Author manuscript, submitted to the AIAA Journal of Aerospace Information Systems. Preprint DOI: 10.2139/ssrn.6750148. Dataset and code: https://zenodo.org/records/18904147

Global Traffic-Normalized Baseline for ADS-B Kinematic Integrity and Ghost Records

Eugene Pik*
Mevocopter Aerospace, Vaughan, Ontario, Canada

João S. D. Garcia† and Timothy A. Smith‡
Embry-Riddle Aeronautical University, Daytona Beach, Florida 32114

The integrity of global air traffic surveillance is increasingly threatened by latent data anomalies that bypass standard monitoring filters. As aviation transitions to highly automated collision avoidance and traffic management systems, validating unencrypted Automatic Dependent Surveillance-Broadcast (ADS-B) positional data becomes critical. This study establishes a global, traffic-normalized baseline for ADS-B kinematic consistency across spatial and temporal dimensions through a retrospective analysis of 246.7 billion messages from the 2025 calendar year. To address the survivorship bias inherent in legacy data pipelines, the novel Global Envelope Thresholding framework is introduced, which enables comprehensive system observability and identifies 3,818,743 "Ghost Records": a newly defined class of severe anomalies with partially missing telemetry that traditional filters systematically discard. Validated across 77.5 million flight hours, the framework mitigates pseudoreplication caused by localized burst artifacts by defining the statistical unit as the unique aircraft impacted. Negative binomial regression reveals substantial macro-regional disparities in surveillance consistency. Scoped as a descriptive system-observability baseline, the methodology evaluates internal statistical consistency, including extreme positional jumps and sensor divergences. The resulting benchmark dataset is open-sourced with a Python interface, enabling researchers and air navigation service providers to assess infrastructure degradation and enhance probabilistic collision risk models.

*CEO, Mevocopter Aerospace; corresponding author, eugene.pik@mevocopter.com. ORCID 0000-0001-6296-919X.
†Assistant Professor, College of Aviation. ORCID 0000-0001-6802-6889.
‡Professor and Chair, Department of Math, Science, and Technology.

Nomenclature

$d_{\text{haversine}}$	=	great-circle distance between sequential coordinates, m
E_f	=	total observed flight-hours (exposure) for a given flight information region, h
$h_{\text{baro},t}$	=	broadcast barometric altitude at time t, m
$h_{\text{geo},t}$	=	broadcast geometric (GNSS) altitude at time t, m
L_t	=	spatial coordinate vector (ϕ_t, λ_t) at time t
R	=	Earth's mean radius, km
v_{derived}	=	geometrically derived speed (Teleportation), m/s
v_t	=	broadcast ground speed (Avionics Velocity) at time t, m/s
VR_t	=	broadcast vertical rate at time t, m/s
Y_f	=	observed count of unique anomalous aircraft in a flight information region
β	=	regression coefficient
Δh	=	altitude difference (Sensor Divergence), m
Δt	=	temporal update interval, s
Δv	=	speed difference (Avionics Velocity Jump), m/s
ΔVR	=	vertical rate difference (Vertical Jolt), m/s
$\Delta \psi$	=	directional shift (Directional Jitter), deg
θ	=	overdispersion parameter
λ	=	longitude coordinate, rad
μ_f	=	expected mean of the event count
ϕ	=	latitude coordinate, rad
ψ	=	heading at time t, deg

I. Introduction

A. Background and Motivation

THE modernization of global Air Traffic Management (ATM) relies on the transition from centralized, ground-based radar infrastructure to decentralized, cooperative surveillance systems [1]. Foremost among these is Automatic Dependent Surveillance-Broadcast (ADS-B), a technology mandated by airspace modernization programs such as NextGen in the United States and SESAR in Europe [2–4]. ADS-B allows aircraft to broadcast their state vectors (position, velocity, altitude, and intent) over the 1090 MHz Radio Frequency (RF) spectrum, including data obtained from Global Navigation Satellite Systems (GNSS).

While ADS-B significantly enhances airspace capacity and situational awareness, its foundational architecture lacks native encryption and message authentication. Consequently, the network is fundamentally vulnerable to both unintentional degradation, such as RF spectrum saturation, packet collisions under the ALOHA protocol [5], and multipath interference, and intentional disruptions, including GNSS spoofing and signal jamming [6–8]. As aviation infrastructure becomes increasingly dependent on this unauthenticated broadcast network, the internal consistency of these data streams becomes a critical operational requirement. Specifically, as automated systems like Airborne Collision Avoidance System X (ACAS X) ingest these continuous state updates, unmitigated data corruption poses a direct risk of triggering unwarranted Resolution Advisories.

To proactively address this vulnerability, the identification of "Ghost Records" represents a critical new safety indicator. These anomalies are formally defined as incomplete broadcast state vectors that omit standard avionics telemetry (e.g., null velocity fields) yet exhibit kinematic values—geometric positional shifts, altitude divergence, or vertical-rate jumps—that exceed standard flight envelopes. Currently, the industry lacks a globally standardized, traffic-normalized baseline to objectively measure and compare the occurrence of these phenomena across different administrative regions and temporal conditions.

As the aviation industry concurrently expands into Advanced Air Mobility (AAM) and Unmanned Aircraft Systems Traffic Management (UTM) frameworks [9], these low-altitude, highly automated vehicles will be required to integrate with this distributed approach. Consequently, mapping the natural occurrence of "Ghost Records" and sensor divergences will be critical for developing probabilistic collision risk models for autonomous operations. Because these vehicles lack traditional human pilots to visually verify conflicting traffic, ensuring the strict integrity of this digital surveillance data serves as the ultimate safety boundary against mid-air collisions.

B. Analytical Framework and Scope

Understanding the factors that contribute to the accurate characterization of ADS-B data will be a prerequisite for the safe modernization of air transport systems. The present research effort supports this objective by establishing a global characterization of ADS-B kinematic self-consistency across both space and time, based on a system-level retrospective analysis of operations in the 2025 calendar year.

More specifically, this analysis focuses on describing the frequency and location of network anomalies. The study provides a global, traffic-normalized characterization of ADS-B kinematic self-consistency, setting a practical benchmark for future studies on timing patterns and multi-variable errors.

Rather than an attempt to diagnose their underlying causes, the observed anomalies are interpreted solely as reflections of variability in surveillance consistency and coverage conditions. The repository, however, can support studies on the identification of the associated localized contributing factors (e.g., cyber-attacks, GNSS spoofing, Air Navigation Service Provider (ANSP) performance failures). Metrics such as ground speed percentiles, heading changes, and altitude differences are considered strictly localized instability indicators rather than definitive systemic anomalies.

C. Study Objectives and Organization

This study is driven by three primary objectives to establish a comprehensive baseline for ADS-B kinematic consistency. First, it seeks to quantify the spatial variability of kinematic inconsistencies across global Flight Information Regions (FIRs) after normalizing for traffic volume. Second, it aims to evaluate the temporal modulation of elevated kinematic instability across diurnal cycles and seasonal shifts at a global scale. Finally, it assesses the dimensional robustness of these observed patterns across physically interpretable kinematic indicators, including velocity, altitude, and directional (heading) anomalies. In pursuing these objectives, the study contributes a scalable big-data analytics methodology, together with an open benchmark dataset and Python API, for the verification and health monitoring of safety-critical aerospace surveillance data feeds.

The remainder of this paper is organized as follows. Section II reviews the existing literature and attempts to characterize the systemic observability gap. Section III details the data architecture, the proposed Global Envelope Thresholding (GET) framework, and the statistical modeling approach. Section IV presents the temporal, spatial, and dimension-specific results. Section V interprets these findings, discusses operational implications, and addresses methodological limitations. Finally, Section VI concludes the study and presents directions for future work.

II. Literature Review

The expansion of crowdsourced ADS-B networks has enabled extensive research into global flight operation dynamics. Open initiatives like the OpenSky Network, supported by foundational open-source decoders and processing libraries, have been particularly important for research [10, 11]. However, the majority of existing literature leveraging this data has focused on highly localized security threats or isolated operational efficiency metrics. Numerous studies have successfully proposed anomaly detection frameworks utilizing Deep Neural Networks and Machine Learning classifiers [12], Gaussian Mixture Models [13], and Kalman filters [14] to identify GNSS spoofing and trajectory deviations. While these highly localized models are effective, they frequently rely on heavily filtered, "clean" training datasets, manually bounded to specific airspace sectors (e.g., a 100 NM radius around a specific airport) over short temporal windows.

A particularly relevant application of flight data analysis is anomaly detection. Historically, such analysis relied heavily on post-flight analysis of on-board flight data recorder parameters. For instance, early advancements successfully applied one-class classifiers in a two-phase approach to mathematically quantify and rank kinematic abnormalities during aircraft descents [15]. The study demonstrated that such classifiers applied to post-flight flight data monitoring (FDM) data are highly effective at enhancing traditional event-based approaches, specifically by identifying deviations from standard operating procedures and detecting precursors to unstable approaches that fail to trigger standard alerts. However, the current literature is limited by a lack of foundational methods capable of scaling these kinematic assessments to support the real-time monitoring of system-wide ADS-B data collected by distributed networks. While advanced ML classifiers and Gaussian Mixture Models offer superior predictive capabilities for localized, tactical anomaly detection, their reliance on heavy feature engineering and pre-cleaned datasets makes them computationally prohibitive for global-scale, continuous ingestion [16]. This leaves a need for highly scalable, computationally lightweight methodologies capable of establishing a comprehensive, post-hoc global baseline, rather than focusing solely on localized predictive sensitivity.

Furthermore, by quantifying the inherent systemic variance of the surveillance architecture, the framework provides an empirical baseline of nominal network noise to support future machine learning research. Recent exploratory investigations have successfully utilized ADS-B data to correlate specific GPS anomalies, such as coordinate gaps and deviations, with localized navigational warnings and traffic density [17]. However, before advanced ML classifiers can definitively attribute an anomaly to malicious GNSS spoofing or intentional cyber-attack, algorithms must first demonstrate that the event statistically significantly deviates from the baseline systemic noise quantified in this dataset.

Consequently, a critical gap remains in the literature regarding large-scale system observability. When existing data pipelines ingest raw ADS-B messages, standard pre-filtering techniques systematically discard incomplete or malformed packets (e.g., enforcing `velocity IS NOT NULL` constraints) to ensure data reliability for modeling [18]. This introduces important *survivorship bias* into the data ecosystem as the very anomalies that potentially indicate critical degradation in information quality, equipment failure, or RF interferences are systematically excluded before analysis can occur. To the authors' knowledge, no study has yet quantified the systemic kinematic instability of positional information broadcast and collected by the global ADS-B network across longer periods to account for seasonal variability while explicitly preserving and evaluating these malformed records. This study directly addresses this gap by establishing a traffic-normalized baseline of kinematic consistency across all global FIRs while accounting for survivorship bias.

III. Data Architecture and Methodology

A. Database Architecture and Survivorship Bias

Processing the 246.7 billion messages collected from the global ADS-B network throughout the 2025 calendar year required a scalable data engineering architecture. The ingestion pipeline utilized a hybrid computational architecture. Initial data sanitation, temporal sorting, and sequential state-pairing via SQL window functions were executed directly within the OpenSky Trino `state_vectors_data4` repository. These paired vectors are subsequently extracted to a local Python environment, where optimized vectorized libraries perform the trigonometric and kinematic calculations. To ensure the integrity of the downstream statistical analysis, the initial database query implements several input sanitation protocols. Records are immediately excluded if they contain null latitude or longitude coordinates, if either geometric or barometric altitude fields are missing, or if the aircraft's unique hexadecimal identifier does not strictly meet the required 6-character length parameter. Furthermore, the ingestion logic excludes ground-based transmissions by enforcing `onground = false`.

It is critical to note that the extracted data originates from OpenSky's decoded state vector tables rather than raw hexadecimal frames. Consequently, the dataset inherits OpenSky's foundational preprocessing logic, including Compact Position Reporting (CPR) decoding algorithms and timestamp binning [19]. While the GET framework explicitly circumvents downstream database filters (such as enforcing non-null telemetry) to recover anomalous records, the baseline observability remains fundamentally bounded by the ingestion and aggregation infrastructure of the source network.

The data processing architecture implemented for the current analysis addressed the survivorship bias by adopting a sequential logic for the data pipeline described in Fig. 1 as the *Global Envelope Thresholding (GET)* framework. This visualizes the contrast between the legacy "Survivorship Bias Trap" (which can incorrectly discard anomalous pre-flight data) and the "Wide-Net Optimization" architecture utilized in this study to ensure comprehensive system observability. More specifically, the architecture explicitly distinguishes between the initial SQL ingestion and pairing layer (executed within the OpenSky Trino cluster) and the subsequent Python vectorized kinematic evaluation layer.

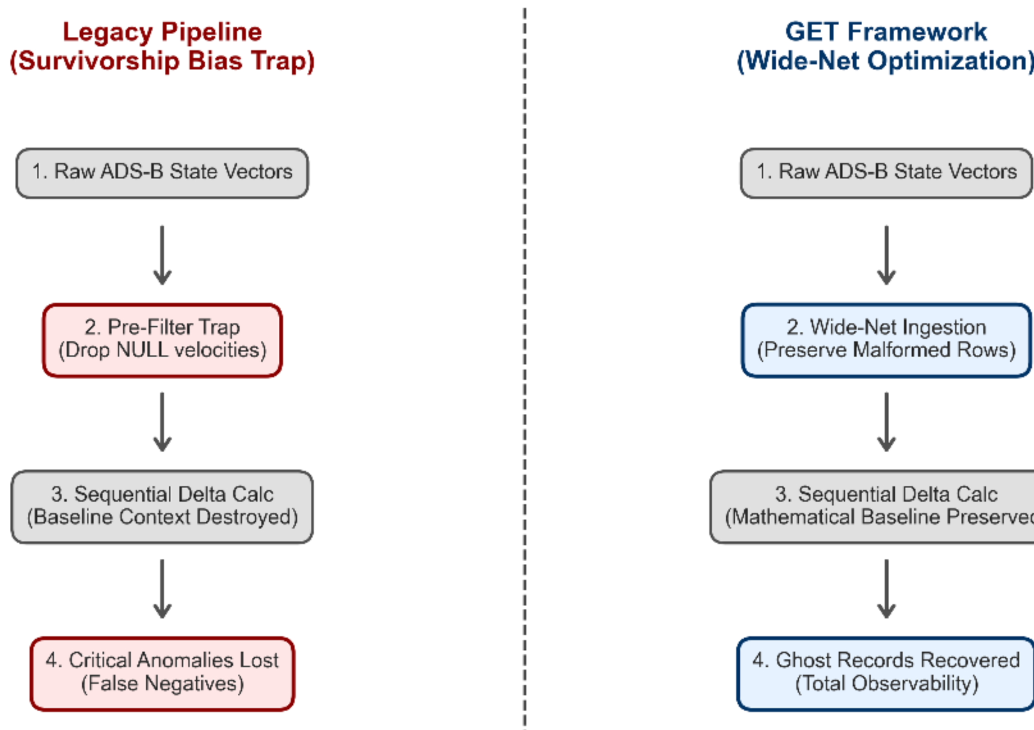

Fig. 1 The Global Envelope Thresholding (GET) Framework.

