *17*, 744. [CrossRef]

Zheng, Y., & Nikolopoulos, K. (2025). Forecasting tourism recovery amid COVID-19: A comparison between bottom-up and non-overlapping temporal aggregation. In *Forecasting, planning and strategy in a turbulent era* (pp. 96–133). Edward Elgar Publishing. [CrossRef]

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*, 301–320. [CrossRef]

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power* (1st ed.). PublicAffairs.

Zuiderwijk, A., & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, *31*, 17–29. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.