In prior iterations of extraction architectures, a critical structural flaw existed in how databases handled normal versus anomalous flight states. Standard SQL filters discard the "previous" normal state when evaluating an anomalous "current" state, thereby severing the causal link required for kinematic validation. Legacy pipelines applied restrictive base-level filtering, utilizing logic designed to immediately discard rows that did not meet standard flight profile criteria [18–20]. Because ADS-B kinematic consistency is determined by evaluating the mathematical delta between a current state and a previous state, deleting normal flight data removed the baseline context necessary for sequential mathematical evaluation. If an aircraft transitioned from a stable state of 400 knots to an anomalous state of 800 knots, the pre-filter logic deleted the 400-knot row before the SQL window function (`LAG()`) could execute. Consequently, the `LAG()` function was *logically compelled* to skip the deleted anomalous rows and pull data from the most recent "normal" record, often creating a massive, incorrect time gap of minutes or hours. This caused the algorithm to fail on anomalies emerging from steady flight states, artificially suppressing reported instability rates to 9%.

To resolve this critical flaw, the GET computational framework implemented a revised ingestion protocol which removed the restrictive pre-filters, forcing a more comprehensive ingestion of airborne traffic data. By doing so, a window function successfully calculates the kinematic delta for every single second of recorded flight sequentially before any filters are applied. When applying the primary GET thresholds, the corrected methodology confirms that approximately 78% to 80% of unique aircraft will experience at least one moment of elevated noise, sensor discrepancy, or dynamic maneuvering during their operational cycle.

Executing these calculations across hundreds of millions of state vectors per day represents a significant computational bottleneck. To achieve the necessary throughput, earlier iterations relying on standard row-by-row application functions (e.g., `df.apply()`) were replaced by vectorized approaches utilizing Pandas boolean masking and NumPy conditional statements. This transition yielded a 100-fold computational speedup, allowing the extraction to proceed efficiently. The final aggregated dataset is further optimized to 100 GB in Parquet format by employing precision rounding, serialized data type casting (Float32 for spatial coordinates and Decimal[18] for kinematics), categorical dictionary encoding for high-cardinality string identifiers (e.g., `icao24`, `callsign`), and max-level Zstandard compression.

B. Ghost Record Detection and Deferred Kinematic Computation

As mentioned previously, the GET extraction pipeline intentionally relaxes standard source filters to ensure comprehensive system observability, allowing the detection of *Ghost Records*. If a transponder broadcasts discontinuous positional coordinates while omitting telemetry, generally accepted as a signature of severe transponder degradation or non-standard configuration, the process still captures the event by computing the derived speed geometrically. In practice, a "Ghost Record" occurs when an aircraft's transponder loses connection to the air data computer (omitting standard avionics telemetry such as `velocity` or `heading`), yet the GPS module continues to broadcast 3D positional vectors. This allows the pipeline to detect severe geometric teleportation despite the missing speed fields, building upon prior methodologies used to detect fundamental coordinate gaps [21]. It is important to note that the ingestion protocol explicitly mandates the presence of both geometric and barometric altitudes to enable the continuous evaluation of vertical sensor divergence.

The pipeline explicitly distinguishes between standard positional broadcasts and true "Ghost Records" by mandating that the incomplete state vector must also trigger the excessive geometric derived speed, altitude divergence, or vertical-rate thresholds to be extracted. Consequently, the term "Ghost Record" denotes a severe *incomplete state reconstruction* at the network level rather than definitively indicating a physical "ghost aircraft" or malicious spoofing event. Crucially, aligning with the study's non-attributional mandate, the detection logic classifies the event strictly as a degradation in systemic observability, purposefully avoiding the distinction between a legitimate onboard avionics failure and a severe network reception gap (RF packet collision) that isolated the positional message.

Simultaneously, to optimize Trino cluster computational load over billions of rows, the system employs a *Deferred Kinematic Computation* technique via a Bounded Geometry Filter. Before executing complex trigonometry, the system performs a computationally lightweight, high-speed check of the change in either latitude or longitude ($|\phi_t - \phi_{t-1}| \geq 0.004^\circ \vee |\lambda_t - \lambda_{t-1}| \geq 0.004^\circ$). This bounding box represents the maximum physical distance a commercial aircraft can travel in one second. By bypassing intensive calculations for routine subsonic flights that remain strictly within this spatial envelope, the pipeline concentrates computational resources exclusively on high-velocity kinematic events. At the equator, this angular threshold equates to a physical distance of approximately 440 meters, establishing a theoretical evaluation baseline of 440 m/s (approximately Mach 1.3 at the equator) for a standard update rate of $\Delta t = 1$ second. This theoretical maximum bound serves as a conservative global ceiling, inherently becoming more sensitive at higher latitudes as the physical distance of 0.004° longitude naturally shrinks (multiplied by $\cos(\phi)$). Crucially, this filter does not discard high-speed data; rather, any state vector exceeding this geometric boundary (including artificial trans-Pacific coordinate shifts wrapping around the $\pm 180^\circ$ antimeridian) is actively routed into the Haversine evaluation pipeline, reserving maximum processing power exclusively for high-velocity kinematic events. Consequently, this architectural optimization explicitly renders the GET framework blind to low-magnitude GNSS spoofing, gradual sensor drift, and sub-Mach 1.3 trajectory deviations, reinforcing its specific mandate as a detector of severe, structural teleportation events.

C. Kinematic Error Taxonomy and Formulations

To objectively quantify surveillance consistency while explicitly distinguishing infrastructure coverage limits from avionics degradation, mathematical definitions must map to physical aviation parameters. This study utilizes the following three categories of sequential delta formulations across independent spatial dimensions, establishing a robust error taxonomy that naturally decouples these phenomena. Table 1 summarizes the five error modes, the physical failure each indicates, and the infrastructure layer to which the GET-25 results predominantly attribute each mode.

Teleportation (quantified by the derived speed metric, v_{derived}) identifies anomalous coordinate shifts mathematically induced by extended data latency, line-of-sight coverage gaps, or geometric signal disruption rather than physical aircraft

Table 1 Kinematic error taxonomy: modes, physical interpretation, and dominant attribution layer.

Error Mode (Metric)	Physically Indicates	Dominant Attribution Layer (Observed)
Teleportation (v_{derived})	Extended data latency, line-of-sight coverage gaps, geometric signal disruption	Network/coverage (oceanic and remote corridors)
Sensor Divergence (Δh)	Barometric failure or EMI-induced 3D-GNSS divergence	Avionics/environment (terrestrial interference zones)
Directional Jitter ($\Delta\psi$)	Gyroscope decoupling, sensor spoofing, high-noise degradation	Mixed (coverage gaps and localized interference)
Avionics Velocity Jump (Δv)	Inertial reference degradation or transmission corruption	Avionics/transmission
Vertical Jolt (ΔVR)	Transponder air-data faults preceding full altitude divergence	Avionics/air-data chain

acceleration. Evaluated using the Haversine formula over the temporal delta:

$$d_{\text{haversine},t} = 2R \arcsin\left(\left(\sin^2\left(\frac{\phi_t - \phi_{t-1}}{2}\right) + \cos(\phi_{t-1})\cos(\phi_t)\sin^2\left(\frac{\lambda_t - \lambda_{t-1}}{2}\right)\right)^{\frac{1}{2}}\right) \tag{1}$$

$$v_{\text{derived},t} = \frac{d_{\text{haversine}}(L_t, L_{t-1})}{\Delta t} \tag{2}$$

where R is the Earth's mean radius (6,371 km, with the resulting great-circle distance converted to meters prior to rate evaluation), $L_t = (\phi_t, \lambda_t)$ represents the spatial coordinate vector at time t (with latitude and longitude converted to radians for mathematical evaluation), and Δt represents the temporal update interval in seconds between the two states. Operationally, this mathematical formulation is executed via the highly optimized, native `great_circle_distance` spatial functions within the Trino database engine, which natively interpret decimal degrees and internally resolve spherical ±180° antimeridian wrap-around constraints [22]. This architectural reliance ensures that trans-Pacific trajectory continuity does not trigger false teleportation flags due to manual calculation errors. Values exceeding the GET P99 envelope indicate geometric teleportation regardless of the onboard avionics velocity report.

Sensor Divergence (Altitude Instability) measures severe vertical discrepancies between independent onboard altitude sensors, indicating barometric failures or external electromagnetic interference (EMI) causing 3D-GNSS divergence. This instantaneous sensor variance is calculated as:

$$\Delta h = \left|h_{\text{geo},t} - h_{\text{baro},t}\right| \tag{3}$$

where $h_{\text{geo},t}$ represents the geometric (GNSS) altitude and $h_{\text{baro},t}$ represents the barometric altitude broadcast within the exact same state vector at time t.

Directional Jitter (Directional Instability) measures anomalous orientation changes without corresponding geometric trajectory shifts, highlighting gyroscope decoupling, sensor spoofing, or high-noise data degradation. To account for the circular nature of compass headings (where 359° and 1° are mathematically 358° apart but physically only 2° apart), the directional shift $\Delta\psi$ is calculated using a modulo operator to find the shortest angular distance:

$$\Delta\psi = |((\psi_t - \psi_{t-1} + 540^\circ) \bmod 360^\circ) - 180^\circ| \tag{4}$$

where ψ_t and ψ_{t-1} represent the broadcast aircraft heading at the current and previously recorded timestamps, respectively.

Avionics Velocity Jump (Speed Difference, Δv) measures abrupt, unphysical changes in the broadcast velocity vector between sequential messages, indicating severe inertial reference degradation or transmission corruption independently of geometric coordinate shifts.

$$\Delta v = |v_t - v_{t-1}| \tag{5}$$

where v_t and v_{t-1} represent the broadcast ground speed at the current and previously recorded timestamps.

Vertical Jolt (Delta Vertical Rate) isolates extreme fluctuations in the broadcast vertical climb or descent rate, identifying transponder air-data faults prior to full altitude divergence.

$$\Delta \mathrm{VR} = |\mathrm{VR}_t - \mathrm{VR}_{t-1}| \tag{6}$$

where VR_t and VR_{t-1} represent the broadcast vertical rate at the current and previously recorded timestamps.

It should be noted that while this manuscript formally evaluates the five primary kinematic error modes detailed above, the open-source GET-25 pipeline concurrently calculates secondary geometric metrics, such as the geometric bearing change between sequential spatial coordinates. These secondary metrics are preserved in the published datasets to support future machine learning and UTM applications, but are explicitly excluded from the current baseline analysis.

D. Phases of the GET Computational Framework

To ensure the non-attribution principle and descriptive rigor, the analytical pipeline explicitly separates the “characterization” of the data from the “extraction” of events using a four-phase process described below.

1. Phase I: Descriptive Sensitivity Analysis

The first step involves executing the hourly analysis script, in which the pipeline acts strictly as a passive observer. A resilient OAuth2-authenticated connection to the OpenSky Trino database supports the extraction of successive position pairs while enforcing a strict 60-second maximum gap constraint for data continuity. The script calculates hourly distributional properties for the entire global dataset across primary kinematic dimensions.

A key computational innovation in this stage is the implementation of Trino's `approx_percentile` array logic. This allows the system to compute distributional markers (the 25th, 75th, 95th, 99th, and 99.9th percentiles) and flatten them into scalar columns. Beyond raw percentiles, the script generates derived statistical summaries to characterize the density of the outlier tails, including Coefficients of Variation (e.g., `std_speed / mean_speed`). The output, compiled into `sensitivity_analysis_V2.5-2025_hourly.csv`, consists entirely of distributional summaries without performing discrete anomaly labeling.

2. Phase II: Global Threshold Consolidation

Initial expectations were that relying on a single hour's threshold would be overly sensitive to local noise, while utilizing a simple mean would be too permissive. To resolve this, Phase II systematically derives the *Maximum Observed Hourly Percentile* across the entire study period. This approach established a robust, conservative upper bound, essentially defining a *Maximum Hourly Envelope Threshold*. By selecting the absolute maximum observed percentile across all 8,760 hours in the year, the pipeline caps the absolute maximum extraction rate at approximately 5% (P95 envelope) or 1% (P99 envelope) during the network's most volatile peak hours. This highly conservative global ceiling ultimately yields an annualized baseline rejection rate of approximately 2.3% across standard operations. A common sensitivity critique is that a singular, extreme "black swan" event (e.g., a massive solar flare or localized interference) could dictate the threshold for the entire year, risking Type II errors during stable periods. While anchoring the threshold to the 99th percentile of hourly distributions would theoretically increase sensitivity, dropping the top 1% of operational hours means the framework would inevitably fail—generating massive false-positive cascades—during the ~87 most congested, volatile hours of the year. Because the objective is to establish a worst-case design envelope, empirical resilience during peak saturation was prioritized over heightened sensitivity during quiet operations, yielding a strictly conservative, mathematically flat P95 "global ceiling" that remains constant throughout the year. By locking this threshold to the 95th percentile of the year's most volatile hour (Fig. 2), the framework ensures that any flagged event is a significant deviation from standard operations. The sustained peak in message volume observed from late March to May visually coincides with the highest kinematic volatility, indicating that the network becomes less stable as global flight volumes expand during the spring transition.

While metrics such as $v_{\text{derived},t}$ and Δv are inherently time-normalized (rates per second), the extraction relies on raw geometric deltas for Δh and $\Delta\psi$. Although the pipeline enforces a 60-second maximum gap ceiling, the overwhelming majority of evaluated sequential pairs operate on standard high-frequency ADS-B update rates ($\Delta t < 2$ seconds). Note that raw $\Delta\psi$ and Δh thresholds are naturally sensitive to data gaps. In oceanic regions (where gaps approach 60s), these flags represent genuine degradations in systemic observability.

The resulting empirical constants (Table 2) represent the absolute maximum hourly percentiles derived across the complete continuous sample of 8,760 operational hours, evaluating the full 246.7 billion state vectors. It is crucial

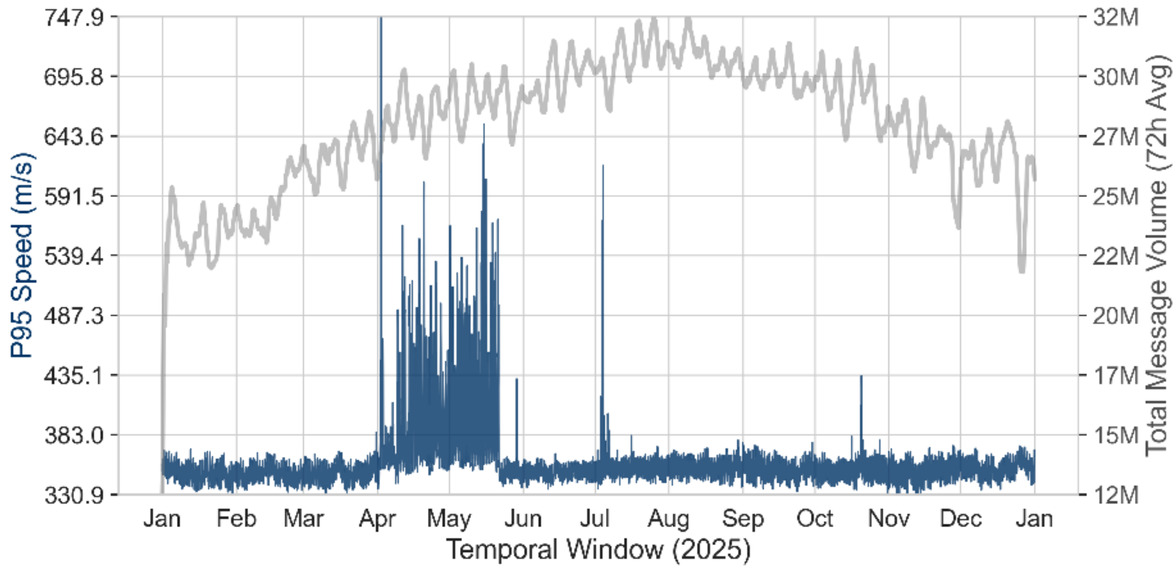

Fig. 2 Stability of the GET Maximum Hourly Envelope Threshold.

Note: The solid blue line (corresponding to the left y-axis) tracks the global hourly P95 derived speed (v_{derived}), revealing volatile peaks during the spring transition. The thick gray line (corresponding to the right y-axis) represents the 72-hour smoothed total message volume. The static GET-25 functional envelope is anchored to the absolute maximum observed volatility during these peak operational periods.

to note that thresholds such as a 170.7-degree directional shift ($\Delta\psi$) within a standard 1-to-2-second update interval are physically impossible for commercial aviation. Therefore, these empirical bounds successfully capture gross data corruption, severe sensor decoupling, and critical information gaps rather than actual aerodynamic maneuvering.

Table 2 Empirical Kinematic GET Envelope Thresholds.

Kinematic Metric	Physical Interpretation	P95 Envelope	P99 Envelope
Derived Speed (v_{derived})	Teleportation	747.9 m/s	6670.3 m/s
Altitude Difference (Δh)	Sensor Divergence	785.8 m	2323.1 m
Directional Jitter ($\Delta\psi$)	Directional Instability	14.7 degrees	170.7 degrees
Delta Speed (Δv)	Avionics Velocity Jump	14.9 m/s	84.9 m/s
Vertical Jolt (ΔVR)	Vertical Rate Instability	3.7 m/s	13.0 m/s

3. Phase III: Event Extraction

The event extraction logic works through a three-step filtering process to catch data errors. First, a simple distance check (the Bounded Geometry Filter) quickly spots any aircraft moving faster than a commercial jet's physical limit. Second, the system "relaxes" standard database rules, allowing it to catch "Ghost Records", messages that have location data but are missing other standard flight telemetry. Finally, the system compares these filtered results against the fixed GET limits (Table 2), saving only the extreme cases of jumping coordinates, altitude, heading, avionics velocity jump, or vertical rate. By saving both the current and previous locations (L_t and L_{t-1}) together, the final dataset allows researchers to instantly map the "jump vector" of every anomaly.

4. Phase IV: Spatial Enrichment

Following extraction, the daily Parquet files are enriched through a highly optimized Polars and GeoPandas architecture. To guarantee spatial accuracy and eliminate the ~4% error rate typically caused by simplified or low-resolution open-source boundaries, the script ingests a high-fidelity, high-vertex-count local map (`icao_firs_global.geojson`). Accuracy is secured through a two-stage spatial join. First, a strict "Point-in-Polygon" check determines if a coordinate falls within a defined FIR boundary. Because standard maps often have small gaps along coastlines or international borders, a second "Coastal Snapping" stage uses a nearest-neighbor algorithm. This ensures that any point falling just outside a boundary is snapped to the closest administrative region within a 0.5-degree (~30 NM) radius rather than being discarded as an orphan record. This dual-layer approach ensures that every extracted kinematic anomaly is correctly attributed to the responsible Air Navigation Service Provider (ANSP).

E. Formal Statistical Modeling Framework

To formally answer the primary research question regarding spatial variability, a statistical prediction model is required, in which the absolute counts of extracted events are mathematically transformed into probabilities. This requires a formal statistical modeling framework capable of handling highly dispersed count data across distinct geographic territories characterized by highly disparate traffic densities.

The Poisson regression model is commonly used in similar applications, but it requires that the mean of the response variable is approximately the same as its variance [23]. The distribution of kinematic inconsistency events in the global ADS-B network data exhibits variance that heavily exceeds its mean. A Negative Binomial regression is a commonly used model to resolve this overdispersion and is what we adopt instead. As a result, the distributional assumption for the problem is defined as $Y_f \sim \text{NegBin}(\mu_f, \theta)$, where μ_f denotes the expected value (mean) of the event count Y_f for a given FIR. In order to mitigate statistical pseudoreplication caused by localized "burst artifacts" (instances where a single malfunctioning transponder broadcasts thousands of sequential anomaly flags), the dependent variable Y_f is defined as the count of *unique aircraft* experiencing an extreme kinematic event within a specific Flight Information Region (f) over the 2025 calendar year, rather than raw message volume. This shift in the unit of analysis from messages to aircraft prevents a single malfunctioning transponder from skewing regional safety metrics. To explicitly demonstrate how the model normalizes for vastly different traffic volumes, the Log-Linear Link function incorporates an exposure offset, defined as:

$$\log(\mu_f) = \log(E_f) + \beta_0 + \beta_1 \text{Region}_f, \tag{7}$$

where the offset $\log(E_f)$ represents the natural logarithm of total observed flight-hours within FIR f. This mathematical inclusion is critical, as it formally transforms the model from predicting raw anomaly counts into predicting the traffic-normalized incidence rate. To calculate this continuous exposure from a discrete message stream, E_f is derived by aggregating the sequential time deltas (Δt) between positional updates, strictly bounded by the 60-second maximum gap constraint to halt accumulation during oceanic unobservability. For the 2025 study period, the total global exposure ($\sum E_f$) was calculated at *77,502,847 flight hours*. This exposure metric serves as the critical offset ($\log(E_f)$), ensuring that reported spatial disparities are independent of traffic density and preventing high-density airspaces like European central corridors from appearing artificially unstable compared to remote regions.

The decision to aggregate FIRs into seven predefined Macro Regions, directly mapping to standard International Civil Aviation Organization (ICAO) administrative and statistical groupings [24], was explicitly chosen to establish a high-level global baseline of structural disparities while preventing geographical selection bias. While intra-regional variance is substantial (e.g., dense Western European sectors versus volatile Eastern European conflict zones), modeling individual FIR characteristics such as receiver density or routing complexity falls outside the descriptive mandate of this study. The Macro Region predictor serves to quantify broad geographic risk, setting the stage for future microscopic models to nest specific FIRs within these regional baselines.

Statistical modeling was performed utilizing the `statsmodels` library in Python, with data engineering executed via Trino, Polars, and GeoPandas. Goodness-of-fit testing revealed a Pearson dispersion statistic of $\chi^2/df = 2.71$, indicating significant overdispersion (variance approximately 2.7 times the mean) and justifying the Negative Binomial specification over Poisson regression. The substantial overdispersion reflects the inherent heterogeneity of crowdsourced ADS-B surveillance, wherein variance sources include terrestrial receiver density gradients, ALOHA protocol collision probabilities under varying traffic load, line-of-sight coverage gaps in oceanic corridors, and differential GNSS interference. As is standard practice for modeling heavy-tailed crowdsourced count data, the NB2 formulation with a fixed overdispersion parameter of $\alpha = 1.0$ was specified. The overdispersion parameter was fixed at 1.0 following preliminary estimation attempts that failed to converge, likely attributable to the high proportion of zero-count FIRs in the dataset. Sensitivity analyses with alternative α values in the range [0.5, 2.0] yielded qualitatively identical regional rankings; the complete sensitivity outputs are archived alongside the GET-25 dataset in the Zenodo repository. This formulation explicitly relaxes the strict Poisson variance assumption by modeling the variance as a quadratic function of the mean, successfully absorbing the extreme overdispersion caused by localized burst artifacts while maintaining interpretability of the incidence rate ratios. The primary analytical output generated by this negative binomial regression is a series of Incidence Rate Ratios (IRRs). Calculated by exponentiating the regional coefficients ($\exp(\beta_1)$), an IRR functions operationally as a relative rate multiplier. Rather than representing a bounded probability, these IRRs directly quantify how many times more (or less) frequently an anomalous aircraft is expected to be observed per flight-hour in a given region compared to the European baseline.

IV. Results

A. Methodological Validation

To empirically validate the efficacy of the methodological relaxations described in Sec. III, a supplementary validation routine was executed on the extracted GET-25 dataset to explicitly isolate "Ghost Records." By applying the `validate_ghost_records.py` script (included in the open-source repository) across the full-year extraction, the verification revealed that the system captured *3,818,743 ghost records* globally. These represent unique events where the standard avionics velocity telemetry was completely null or missing upon network reception, yet the broadcast included sufficient spatial information to trigger the geometric, altitude-divergence, or vertical-rate anomaly thresholds. Compositionally, altitude-divergence triggers dominate the Ghost Record population (89.7%), followed by vertical-rate jolts (7.3%), with derived-speed (teleportation) triggers accounting for approximately 2.0%—indicating that severe telemetry loss most frequently co-occurs with vertical-channel corruption rather than horizontal positional jumps.

Table 3 Sample Structure from `adsb_2025_filtered_daily_YYYYMMDD.parquet`.

icao24	callsign	lat	lon	prev lat	prev lon	velocity	derived speed	delta heading	regime	metric triggered
151dee	SDM6573	56.59	89.87	56.60	89.86	205.76	1012.10	0.00	p95	Derived Speed
151dee	SDM6573	56.57	89.98	56.57	89.96	206.44	1643.80	0.00	p95	Derived Speed
3a46b3	VT596	-14.94	-148.29	-14.96	-148.18	138.72	12288.00	0.00	p99	Derived Speed
3dd577	PBB17	52.31	13.47	52.31	13.47	5.06	4.60	76.19	p95	Directional Jitter

An evaluation of a sample daily extract (Table 3) illustrates how the thresholds actively classify real-world state vectors, demonstrated through three representative checks. First, for the *Threshold Logic Check*, observations of aircraft `3a46b3` (callsign VT596) reveal a v_{derived} of 12,288 m/s. Because this value significantly exceeds the standard flight envelope (equivalent to approximately Mach 36), the pipeline correctly classifies it under the `p99` regime, flagging a severe position jump where the geometric distance covered vastly exceeds physical aerodynamic capabilities. Furthermore, in *High-Velocity Filtering*, aircraft `151dee` illustrates the capture of `p95` speed anomalies. With a v_{derived} of 1,643.80 m/s against a recorded ground `velocity` of 206.44 m/s, the system identifies a substantial discrepancy. Finally, in *Directional Variance*, the entry for aircraft `3dd577` demonstrates the sensitivity of the filters to non-speed metrics. Despite a low `velocity` (5.06 m/s), it triggered a `p95` flag for $\Delta\psi$, allowing the pipeline to isolate erratic directional changes or sensor "jitter." These demonstrate that actual cases were flagged using the proposed rules as designed.

B. Temporal Modulations and Volatility

An analysis of the detailed extraction statistics across the 2025 calendar year provides critical insights into the temporal modulation of both network volume and kinematic instability. The dataset demonstrates that the absolute scale of the ADS-B network expands and contracts dynamically across both diurnal cycles and major seasonal shifts. The following sections present these results in added detail.

1. Daily Load and Systemic Elasticity

During the initial January and February observation period, the global network exhibited a standard diurnal curve predominantly driven by North American and European operations. The temporal analysis curves (Fig. 3) indicate the modulation of kinematic volatility across the 24-hour UTC cycle, with values representing the 95th percentile of Derived Speed (v_{derived}) as the primary tracking metric. Because infrastructure bottlenecks and packet collisions are more likely to be captured in expanding temporal data blocks, v_{derived} serves as the most sensitive and direct proxy for systemic network latency under load. Conversely, metrics such as altitude and directional divergence are primarily driven by localized, external electromagnetic interference rather than global traffic volume. Because this global aggregation references common timestamps in UTC, the observed oscillations may more directly reflect the aggregated behavior of the world's densest airspaces. For instance, the absolute global minimum for operational traffic occurred on January 19 at 05:00 UTC, processing just 12,533,825 messages—a window that aligns with the deep pre-dawn downtime across Europe (where crowdsourced receiver density is exceptionally high) coinciding with the post-midnight lull of North American East Coast operations. Conversely, peak hours during these baseline early-year months typically see volumes exceed 28 million messages as they approach 15:00 UTC, the critical daily window where the European afternoon peak directly overlaps with the North American morning expansion. Despite these substantial intra-day increases in volume, the overall proportional rate of kinematic instability remains highly consistent within the winter operational context, demonstrating that raw anomaly counts scale linearly with daily traffic patterns.

To assess the potential effects of load tolerance and performance degradation of the system under load, a regression analysis was conducted between hourly global traffic volume and the aggregate anomaly filter rate across the full year (Fig. 4). The data demonstrates that general degradation, as measured by the *Anomaly Filter Rate*, remains remarkably flat and consistent across all volume profiles, though the variance of that hourly rate widens at absolute capacity limits. After removal of 21 exactly duplicated ingestion rows (9 January, 00:00–20:00 UTC), the resulting correlation across all $N = 8{,}760$ hourly observations ($r = 0.0372$) indicates virtually zero correlation ($R^2 = 0.0014$). While the massive sample size mathematically generates a significant p-value, the operational effect size is completely negligible. This finding provides no evidence of a "saturation zone" and indicates that the network's baseline rejection rate is highly resilient, remaining independent of network volume and anchored near 2.3% regardless of network load. While the mean filter rate remains stable, the variance (the vertical spread of outlier hours in Fig. 4) widens slightly as volume exceeds 35 million messages. Rather than indicating architectural server limits or centralized throughput bottlenecks, this slight widening suggests that extreme capacity peaks simply introduce a higher absolute number of aircraft subject to natural stochastic noise, localized RF packet collisions, and high-density terminal maneuvering.

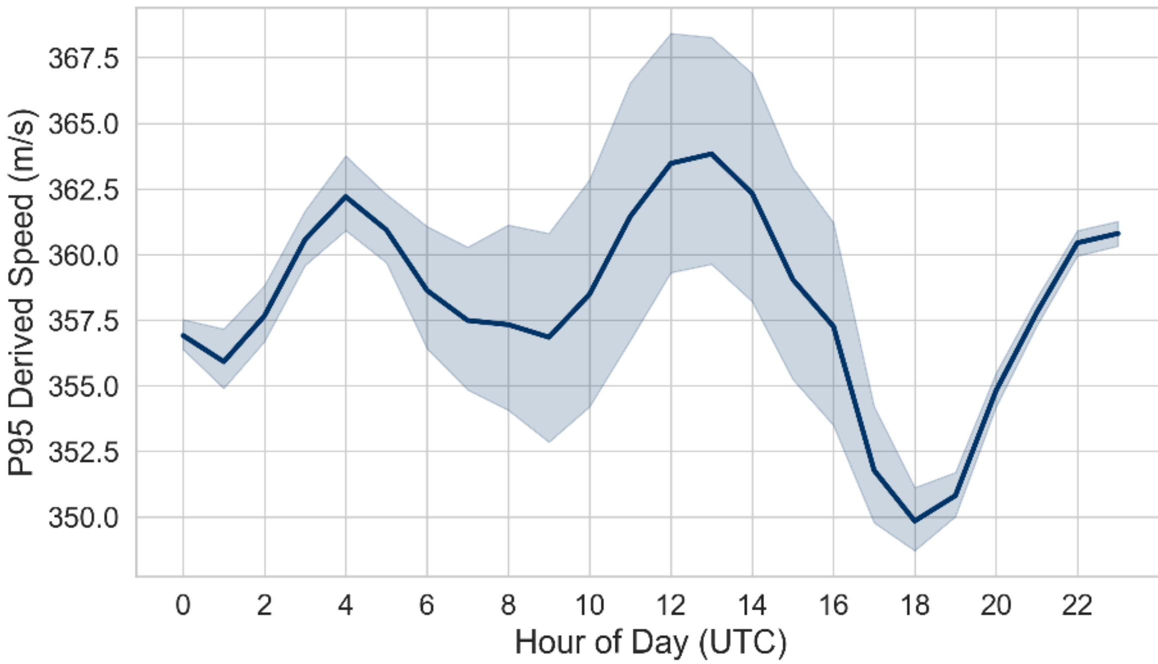

Fig. 3 Temporal Analysis of Diurnal Fluctuations.

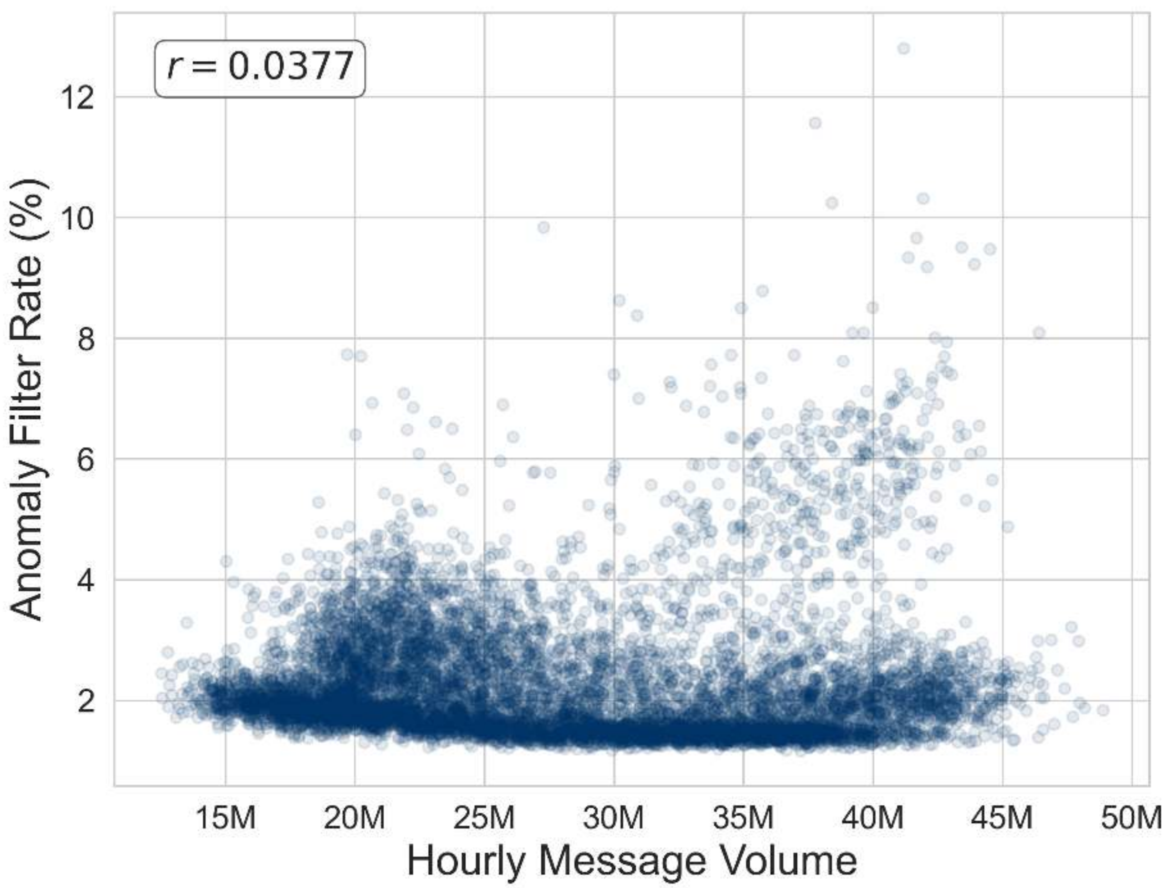

Fig. 4 Systemic Elasticity and Bandwidth Saturation.

To further decompose the systemic variance identified in the global aggregate (Fig. 4), the regression was stratified by macro-region (Figs. 5 and 6). As explicitly visualized in the accompanying two-panel global volume and regional anomaly distributions (Fig. 7), message volumes in the African (AFI) macro-region are concentrated at the lower bounds, whereas North America (NAM) and Europe (EUR) routinely process tens of millions of messages concurrently.

In mature, high-density airspaces such as NAM, EUR, ASIA, and CARSAM (Fig. 5), the relationship between global volume and the regional anomaly count is highly predictable. While axis auto-scaling visually exaggerates the plotted trajectories, the mathematical slopes remain exceptionally low (Table 4). For example, North America (NAM) exhibits a strong positive correlation ($r = 0.864$), but its Ordinary Least Squares (OLS) linear regression slope is just 0.000328. In operational terms, for every massive influx of 10 million additional global messages, the NAM region only experiences an increase of approximately 3,280 anomalous aircraft. This quantitative ceiling confirms that the absence of any severe non-linear acceleration in the global aggregate (Fig. 4) holds true at the regional level; massive expansions in global data do not trigger non-linear packet loss or centralized saturation. Furthermore, the strong negative correlation observed in the ASIA macro-region ($r = -0.829$) perfectly reflects diurnal timezone shifts; as global volume peaks during the Euro-American operational day, Asian regional traffic concurrently reaches its nocturnal minimum.

Conversely, developing and oceanic regions like AFI, MID, and the NAT corridor (Fig. 6) exhibit weak correlation to global volume and high baseline variance. The Middle East (MID), for instance, shows a highly dispersed, negligible correlation to global load ($r = 0.218$, $R^2 = 0.047$). Together with the volume and anomaly distributions in Fig. 7, this reinforces that instability in these regions is driven by severe terrestrial receiver sparsity, line-of-sight gaps, and localized interference rather than volume-induced capacity limits.

Table 4 Regional Systemic Elasticity: Correlation between Global Message Volume and Regional Anomaly Counts.

Macro Region	**Slope (β)**	**Pearson r**	R^2	**P-value**
NAM (North America)	0.000328	0.8640	0.7465	< 0.001
CARSAM (Caribbean/S.A.)	0.000102	0.8966	0.8038	< 0.001
EUR (Europe)	0.000078	0.4005	0.1604	< 0.001
NAT (North Atlantic)	0.000001	0.4271	0.1824	< 0.001
AFI (Africa)	0.000006	0.4368	0.1908	< 0.001
MID (Middle East)	0.000002	0.2180	0.0475	< 0.001
ASIA (Asia)	−0.000063	−0.8294	0.6878	< 0.001

Note: The OLS linear regression slope (β) represents the absolute increase in regional anomalous aircraft per single unit increase in global message volume.

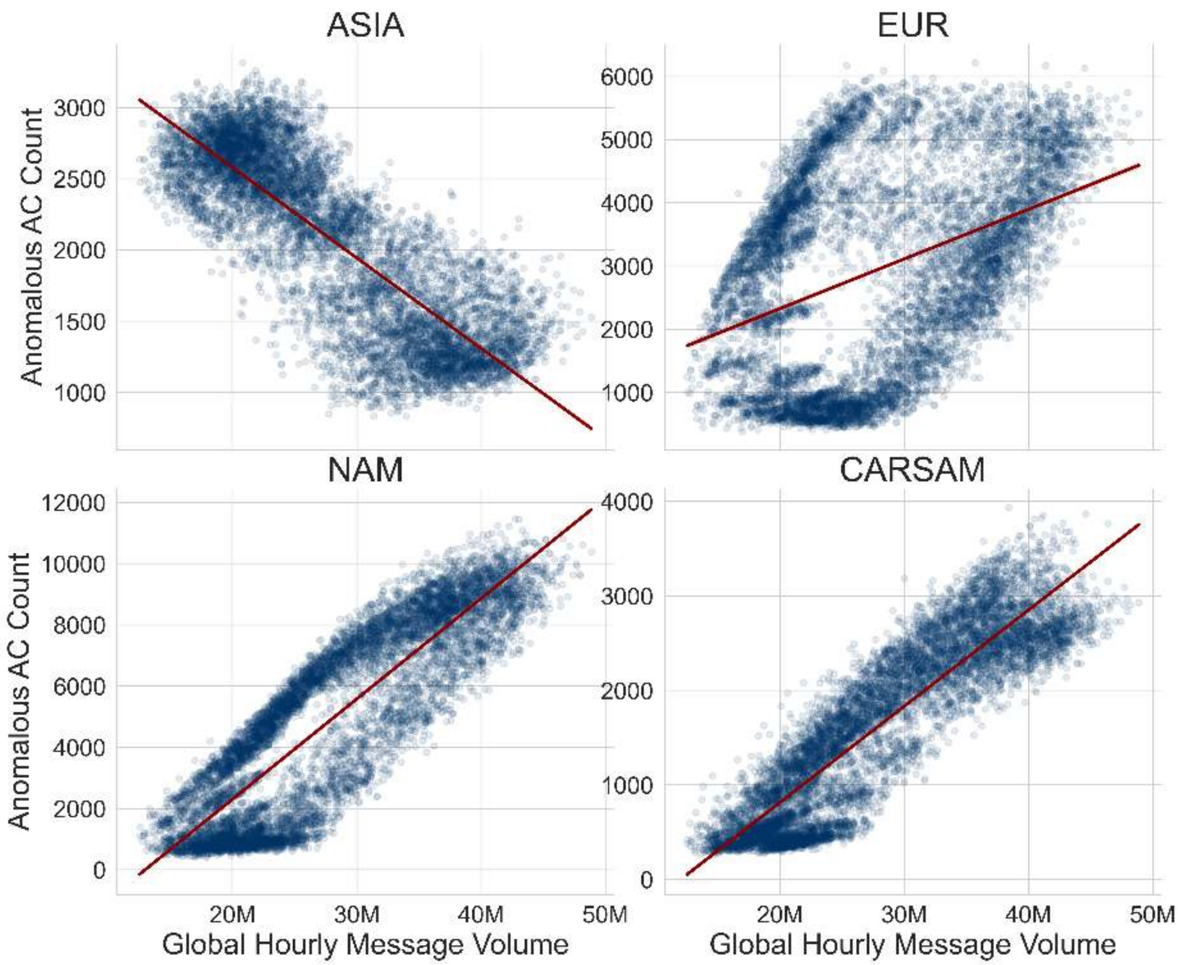

Fig. 5 Stratified Systemic Elasticity by Macro-Region (Mature Markets).

Note: These panels display regions with high baseline traffic volumes and dense terrestrial receiver networks.

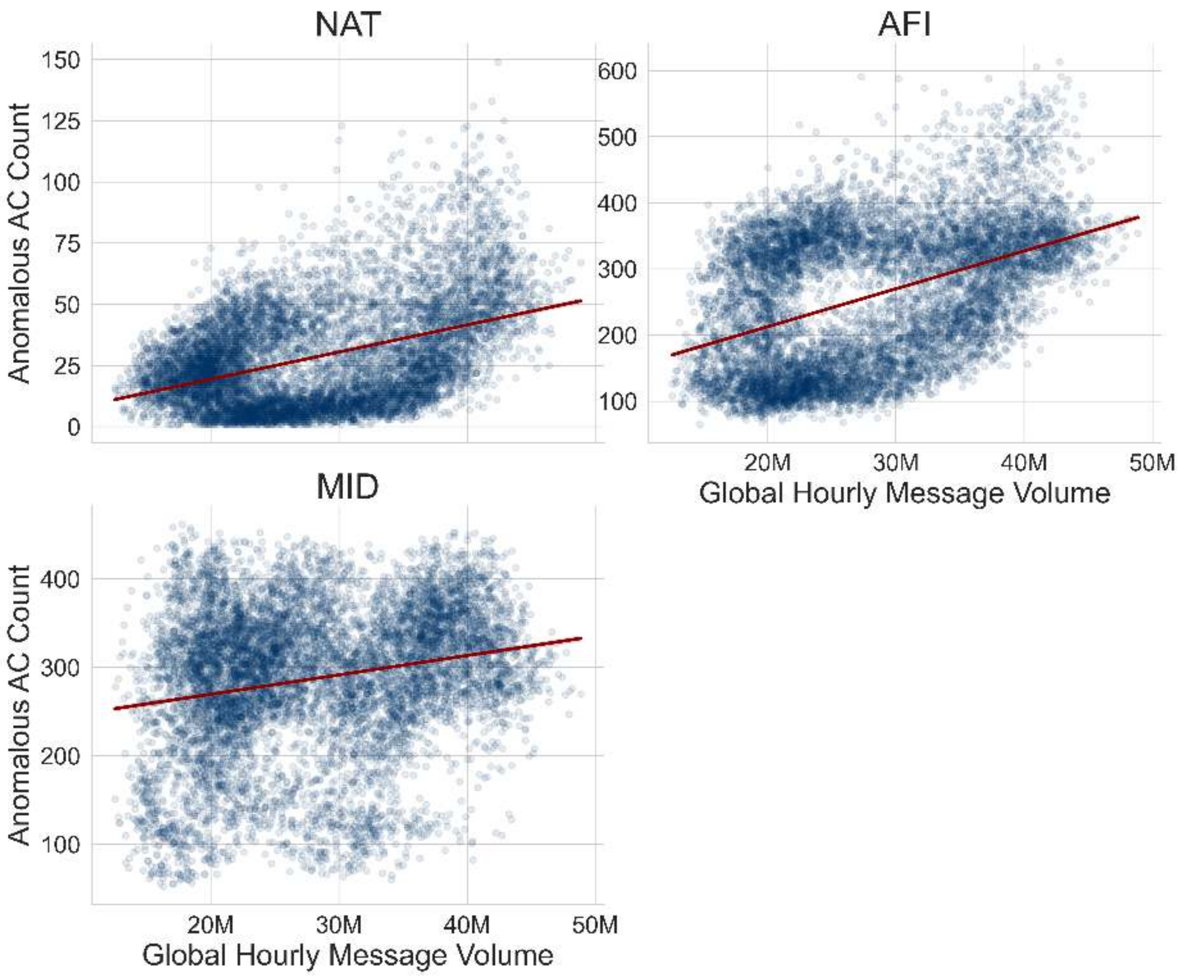

Fig. 6 Stratified Systemic Elasticity by Macro-Region (Developing and Oceanic Regions).

Note: These panels demonstrate regions characterized by lower relative message volumes and distinct line-of-sight coverage gaps.

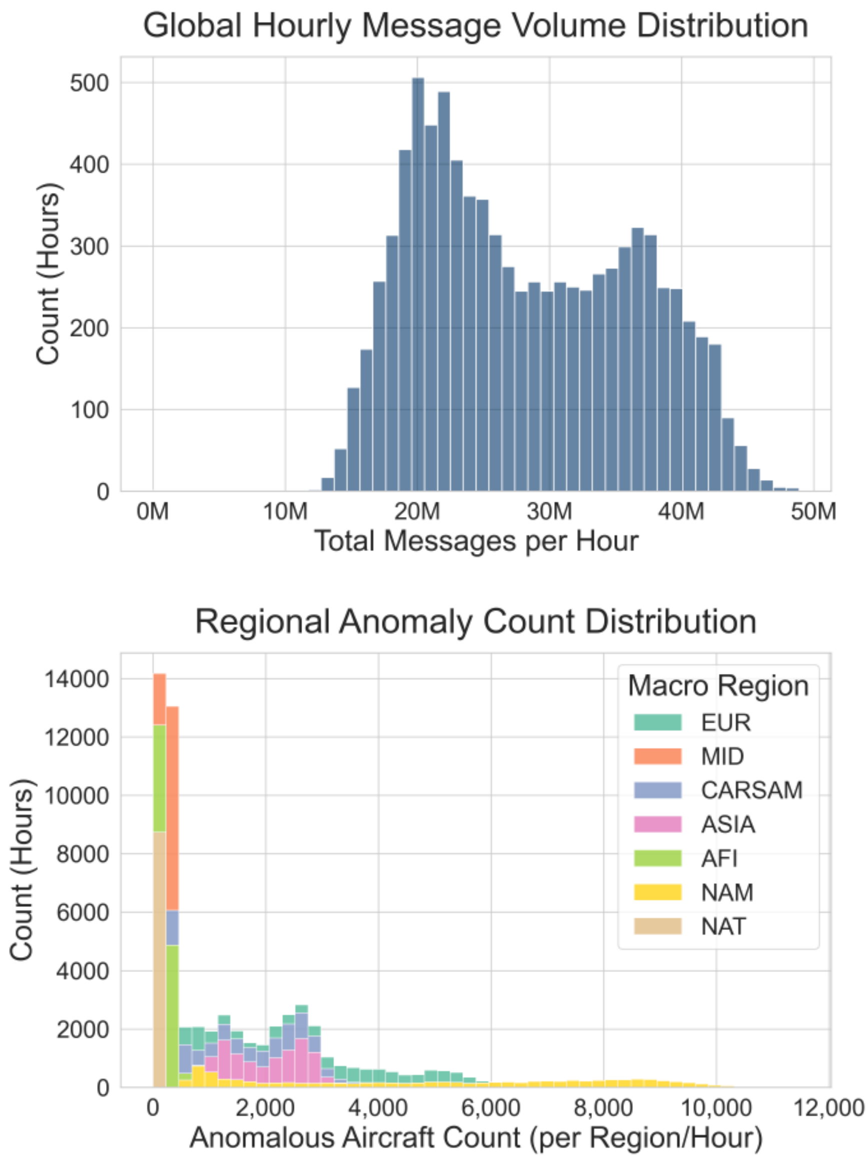

Fig. 7 Distribution of Global Message Volume and Regional Anomaly Counts.

Note: The upper panel displays the distribution of total global message volume. The lower panel displays the stacked distribution of unique anomalous aircraft counts across the seven primary macro-regions.

2. Macro-Seasonal Shifts and Intersections

Critical monthly shifts in traffic volume drive the methodological requirement for traffic normalization in spatial analysis. Monthly processing volumes ranged from 16.70 billion messages in February to a peak of 23.05 billion messages in July, representing a substantial 38.0% intra-year variation.

As the calendar advances from the relative stability of January and February into the northern hemisphere's spring season, the global network experiences a substantial expansion in both raw data volume and the severity of kinematic disruptions. The monthly boxplots (Fig. 8) demonstrate the stable baseline distributions of January/February giving way to more severe variance during April and May. The escalation intensified through early April: on April 2, the temporal analysis identified the absolute peak volatility hour of the 8,760-hour observation period, with the filtered share of messages reaching a year-high 12.80% at 20:00 UTC (Table 5). Elevated instability persisted on April 7, when the 15:00 UTC filter rate stood at 3.93% against the 1.53% mid-March baseline, and into early May: at 15:00 UTC on May 1, 46,395,858 analyzed messages yielded a sustained filter rate of 8.09%, reflecting the simultaneous triggering of multiple independent kinematic thresholds (speed, altitude, and heading).

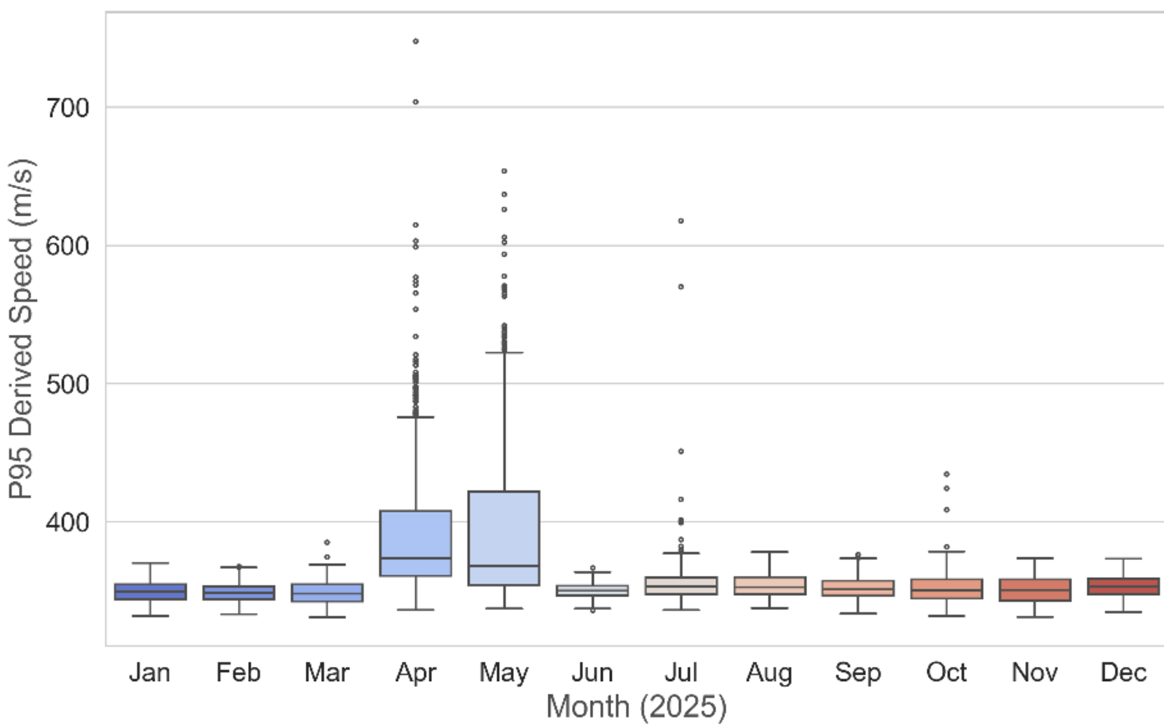

Fig. 8 Macro-Temporal Shifts and Seasonal Volatility.

Figure 9 clarifies temporal variation in P95 derived speed by combining the diurnal and seasonal findings into a two-dimensional matrix, revealing intersectional vulnerabilities within the surveillance architecture. The heatmap confirms that the absolute highest risk of kinematic instability occurs specifically during the 10:00–14:00 (UTC) periods within April and May. Because this visualization represents the P95 Derived Speed (v_{derived}) rather than raw anomaly

Table 5 Spring Volatility Extraction Statistics.

Date (Peak Hour)	Total Messages	% Filtered	Unique AC
March 13, 15:00 UTC	39,181,689	1.53%	16,270
April 2, 20:00 UTC	41,183,733	12.80%	16,938
April 7, 15:00 UTC	38,240,203	3.93%	16,227
May 1, 15:00 UTC	46,395,858	8.09%	19,103

counts, these high-risk clusters highlight true systemic latency and geometric teleportation rather than mere artifacts of increased flight volume. It visualizes a compounding degradation effect where daily traffic peaks intersect with specific seasonal network vulnerabilities.

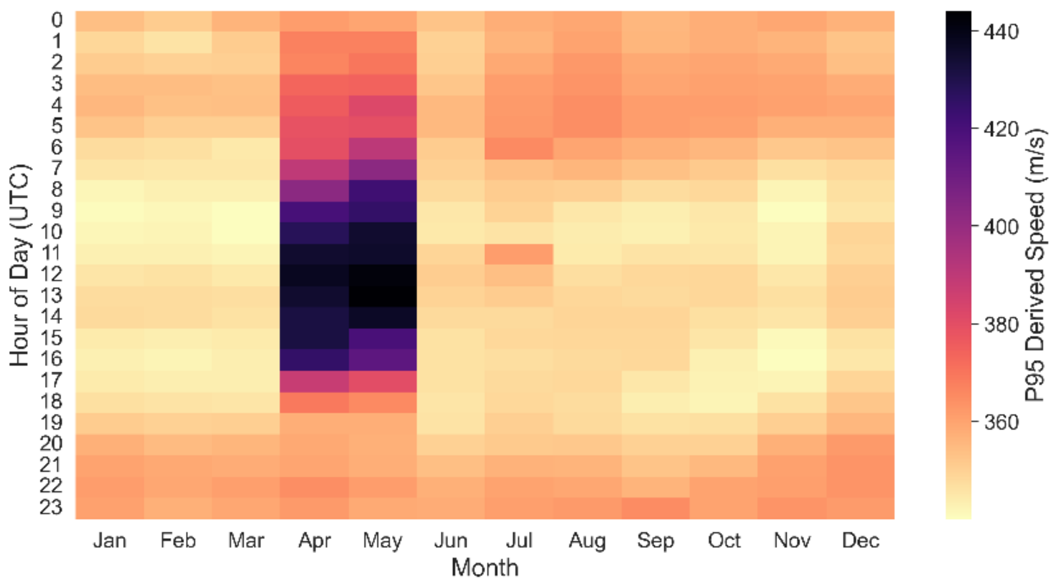

Fig. 9 2D Spatiotemporal Heatmap.

Moving into the peak mid-year travel months, the network sustains extreme volume but generally experiences a stabilization in the percentage of anomalous messages. Within the evaluated GET-25 dataset, the network reached its absolute global maximum processing volume for the year on July 18, 2025, at 15:00 UTC, ingesting *48,871,605 messages* in a single hour. Throughout July and August, the total hourly volumes frequently exceeded 40 million messages, yet the filter rate largely remained controlled. The fact that July volumes routinely exceed March-to-May transitional volumes yet exhibit lower volatility confirms that systemic instability is not merely a function of absolute message count, but rather reflects the network's vulnerability to major seasonal operational transitions. As the year shifts into its final quarter, the network undergoes a predictable contraction, returning toward the highly stable baseline parameters observed in January.

C. Spatial Disparities and Systemic Stability

This section addresses the initiative to quantify the spatial disparities in position information reported through ADS-B data and presents results at the FIR level.

1. FIR-Level Traffic-Normalized Rates

The traffic-normalized summary statistics (Table 6) provide practical support to the proposed methodological framework. To mitigate adverse sampling effects, the tabulation is restricted to FIRs processing a minimum of 10,000 annual flight hours. It is also important to distinguish the unit of analysis presented in this descriptive summary from the regression model discussed in Sec. III. The metric in Table 6 denotes the traffic-normalized rate of unique aircraft experiencing extreme kinematic anomalies per 1,000 flight hours. Rates exceeding 2,000 unique aircraft per 1,000 flight hours (e.g., the 2,521.60 rate observed in KZHU) highlight airspaces where a significant proportion of transit involves at least one severe anomaly flag. By defining the statistical unit as the unique aircraft impacted, this metric effectively mitigates pseudoreplication caused by multi-trigger events or rapid sequential burst artifacts. The pipeline independently identified regions widely documented as kinematically volatile airspaces in 2025 (e.g., Tripoli, Lahore, Baku, and Amman), geographically aligning with reported interference patterns, further supporting the sensitivity approach used in the GET baseline [25]. This also adds to literature assessing the relationship between ADS-B/GNSS interferences and conflict [26]. Figure 10 presents these rates on a logarithmic scale, highlighting structural risk density in dark red across Eastern Europe, Africa, and oceanic corridors. Regional inset maps (Fig. 11) provide additional detail of the volatility clustering over Eastern Europe and the Middle East.

Conversely, while the North American (NAM) macro-region demonstrates the highest overall systemic consistency across its aggregate airspace (as modeled in the subsequent section), absolute zero counts (0.00 rates) of extreme threshold breaches were observed specifically in highly redundant European sectors (e.g., LFRR, LIMM, and LFEE). This establishes these individual FIRs as empirical *High-Integrity Reference Airspaces.* It must be explicitly clarified that these zero counts do not imply a physical absence of minor sub-threshold noise, standard atmospheric interference, or micro-latency; rather, they confirm that these regions did not record a single instance of the severe kinematic anomalies (e.g., geometric teleportation exceeding 747.9 m/s) that persistently trigger the GET filters elsewhere. Because the static GET envelope is dominated by peak global volatility, it isolates severe structural failures while remaining insensitive to subtle degradation within already-stable noise floors (a trade-off examined in the Limitations). Consequently, these resilient regions serve as critical benchmarks for infrastructure capability, proving that highly optimized, dense ground receiver architecture successfully absorbs and mitigates extreme kinematic volatility before it manifests as a macro-level tracking failure.

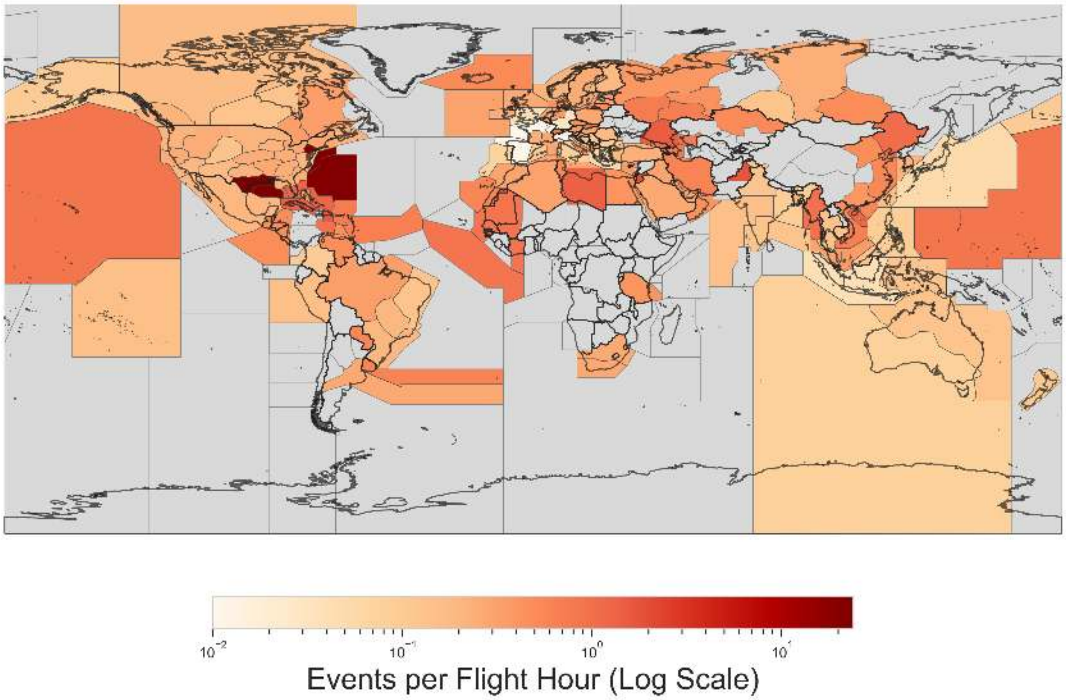

Fig. 10 Global Spatial Distribution of Traffic-Normalized Kinematic Instability.

Note: To provide comprehensive geographic coverage, this spatial projection includes all observed FIRs without applying the 10,000 flight-hour minimum utilized in the tabular summaries. Due to the logarithmic color scale, High-Integrity Reference Airspaces with absolute zero anomaly rates are represented by the lightest heatmap shade, while unobserved regions completely lacking data are explicitly mapped in solid gray.

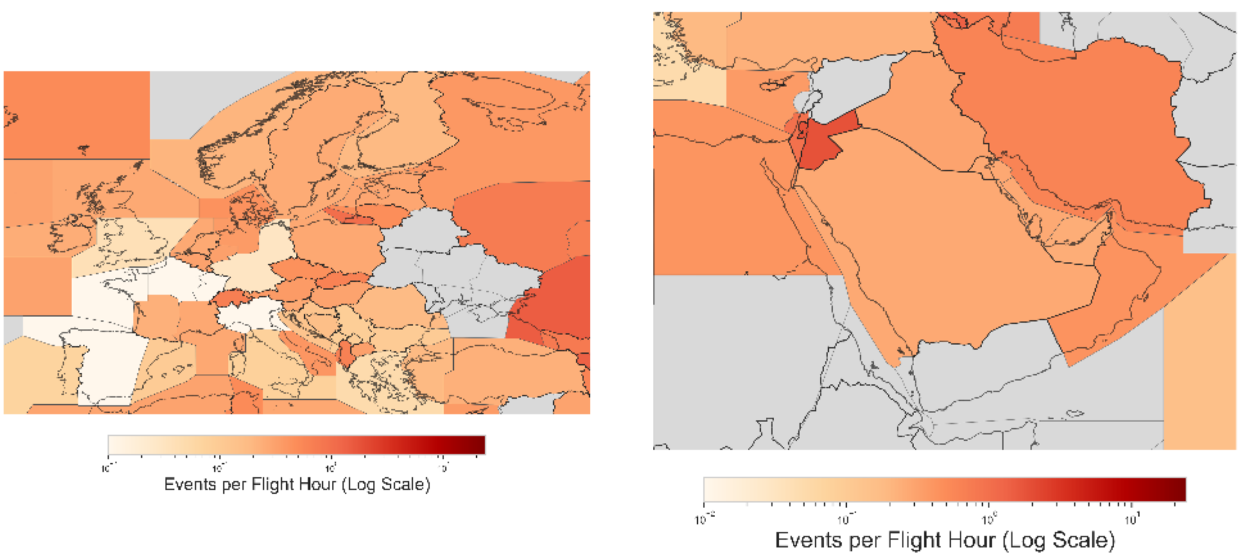

Fig. 11 Regional Inset Maps.

Table 6 Summary Statistics of Traffic-Normalized Kinematic Inconsistency.

Top 10 Most Volatile FIRs (Highest Anomaly Rates)				
FIR	**Macro Region**	**Unique AC Count**	**Flight Hours**	**Rate /1k Hrs**
KZHU (Houston Oceanic)*	CARSAM	48,775	19,343	2,521.60
KZWY (New York Oceanic)*	CARSAM	42,294	19,474	2,171.83
MUFH (Havana)	CARSAM	10,362	21,692	477.69
HLLL (Tripoli)	MID	5,267	12,247	430.06
OPLR (Lahore)	ASIA	3,970	10,403	381.61
UDDD (Baku)	EUR	3,977	10,637	373.88
MDCS (Santo Domingo)	CARSAM	8,071	28,853	279.73
ZYSH (Shenyang)	ASIA	3,323	14,312	232.18
OJAC (Amman)	MID	6,712	28,999	231.45
ZJSA (Sanya)	ASIA	4,757	20,590	231.04

Top 10 High-Integrity Reference Airspaces (Lowest Anomaly Rates)				
FIR	**Macro Region**	**Unique AC Count**	**Flight Hours**	**Rate /1k Hrs**
LFRR (Brest)	EUR	0	474,532	0.00
LIMM (Milan)	EUR	0	729,493	0.00
LFEE (Reims)	EUR	0	191,433	0.00
LFFF (Paris)	EUR	22,629	3,618,318	6.25
RJJJ (Fukuoka)	ASIA	9,423	1,389,551	6.78
LECM (Madrid)	EUR	14,683	2,112,649	6.95
EGTT (London)	EUR	23,793	3,118,993	7.63
WIIZ (Jakarta)	ASIA	5,631	574,920	9.79
LIRR (Rome)	EUR	20,722	2,090,471	9.91
EDUU (Rhein UIR)	EUR	14,090	1,412,610	9.97

*High rates in oceanic regions (KZHU, KZWY) are predominantly driven by line-of-sight infrastructure latency and spatial coverage gaps rather than physical kinematic instability (see Sec. IV.D.1). Note: KZWY (New York Oceanic) is formally mapped to the CARSAM macro-region in accordance with ICAO GIS administrative boundaries to prevent subjective geographic selection bias.

2. Macro-Regional Incidence Rate Ratios

While raw anomaly counts effectively highlight airspaces with absolute, volume-driven errors, they fail to isolate the underlying structural integrity of the infrastructure. To definitively answer the primary spatial research question—*Does the intrinsic risk of an aircraft experiencing a severe kinematic anomaly change based solely on the global region it traverses, independent of localized traffic density?*—a Negative Binomial regression model was evaluated. By utilizing continuous flight exposure as a mathematical offset, this model effectively strips away traffic volume, isolating the pure infrastructure-driven disparities in surveillance stability. The model (N = 260 FIRs) successfully converged, revealing pronounced macro-regional risk differentials across the global airspace (Table 7).

It is critical to mathematically contextualize the selection of the European (EUR) baseline. As demonstrated previously in Table 6, the EUR macro-region exhibits a severely bimodal distribution, simultaneously containing the world's most stable, highly redundant airspaces (e.g., LFRR, LIMM) and the most volatile sectors subjected to intense electromagnetic interference (e.g., UUWV, URRV). While this bimodal nature mathematically establishes the EUR baseline (IRR 1.000) as an "average of extremes," this characteristic deliberately positions it as the optimal global anchor for a descriptive infrastructure model. Because the study strictly adheres to standard ICAO administrative boundaries to maintain nonattributional objectivity, explicitly modeling "conflict zones" as separate categorical predictors or subdividing Europe into custom geopolitical sectors would violate the methodological mandate. Consequently, while the EUR aggregate does inflate the baseline relative to pristine airspaces, it serves as a highly representative global median. This deliberate choice allows regions with extreme stability (such as North America, which emerges as a statistical outlier with an IRR of 0.006) and regions with extreme volatility (such as Africa) to be accurately quantified against a realistic, real world operational average rather than an idealized, perfectly stable baseline. To contextualize the sensitivity of this baseline selection: if the regression were anchored to the most stable macro-region (NAM), the relative risk profile of volatile regions would mathematically escalate by two orders of magnitude (for instance, the African IRR would exceed 900 relative to NAM). Therefore, the bimodal European baseline successfully centers the regression, preventing extreme mathematical scaling while preserving proportional comparability across the global airspace.

Evaluated against this aggregate baseline, the Negative Binomial regression model establishes North America (NAM) as the empirical standard for global consistency (IRR 0.006, $p < 0.001$). Notably, the North Atlantic (NAT), which often appears highly anomalous under raw volumetric counts due to sequential oceanic latency errors, yields an Incidence Rate Ratio of 0.966 ($p = 0.929$), rendering its difference from the European baseline not statistically significant. This indicates that the fundamental baseline risk of an aircraft experiencing a kinematic anomaly in the NAT corridor is essentially identical to navigating Europe. The drastic reduction in apparent risk compared to raw anomaly counts confirms that oceanic teleportation is heavily driven by burst artifacts from individual aircraft over extended data gaps, which the unique-aircraft normalization successfully mitigates. By contrast, the African macro-region (AFI) emerges as structurally volatile, exhibiting a rate 5.87 times higher for unique aircraft experiencing kinematic inconsistencies per exposure hour compared to the aggregate European baseline (IRR 5.867, $p < 0.001$). It is critical to interpret this specific IRR not as a measure of poor avionics health, but as a direct quantification of sparse terrestrial receiver density. The lack of overlapping crowdsourced receivers in this region inherently causes sequential packet loss, resulting in elevated rates of geometric teleportation artifacts strictly from the perspective of the ground network.

The exceptionally low IRR for NAM (0.006) also reflects the static GET methodology: because the threshold is derived from the most volatile global hours, it suppresses detection in high-integrity airspaces, establishing NAM as the empirical "floor" of global observability. The forest plot (Fig. 12) illustrates this relative risk by macro-region compared to the European baseline, utilizing a log scale to highlight the extreme stability in North America versus volatility in Africa.

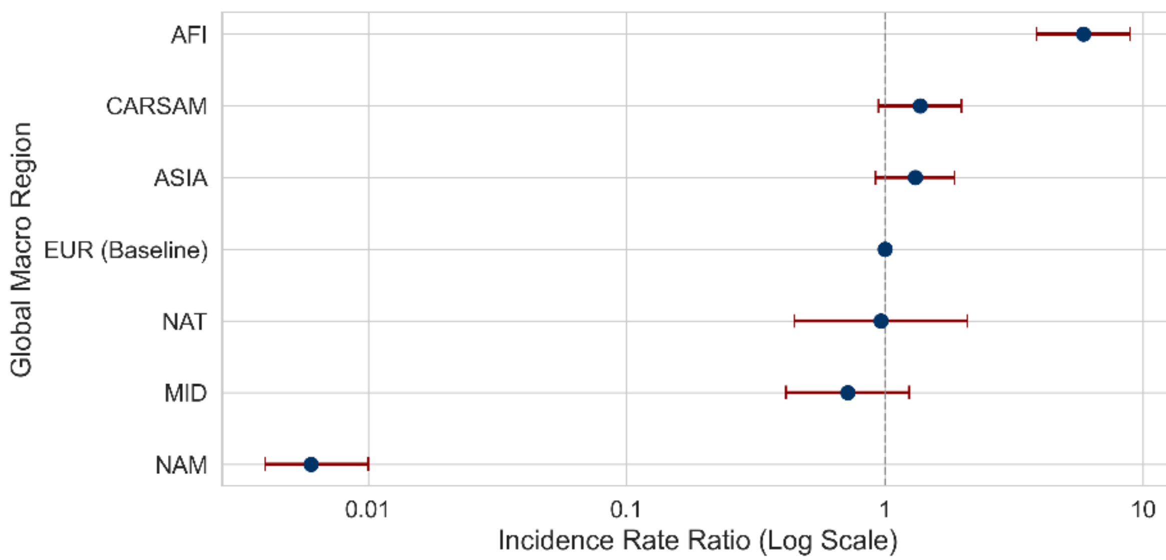

Fig. 12 Forest Plot of Incidence Rate Ratios.

Table 7 Regional Incidence Rate Ratios (IRRs) derived from negative binomial regression with categorical predictors.

Macro Region	**IRR** ($\exp(\beta)$)	**95% CI**	**P-Value**	**Sig.**
EUR (Baseline)	1.000	—	—	—
ASIA (Asia)	1.310	[0.921, 1.861]	0.1326	ns
NAT (North Atlantic)	0.966	[0.447, 2.084]	0.9289	ns
NAM (North America)	0.006	[0.004, 0.010]	< 0.001	***
AFI (Africa)	5.867	[3.870, 8.895]	< 0.001	***
MID (Middle East)	0.718	[0.415, 1.241]	0.2354	ns
CARSAM (Caribbean/S.A.)	1.369	[0.945, 1.983]	0.0963	ns

Note: $N = 260$ FIRs. All IRRs are expressed relative to the EUR reference region (IRR = 1.000). Significance codes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns = not significant.

D. Dimensional Robustness and Metric Coupling

To comprehensively address the third study objective, the analysis must evaluate how kinematic anomalies manifest and interact across different physical dimensions of flight.

1. Metric-Specific Spatial Decomposition

While overall systemic volatility models account for all five extracted kinematic flags, to explicitly separate infrastructure-driven coverage gaps from safety-critical avionics errors, the spatial decomposition focuses strictly on the three primary spatial dimensions: derived speed (teleportation), altitude (sensor divergence), and directional instability (Table 8). The spatial distribution of these metrics confirms that surveillance instability is highly dimension-dependent. The 100% stacked bar chart (Fig. 13) decomposes the top 10 volatile FIRs by error mode, demonstrating a stark visual divergence between oceanic latency gaps (Speed) and terrestrial sensor interference (Altitude).

Extreme derived speed artifacts (teleportation) are predominantly concentrated in oceanic and remote corridors (e.g., KZHU Houston Oceanic and KZWY New York Oceanic). As detailed concurrently in Table 8, isolating these top anomalous FIRs by individual dimensions explicitly confirms that line-of-sight coverage gaps, rather than true kinematic instabilities, mathematically induce these geometric positional jumps, empirically confirming a degradation of surveillance continuity in these regions. It is important to note that the Teleportation flag triggers strictly when the calculated v_{derived} exceeds the static GET envelope regardless of the exact temporal interval (Δt). This successfully captures both localized coordinate jumps where Δt is small, as well as extended oceanic latency gaps up to the 60-second limit. Altitude divergence and directional instability similarly exhibit high rates in these oceanic regions, reflecting how extended data latency can simultaneously corrupt multiple state parameters upon reconnection. However, they also reveal severe sensor-specific anomalies in terrestrial airspaces subject to significant electromagnetic interference (e.g., OPLR Lahore). This indicates that while oceanic gaps affect all kinematic dimensions equally through unobservability, discrete terrestrial events reflect distinct physical realities of the global airspace.

2. Anatomy of the Compound Effect: Metric Co-Occurrence

The previously detailed spring volatility analysis (Table 5) highlighted instances where messages triggered multiple kinematic thresholds simultaneously. To quantify this structural coupling, a co-occurrence matrix was generated for a peak spring volatility hour (May 1, 15:00 UTC).

The analysis reveals a substantial coupling across multiple kinematic dimensions, most notably between *Directional Instability* and *Vertical Jolt*, with 44.8% of flagged aircraft exhibiting simultaneous errors. By visualizing the probability of simultaneous threshold breaches (Fig. 14), further severe co-occurrences emerge, including Altitude Divergence and Directional Instability (23.2%). This intense off-diagonal clustering suggests a shared vulnerability to GNSS integrity loss, simultaneously corrupting vertical rate, barometric, and horizontal state estimations. Physically, this

Table 8 **Decomposition of Kinematic Instability.**

Metric / Error Mode	Top FIRs	Macro Region	Unique AC Count	Rate /Hr
Derived Speed (v_{derived}) (Teleportation)	KZHU (Houston Oceanic)	CARSAM	48,148	2.49
	KZWY (New York Oceanic)	CARSAM	40,291	2.07
	MUFH (Havana)	CARSAM	10,182	0.47
	HLLL (Tripoli)	MID	4,882	0.40
	OPLR (Lahore)	ASIA	3,780	0.36
	UDDD (Baku)	EUR	3,815	0.36
	MDCS (Santo Domingo)	CARSAM	8,053	0.28
	ZJSA (Sanya)	ASIA	4,708	0.23
	ZYSH (Shenyang)	ASIA	3,256	0.23
	TNCF (Curaçao)	CARSAM	4,130	0.22
Altitude Difference (Δh) (Sensor Divergence)	KZHU (Houston Oceanic)	CARSAM	29,010	1.50
	KZWY (New York Oceanic)	CARSAM	25,334	1.30
	OPLR (Lahore)	ASIA	2,021	0.19
	OJAC (Amman)	MID	5,054	0.17
	URRV (Rostov)	EUR	4,308	0.17
	HLLL (Tripoli)	MID	1,919	0.16
	ZJSA (Sanya)	ASIA	3,010	0.15
	MDCS (Santo Domingo)	CARSAM	4,126	0.14
	UDDD (Baku)	EUR	1,458	0.14
	MUFH (Havana)	CARSAM	3,036	0.14
Directional Jitter ($\Delta\psi$) (Directional Instability)	KZHU (Houston Oceanic)	CARSAM	35,679	1.84
	KZWY (New York Oceanic)	CARSAM	29,588	1.52
	OPLR (Lahore)	ASIA	1,515	0.15
	MUFH (Havana)	CARSAM	2,970	0.14
	OJAC (Amman)	MID	3,355	0.12
	ZJSA (Sanya)	ASIA	2,549	0.12
	ZYSH (Shenyang)	ASIA	1,675	0.12
	MDCS (Santo Domingo)	CARSAM	3,241	0.11
	KZMA (Miami)	CARSAM	39,105	0.09
	UDDD (Baku)	EUR	728	0.07

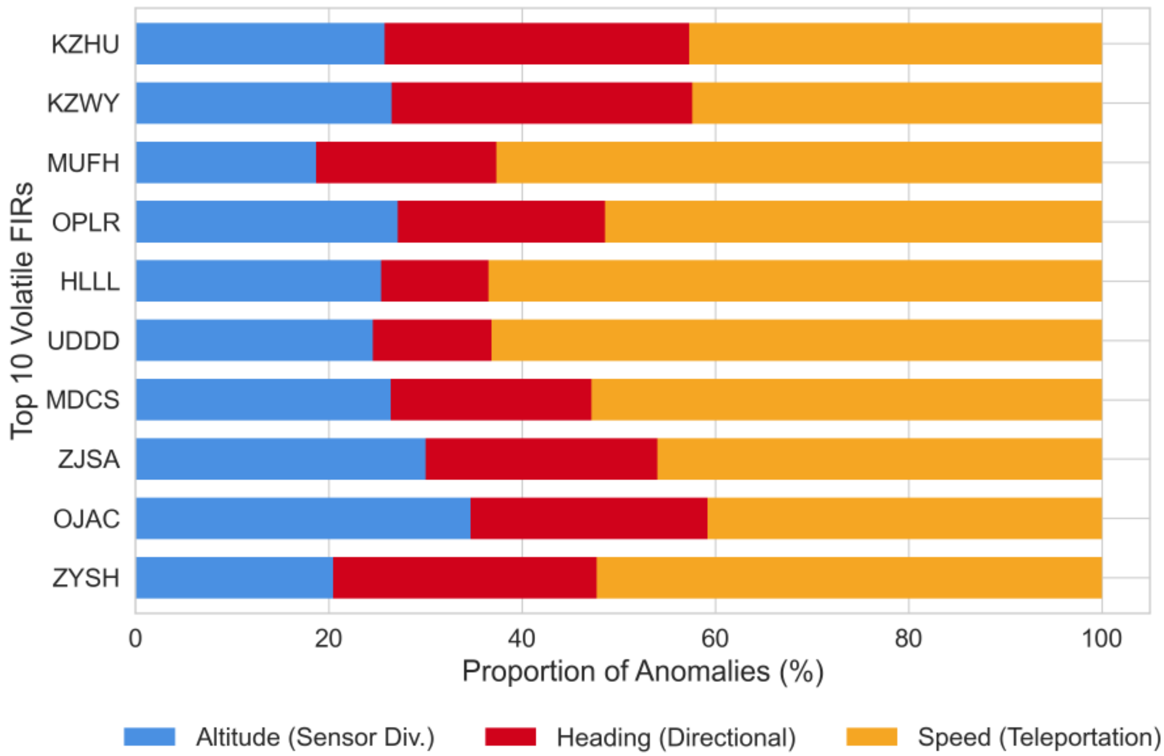

Fig. 13 Kinematic Anomaly Signatures.

mechanism likely reflects scenarios where the sudden loss of high-precision 3D-GNSS forces the avionics to decouple or fall back upon lower-accuracy, uncalibrated inputs (such as raw barometric data or degrading Inertial Reference Systems), resulting in synchronized data divergence across multiple axes. In contrast, *Speed (Teleportation)* anomalies remained largely decoupled from other metrics, indicating that derived speed anomalies are driven by distinct physical mechanisms (latency/coverage gaps) compared to Sensor Divergence anomalies.

V. Discussion

A. Interpretation of Key Findings

The results of this study substantially recharacterize how kinematic instability is distributed within the global ADS-B network. The decomposition of spatial anomalies (Table 8 and Fig. 13) provides compelling evidence that surveillance failures are not monolithic. While this study maintains a strictly descriptive mandate, the concentration of substantial altitude and heading divergences over the Eastern European and Middle Eastern macro-regions aligns geographically with areas subject to externally documented airspace warnings regarding electromagnetic interference (EMI) [27], further supporting recent findings that dense GNSS anomalies serve as robust indicators of localized geopolitical instability [26]. In these regions, terrestrial receivers are successfully capturing data, but the internal geometry of the broadcast state vectors is structurally corrupted. Conversely, the concentration of extreme "derived speed" artifacts in

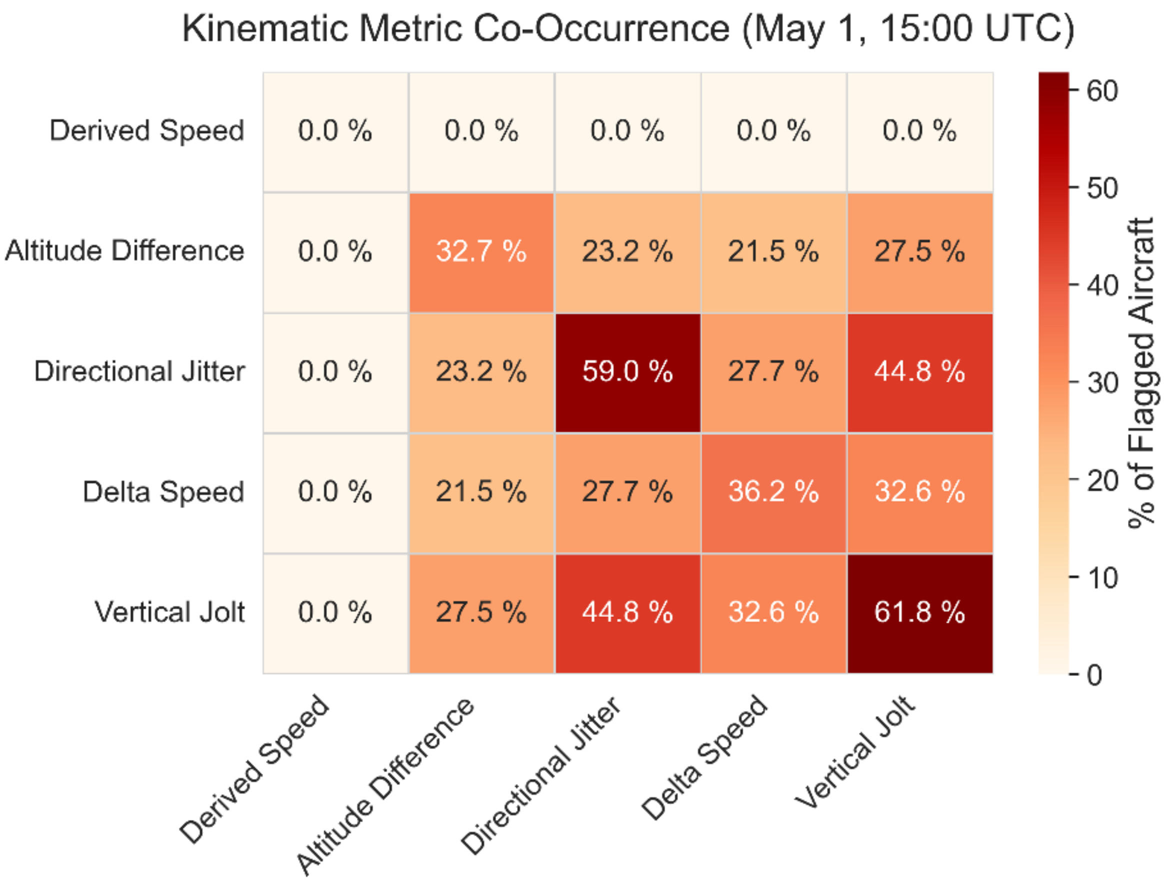

Fig. 14 Kinematic Metric Co-Occurrence Matrix.

remote oceanic corridors such as the western Atlantic (e.g., KZWY New York Oceanic) and the Gulf of Mexico (e.g., KZHU Houston Oceanic) confirms that infrastructure-induced data latency (line-of-sight coverage gaps) functionally manifests as geometric teleportation. This represents a known physical limitation of terrestrial architecture rather than a physical safety risk posed by the aircraft.

The elevated volatility in the African (AFI) macro-region (IRR 5.867) necessitates nuanced interpretation. Due to sparse receiver density, state vector updates are highly fragmented. The Ghost Record detection mechanism geometrically calculates speed over these extended Δt gaps. Consequently, the high anomaly rate reflects genuine systemic unobservability—from the perspective of the ground network, the aircraft is kinematically unstable and unpredictable—even if the onboard transponder is functioning nominally.

Conversely, the North American IRR of 0.006 should be interpreted as a lower-bound estimate of relative stability under the static GET framework. Because re-anchoring the categorical baseline rescales all IRRs proportionally without altering their ordinal ranking, NAM remains the minimum-IRR region under any baseline choice; the finding is therefore not an artifact of baseline selection.

Systemic Elasticity Under Peak Load. Furthermore, the temporal analysis establishes a critical insight into the limits of the analyzed surveillance network. The data do not support the hypothesis of severe non-linear degradation or a "saturation zone" at high volumes. Instead, the near-zero correlation identified in Fig. 4 demonstrates that the OpenSky ingestion infrastructure successfully handles mid-year peak operational loads without a disproportionate spike in the rate of kinematic anomalies. This indicates that current crowdsourced server-side architectures are remarkably elastic, maintaining structural consistency even as message throughput triples from the early-year troughs (January/February) to the peak global travel season in July. This distinction is vital to avoid misleading operational interpretations by Air Navigation Service Providers (ANSPs) who operate professional, dedicated receiver networks with significantly higher throughput capacities. Nevertheless, the observed increase in variance during peak volume periods suggests that stochastic observability degradation remains a factor during rapid operational expansions. This suggests that while server-side ingestion is robust, the inherent limitations of the ALOHA protocol and terrestrial receiver sparsity continue to define the network's global noise floor.

B. Implications for Automated Safety Systems and Policy

The establishment of the GET framework provides Air Navigation Service Providers (ANSPs) with a deployable, objective mechanism to monitor the health of their local airspace. When evaluating the impact of unmitigated sensor divergence and Ghost Record teleportation anomalies, it is critical to distinguish between ground-infrastructure limitations and true airborne avionics failures.

If a Ghost Record is exclusively an artifact of terrestrial line-of-sight gaps or localized ground-receiver RF saturation, the transmitting aircraft likely broadcast a complete state vector. In such scenarios, the risk is strictly confined to ground-based Air Traffic Management (ATM) and Unmanned Aircraft Systems Traffic Management (UTM) trackers that rely on these terrestrial network feeds. However, if the Ghost Record originates from a true onboard transponder failure, or if severe airspace-wide RF congestion prevents air-to-air reception, it poses a direct operational risk to airborne automated safety nets, particularly the Airborne Collision Avoidance System (ACAS X) [28, 29]. ACAS X relies on probabilistic Markov Decision Process (MDP) state models and dynamic programming to estimate aircraft trajectories. If an aircraft physically broadcasts a corrupted state vector indicating a localized 10,000 m/s speed jump, this severely disrupts the specific MDP state-transition matrices used by the surrounding airborne algorithms. The tracker may assign a non-trivial probability mass to a false collision trajectory, subsequently issuing an unwarranted Resolution Advisory (RA) and severely compromising pilot trust.

To proactively mitigate these vulnerabilities across both ground and airborne domains, we propose the following strategic recommendations for global ANSPs and automated safety architectures:

Adopt a Tiered Threshold Architecture: While the static GET envelopes established in this study serve as the necessary first step to bound absolute worst-case systemic degradations, relying solely on these global maximums compromises sensitivity in highly stable airspaces. For proactive, real-time tactical safety monitoring, ANSPs should implement a tiered approach: utilize the static GET baseline as a primary structural filter to immediately quarantine severe, physically impossible failures (e.g., massive geometric teleportation), while simultaneously deploying a secondary, dynamic rolling-mean threshold. This context-aware dynamic layer is critical to minimize Type II errors, allowing the system to detect subtle GNSS drift or low-magnitude spoofing that occurs below the global noise floor.

Implement Kinematic Plausibility Filters: Decentralized tracking algorithms (e.g., ACAS X logic) should integrate pre-processing kinematic plausibility filters based on empirical operational envelopes (such as the GET thresholds) to actively quarantine and reject malformed, unphysical state vectors before track updating occurs.

Adopt Ghost Record Monitoring: Establish the detection and quantification of “Ghost Records” (kinematic anomalies lacking standard telemetry) as a formalized Key Performance Indicator (KPI) for regional airspace health.

Integrate GET Baselines: Integrate the published GET constants (Table 2) into real-time filtering architectures to instantly flag structural degradation exceeding established empirical bounds.

Formalize Benchmarking Standards: Adopt the “GET Zero-Anomaly Rate” exhibited by High-Integrity Reference Airspaces (e.g., LFRR, LIMM) as a formal industry metric for evaluating the maturity and redundancy of terrestrial receiver networks, utilizing these regions to explicitly target infrastructure deficits.

C. Methodological Limitations

To ensure transparency and methodological rigor, the following consolidated sections address the specific limitations inherent to this study.

1. Coverage-Induced Artifacts and Receiver Heterogeneity

This study evaluates the stability of broadcast state vectors strictly as received by the terrestrial ground network (the "network-received state"). Extreme derived speed values induced by coverage gaps (latency) are therefore intentionally retained in the numerator. Because these represent calculation artifacts driven by an expanding Δt term in the Haversine equation rather than physical aircraft movement, flagging them as kinematic anomalies introduces a semantic limitation: an aircraft over the mid-Atlantic is not kinematically unstable, but it is systemically unobserved. Retaining these gaps allows the unified metric to reflect the aggregate degradation of terrestrial surveillance continuity, while the dimensional decomposition (Sec. IV.D.1) mathematically isolates infrastructure limits (Speed artifacts) from safety-critical avionics failures (Altitude and Heading divergences).

This confounding is inseparable from receiver heterogeneity. The OpenSky Network relies on crowdsourced terrestrial receivers constrained by line-of-sight physics (typically ~250 NM), so raw data collection is intrinsically biased toward wealthy, populated landmasses. The exposure offset ($\log(E_f)$) corrects for traffic volume disparities but not for receiver density. A valid critique of the regression model (Table 7) is therefore that it partly captures ground receiver density rather than isolated airspace kinematic integrity. Incorporating receiver count or average range as a covariate could theoretically separate avionics instability from coverage; however, because the study's mandate is to quantify *aggregate systemic observability* from the ground network's perspective—where a coverage gap is functionally indistinguishable from an avionics failure in terms of tracking continuity—controlling for receiver density would obscure the very infrastructure deficits this baseline seeks to highlight. Consequently, the elevated Incidence Rate Ratios in regions such as Africa (AFI) must be interpreted strictly as terrestrial unobservability rather than physical flight safety risk, and this physical limitation underscores the operational utility of space-based augmentation.

2. Spatial and Temporal Autocorrelation

The cross-sectional Negative Binomial regression assumes independence among observations (FIRs). Aviation is a highly networked system; if a single "ghost" aircraft traverses both FIR A and FIR B, their error terms become correlated, potentially compressing standard errors. Temporally, raw anomaly extractions in volatile regions exceeded 3,600 events per flight hour prior to normalization, reflecting "burst" artifacts in which a single malfunctioning transponder generated multiple simultaneous kinematic flags per second over sustained segments.

Both threats are mitigated by the same design decision: the dependent variable is defined as the *count of unique aircraft impacted* rather than raw anomaly volume, severing the high-frequency autocorrelation caused by a single transponder broadcasting thousands of sequential errors across borders. Additionally, the broad macro-regional predictor groups neighbor-effect correlations into unified regional intercepts. Future spatial econometric refinements—clustered standard errors or explicit spatial lag models—could sharpen precision, but the FIR-centric unit remains operationally vital: from an ANSP's perspective, an anomaly introduced by a transiting aircraft still represents a discrete degradation of local observability.

3. Threshold Sensitivity and the Maximum Envelope

Anchoring the threshold to the Maximum Observed Hourly Percentile establishes an explicitly conservative, "worst-case" baseline that prioritizes specificity over sensitivity. By defining normality as the most volatile hour observed, the static threshold is rendered essentially blind to minor degradation within pristine airspaces (such as LFRR or LIMM) where the local noise floor is exceptionally low: subtle sensor drift or low-magnitude spoofing passes unflagged if it falls below the global ceiling, producing an elevated Type II error rate in these specific regions. Estimating the exact miss rate would require concurrent deployment of a dynamic rolling-threshold model.

A dynamic, traffic-normalized threshold (e.g., a rolling mean with a three-sigma band) was rejected deliberately: such thresholds auto-correct for local infrastructure quality, effectively grading volatile and stable regions on different mathematical curves and obscuring the absolute macro-regional disparities the regression model explicitly seeks to quantify. Anchoring to a universal static ceiling ensures that a flagged anomaly in the African macro-region represents the same physical severity as one in European airspace—a prerequisite for fair cross-regional benchmarking. This design positions the framework as a *system-wide infrastructure benchmarking tool* rather than a *real-time tactical safety system*; formally quantifying the maximum credible noise floor is the necessary precondition for later dynamic algorithms to detect subtle spoofing without false-positive cascades during peak load.

An additional mathematical artifact exists within the directional instability threshold ($\Delta\psi$). While the final event extraction resolves 360° compass wrap-around via a modulo operator (Sec. III), the Phase I sensitivity analysis evaluated raw absolute differences to optimize throughput across 246.7 billion rows. Normal maneuvers crossing the True North vector therefore temporarily inflated the extreme tails, yielding a conservative P95 directional threshold of 14.7°. This slightly suppresses detection of minor directional jitter but reinforces the framework's design of flagging only severe, unphysical sensor decoupling.

4. The Etiology of "Ghost Records" and Sub-Frame Decoding

The ADS-B protocol inherently decouples telemetry, typically transmitting airborne position via Comm-B Data Selector (BDS) register 0,5 and airborne velocity via BDS register 0,9. "Ghost Records" can therefore originate from two distinct mechanisms: a true avionics degradation, in which the transponder or air data computer fails to generate the BDS 0,9 sub-frame, or a terrestrial infrastructure limitation, in which the ground receiver decodes the position packet but loses the subsequent velocity packet to ALOHA packet collisions or signal fading.

Because this study utilizes aggregated, decoded state vectors rather than raw Mode S hexadecimal frames, these root causes cannot be definitively disentangled. From the perspective of downstream automated safety systems (such as ACAS X or UTM trackers) ingesting these feeds, however, the distinction is operationally moot at the moment of reception: whether the aircraft failed to transmit the velocity or the network failed to receive it, the resulting state vector is kinematically corrupted and mathematically unobservable. The GET framework therefore validly flags these as severe systemic degradations, leaving forensic root-cause attribution to future localized studies utilizing raw multi-lateration frame analysis.

D. Directions for Future Research

With the global descriptive baseline established, future research must transition from static characterization to rolling-window persistence models. Evaluating the temporal decay rate of anomalies—determining whether an aircraft recovers correct telemetry within 10 seconds versus transmitting false altitudes for 30 minutes—is critical for assessing the severity of the threat to automated collision avoidance systems. Additionally, leveraging the Ghost Record detection framework to build multi-parameter ML classifiers will enable definitive root-cause attribution, mathematically differentiating between benign transponder degradation and malicious GNSS spoofing vectors [30].

Furthermore, the terrestrial baselines established in this study serve as the necessary empirical control group for future research. As global ANSPs increasingly integrate space-based Low Earth Orbit (LEO) ADS-B surveillance (e.g., Aireon) [31, 32], researchers can utilize this GET baseline to empirically quantify the reduction in oceanic "teleportation" anomalies resulting from the closure of these line-of-sight gaps. Finally, while the 2025 constants provide a static historical baseline, ANSPs adopting this framework for active monitoring are recommended to implement a dynamic, trailing-12-month recalibration protocol to continuously account for these structural changes in the surveillance infrastructure.

VI. Conclusion

This study establishes the first global, traffic-normalized baseline for evaluating the kinematic self-consistency of the Automatic Dependent Surveillance-Broadcast (ADS-B) network. By analyzing over 246 billion state vectors recorded throughout 2025 and resolving legacy survivorship bias via the novel Global Envelope Thresholding (GET) framework, the research quantified the baseline structural stability of the global airspace. The resulting models demonstrate substantial elasticity within the system, wherein the anomaly filter rate remains structurally stable under high temporal load. While seasonal traffic shifts expand message volume by 38.0%, the global infrastructure maintains a stable kinematic consistency baseline, and the previously hypothesized "saturation zone" is not observed, though extreme peak capacity increases the variance of outlier events. Furthermore, spatial mapping reveals significant regional disparities, illustrating the decoupling of infrastructure-induced latency in oceanic corridors from localized sensor divergence in terrestrial zones. Ultimately, this foundational characterization facilitates comprehensive system observability, providing the aviation industry with an objective mathematical framework and distinct high-integrity benchmarks to monitor, evaluate, and interpret the kinematic integrity of decentralized surveillance architectures worldwide.

Data Availability Statement

The raw ADS-B state vectors utilized in this foundational study were provisioned by the OpenSky Network. To guarantee total reproducibility and facilitate immediate industry benchmarking, the resulting dataset has been open-sourced as the **GET-25 Benchmark** [33]. The complete Global Envelope Thresholding (GET) configuration files, processed kinematic baselines, and a lightweight Python API allowing researchers to instantly apply the framework to their own local OpenSky or custom terrestrial receiver (e.g., RTL-SDR) datasets are available under a CC BY 4.0 license via the Zenodo repository at `https://doi.org/10.5281/zenodo.18904146`.

Funding Sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgments

The authors would like to explicitly thank the OpenSky Network, its founders, and the global network of receiver operators. The crowd-sourced collection and maintenance of this massive terrestrial surveillance database made this macroscopic systemic analysis possible.

References

[1] Schäfer, M., Strohmeier, M., Lenders, V., Martinovic, I., and Wilhelm, M., "Bringing up OpenSky: A large-scale ADS-B sensor network for research," *Proceedings of the 13th International Symposium on Information Processing in Sensor Networks (IPSN)*, 2014, pp. 83–94. https://doi.org/10.1109/IPSN.2014.6846743.

[2] National Academies of Sciences, Engineering, and Medicine, *A Review of the Next Generation Air Transportation System*, National Academies Press, Washington, D.C., 2015. https://doi.org/10.17226/21721.

[3] Bolić, T., and Ravenhill, P., "SESAR: The Past, Present, and Future of European Air Traffic Management Research," *Engineering*, Vol. 7, No. 4, 2021, pp. 448–451. https://doi.org/10.1016/j.eng.2020.08.023.

[4] Brooker, P., "SESAR and NextGen: Investing In New Paradigms," *Journal of Navigation*, Vol. 61, No. 2, 2008, pp. 195–208. https://doi.org/10.1017/S0373463307004596.

[5] van der Pryt, R., and Vincent, R., "A Simulation of Signal Collisions over the North Atlantic for a Spaceborne ADS-B Receiver Using Aloha Protocol," *Positioning*, Vol. 6, No. 3, 2015, pp. 23–31. https://doi.org/10.4236/pos.2015.63003.

[6] Strohmeier, M., Lenders, V., and Martinovic, I., "On the Security of the Automatic Dependent Surveillance-Broadcast Protocol," *IEEE Communications Surveys & Tutorials*, Vol. 17, No. 2, 2015, pp. 1066–1087. https://doi.org/10.1109/COMST.2014.2365951.

[7] McCallie, D., Butts, J., and Mills, R., "Security analysis of the ADS-B implementation in the Next Generation Air Transportation System," *International Journal of Critical Infrastructure Protection*, Vol. 4, No. 2, 2011, pp. 78–87. https://doi.org/10.1016/j.ijcip.2011.06.001.

[8] Westbrook, T., "A Taxonomy of Radiofrequency Jamming and Spoofing Strategies and Criminal Motives," *Journal of Strategic Security*, Vol. 16, No. 2, 2023, pp. 68–80. https://doi.org/10.5038/1944-0472.16.2.2081.

[9] Prevot, T., Rios, J., Kopardekar, P., Robinson, I., John E., Johnson, M., and Jung, J., "UAS Traffic Management (UTM) Concept of Operations to Safely Enable Low Altitude Flight Operations," *16th AIAA Aviation Technology, Integration, and Operations Conference*, AIAA, Washington, D.C., 2016, pp. AIAA 2016–3292. https://doi.org/10.2514/6.2016-3292.

[10] Sun, J., Vu, H., Ellerbroek, J., and Hoekstra, J. M., "pyModeS: Decoding Mode-S Surveillance Data for Open Air Transportation Research," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 21, No. 7, 2020, pp. 2777–2786. https://doi.org/10.1109/TITS.2019.2914770.

[11] Olive, X., "traffic, a toolbox for processing and analysing air traffic data," *Journal of Open Source Software*, Vol. 4, No. 39, 2019, p. 1518. https://doi.org/10.21105/joss.01518.

[12] Ying, X., Mazer, J., Bernieri, G., Conti, M., Bushnell, L., and Poovendran, R., "Detecting ADS-B Spoofing Attacks Using Deep Neural Networks," *2019 IEEE Conference on Communications and Network Security (CNS)*, 2019, pp. 187–195. https://doi.org/10.1109/CNS.2019.8802732.

[13] Li, L., Hansman, R. J., Palacios, R., and Welsch, R., "Anomaly detection via a Gaussian Mixture Model for flight operation and safety monitoring," *Transportation Research Part C: Emerging Technologies*, Vol. 64, 2016, pp. 45–57. https://doi.org/10.1016/j.trc.2016.01.007.

[14] Leonardi, M., and Sirbu, G., "ADS-B Crowd-Sensor Network and Two-Step Kalman Filter for GNSS and ADS-B Cyber-Attack Detection," *Sensors*, Vol. 21, No. 15, 2021, p. 4992. https://doi.org/10.3390/s21154992.

[15] Smart, E., Brown, D., and Denman, J., "A Two-Phase Method of Detecting Abnormalities in Aircraft Flight Data and Ranking Their Impact on Individual Flights," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 13, No. 3, 2012, pp. 1253–1265. https://doi.org/10.1109/TITS.2012.2188391.

[16] Çevik, N., and Akleylek, S., "Comparison of Machine Learning Based Anomaly Detection Methods for ADS-B System," *Information Technologies and Their Applications. ITA 2024*, Lecture Notes in Networks and Systems, Vol. 1092, edited by G. Mammadova, T. Aliev, and K. Aida-zade, Springer, Cham, 2025, pp. 275–286. https://doi.org/10.1007/978-3-031-73420-5_23.

[17] Pik, E., Berra, M., Yearwood, J., and Garcia, J. S. D., "GPS Anomalies in Aviation: Preliminary Insights from Automatic Dependent Surveillance–Broadcast Data," *Journal of Aerospace Information Systems*, Vol. 22, No. 7, 2025, pp. 570–582. https://doi.org/10.2514/1.I011527.

[18] Tabassum, A., and Semke, W., "UAT ADS-B Data Anomalies and the Effect of Flight Parameters on Dropout Occurrences," *Data*, Vol. 3, No. 2, 2018, p. 19. https://doi.org/10.3390/data3020019.

[19] Strohmeier, M., Olive, X., Lübbe, J., Schäfer, M., and Lenders, V., "Crowdsourced air traffic data from the OpenSky Network 2019–2020," *Earth System Science Data*, Vol. 13, No. 2, 2021, pp. 357–366. https://doi.org/10.5194/essd-13-357-2021.

[20] Olive, X., Sun, J., Figuet, B., and Alligier, R., "Filtering Techniques for ADS-B Trajectory Preprocessing," *Journal of Open Aviation Science*, Vol. 2, No. 2, 2025. https://doi.org/10.59490/joas.2024.7882, URL https://doi.org/10.59490/joas.2024.7882.

[21] Pik, E., Berra, M., Yearwood, J., and Garcia, J. S. D., "Detecting GPS Anomalies in Aviation Using ADS-B: Correlating Coordinate Gaps and GPS Deviations With NOTAM Warnings," *AIAA AVIATION FORUM AND ASCEND 2024*, 2024, pp. AIAA 2024–4640. https://doi.org/10.2514/6.2024-4640.

[22] Karney, C. F. F., "Algorithms for geodesics," *Journal of Geodesy*, Vol. 87, No. 1, 2013, pp. 43–55. https://doi.org/10.1007/s00190-012-0578-z.

[23] Cameron, A. C., and Trivedi, P. K., *Regression Analysis of Count Data*, 2nd ed., Cambridge University Press, Cambridge, UK, 2013.

[24] International Civil Aviation Organization, *Location Indicators (Doc 7910) and Flight Information Region (FIR) Boundaries*, International Civil Aviation Organization (ICAO), Montreal, Canada, 2024. URL https://gis.icao.int/icaofir/, official geospatial database for global Flight Information Regions and Macro-Regional administrative boundaries.

[25] Federal Aviation Administration, "GPS and GNSS Interference Resource Guide: Jamming and/or Spoofing," Resource guide, Federal Aviation Administration, Flight Technologies and Procedures Division (AFS-400), Mar. 2026. URL https://www.faa.gov/about/office_org/headquarters_offices/avs/offices/afx/afs/afs400/afs410/GNSS/GPS_GNSS_Interference_Resource_Guide.pdf.

[26] Pik, E., Garcia, J. S. D., Smith, T. A., Kocaman, I., and Berra, M., "Predicting geopolitical instability through GNSS anomalies and air traffic data," *Journal of Transportation Security*, Vol. 18, No. 1, 2025, p. 29. https://doi.org/10.1007/s12198-025-00323-w.

[27] Felux, M., Fol, P., Figuet, B., Waltert, M., and Olive, X., "Impacts of Global Navigation Satellite System Jamming on Aviation," *Navigation*, Vol. 71, No. 3, 2024. https://doi.org/10.33012/navi.657.

[28] Strohmeier, M., Martinovic, I., and Lenders, V., "Understanding realistic attacks on airborne collision avoidance systems," *Journal of Transportation Security*, Vol. 15, No. 1-2, 2022, pp. 87–118. https://doi.org/10.1007/s12198-021-00238-2, URL https://doi.org/10.1007/s12198-021-00238-2.

[29] Holland, J. E., Kochenderfer, M. J., and Olson, W. A., "Optimizing the Next Generation Collision Avoidance System for Safe, Suitable, and Acceptable Operational Performance," *Air Traffic Control Quarterly*, Vol. 21, No. 3, 2013, pp. 275–297. https://doi.org/10.2514/atcq.21.3.275.

[30] Liu, Z., Lo, S., Blanch, J., and Walter, T., "GNSS Spoofing Detection and Localization Using ADS-B Data," *Proceedings of the 37th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2024)*, 2024, pp. 796–803. https://doi.org/10.33012/2024.19749.

[31] Garcia, M. A., Dolan, J., and Hoag, A., "Aireon's initial on-orbit performance analysis of space-based ADS-B," *2017 Integrated Communications Navigation and Surveillance Conference (ICNS)*, 2017, pp. 4A1–1–4A1–8. https://doi.org/10.1109/ICNSURV.2017.8011919.

[32] Budroweit, J., Eichstaedt, F., and Delovski, T., "Aircraft Surveillance From Space: The Future of Air Traffic Control?: Space-Based ADS-B, Status, Challenges and Opportunities," *IEEE Microwave Magazine*, Vol. 25, No. 12, 2024, pp. 68–76. https://doi.org/10.1109/MMM.2024.3443754.

[33] Pik, E., Garcia, J. S. D., and Smith, T. A., "GET-25: The Global Spatiotemporal ADS-B Kinematic Consistency Benchmark," https://doi.org/10.5281/zenodo.18904146, 2026. [Dataset].