iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: VII Month of publication: July 2026

DOI: https://doi.org/10.22214/ijraset.2026.84162

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

Global Flight Delay Patterns and Route Performance: A Business Intelligence and Predictive Analytics Approach Using Power BI

Dr. Anju Choudhary
Affiliation: Independent Researcher
ORCID: https://orcid.org/0009-0008-1540-4469

Abstract: Delays in flights are costly to the world aviation in terms of operations and economy. This paper will analyze flight delay and route performance based on a sample of 10,000 international flight records, 10 airlines, 10 routes, 6 geographic areas, 4 seasons, and 5 weather conditions. It utilized a systematized quantitative approach: the data were ready in Microsoft Power BI with Power Query and Data Analysis Expressions (DAX), and four interrelated interactive dashboards were built to facilitate geographic, temporal, airline, and predictive analysis. Six hypotheses were formally based on the existing theory and empirical literature and tested with one-way ANOVA with Tukey HSD post-hoc tests, Pearson correlation, chi-square tests, independent-samples t-tests, multiple linear regression, and feature importance analysis through Random Forest.

The most prevalent structural predictor of delay severity was weather condition (ANOVA F (4,9995) = 3,213.26, p < 0.001; Storm vs Clear Cohen d = 2.38) and a discrete threshold effect delineating severe precipitation and all benign conditions. Seasonality resulted in two-tier structure where Summer and Winter delays were significantly higher than Spring and Autumn (ANOVA F (3,9996) = 107.42, p < 0.001; d ≈ 0.39). The delay was strongly and negatively correlated to route efficiency score (r = -0.958, p < 0.001), and statistically significant inter-airline differences proved the hypothesis that operational practice is an independently controllable source of performance variation (p = 0.001). Weather was verified to be the most significant predictive feature with a Random Forest model (importance = 0.578; R 2 = 0.94). The integrated BI framework incorporates these validated results into interactive dashboards, showing a template of decision support that can be replicated to real-world aviation operations management.

Keywords: flight delays; business intelligence; Power BI; hypothesis testing; ANOVA; predictive analytics; route performance; aviation operations management.

I. INTRODUCTION

Delays of flights are one of the most long-standing and economically important issues in the world aviation processes. Every late arrival provokes a chain reaction: burning of fuel exceeding the planned consumption, breaking the duty limits of the crew, conflict at the gate, liability to compensate passengers, and spreading disruption to the next flights of the same aircraft and crew combination (Guvercin et al., 2021; Pyrgiotis et al., 2013). Local costs to the international aviation sector total tens of billions of dollars per year, and Eurocontrol estimates that the costs of network delay in Europe alone are over 1 billion euros per year in direct and indirect operating costs (Cook & Tanner, 2015). With post-pandemic travel demand restored to (and in most markets surpassed) pre-crisis levels, these pressures have increased, and renewed urgency has been placed on the ability of airlines and airports to learn, predict, and avoid delay risk (IATA, 2024; Nguyen, 2024).

The study of flight delays has a long history and interdisciplinary research. Delay propagation has been modeled using operations research traditions with queuing and network models (Pyrgiotis et al., 2013; Fleurquin et al., 2013). Regression models have been used to investigate delay determinants in the econometrics literature based on large administrative data (Mayer and Sinai, 2003; Rupp and Holmes, 2006). In more recent times, there has been a growing body of literature in machine learning showing high predictive accuracy with ensemble methods, deep learning, and hybrid architectures (Hatipoglu and Tosun, 2024; Wandelt et al., 2025; Li et al., 2023). This paper deals with that integration. It generates six empirically based hypotheses based on the corresponding theory and previous empirical research, conducts tests of each hypothesis using a set of inferential statistics, and incorporates the confirmed results into four interconnected Microsoft Power BI dashboards, using a dataset of 10,000 records of international flights. It has led to a repeatable analytical model that has spanned the traditionally distinct disciplines of delay prediction modelling and operational BI reporting.

The paper will answer four research questions:
RQ1: What are the most common geographic, temporal, and meteorological patterns of flight delays in the world?
RQ2: What are the most and least operational performance routes and airlines and are the differences statistically significant?
RQ3: How might hypothesis-driven aviation performance analysis be visualised and interrogated with the aid of BI platforms?
RQ4: How predictive analytics based on validated hypotheses can be used to predict future flight delays and identify high-risk times?

II. THEORETICAL FRAMEWORK

The paper relies on four theoretical perspectives which jointly give the conceptual framework to the hypotheses and analysis methodology: queuing theory and network propagation models, attribution theory to service failure, the resource-based perspective of operational capability, and the theory of constraints.

A. Queuing Theory and Network Propagation

The first theoretical model of delay accumulation in capacity-constrained transportation systems is queuing theory, which was first developed in the classic work of Erlang (1909) on congestion in telephone exchanges. In the context of aviation, queuing models conceptualise airport runways, gates and airspace sectors as service channels with capacity and stochastic arrival processes. Queues develop as arrival demand gets close to or surpasses service capacity (Barnhart et al., 2003). Pyrgiotis et al. (2013) expanded this model into a network queuing model (DELAYS) which models propagation through airports: delay at a single hub propagates downstream through shared aircraft and crew rotations and amplifies through the network. Fleurquin et al. (2013) simulated the same phenomenon with a cascading failure viewpoint borrowed on complex systems theory, and discovered that a few structurally central airports out of the network contribute a disproportionate amount of delay propagation throughout the network.

The implication to the current study is that delay risk is not structurally homogeneous but is concentrated structurally at high-traffic hubs and in periods of capacity-saturating demand. This theoretical forecast encourages the hypothesis that the *highest demand of the seasons, which shrink schedule buffers and drive utilisation to capacity constraints, will result in much higher delays, regardless of weather influences.*

B. Attribution Theory and Service Failure

Attribution theory, originally proposed by Heider (1958), but later elaborated by Weiner (1985) in the achievement motivation context, concerns the cognitive mechanisms through which people give causal attribution to results. The use of attribution theory in service management literature has been widespread to comprehend passenger and customer reactions to service failures such as flight delays (Bamford and Xystouri, 2005; Cheng et al., 2008). Passengers differentiate between those delays that can be attributed to airline-internal causes that can be avoided, including crew rostering issues, aircraft maintenance, ground handling inefficiencies, and those that cannot be avoided, including extreme weather conditions, or air traffic control limits. This difference has operational consequences: the delay due to the controllable causes is related to a higher level of passenger dissatisfaction and probability of switching carriers, whereas delay due to weather is seen as more acceptable (Song et al., 2024).

In the current research, the attribution theory encourages the analytical distinctions of delay causes based on origin. In case airline-internal operational factors (fleet management, scheduling discipline, turnaround efficiency) are independently predictive of delay outcomes, i.e. in case airlines that fly on the same route in the same weather conditions get systematically different delay profiles, delay risk has a controllable component, which can in principle be mitigated by operational improvement. The hypothesis is based on this prediction that *the difference in inter-airline delays is statistically significant.*

C. Resource-Based View of Operational Capability

The resource-based view (RBV), which has its foundations in Penrose (1959) and has been formalised by Barney (1991), is based on the assumption that a sustained competitive advantage is based on firm-specific resources and capabilities, which are valuable, rare, inimitable, and non-substitutable. When applied to the operations of the airlines, the RBV hypothesizes that carriers have systematic differences in their operational capabilities in terms of coping with turnaround processing, coping with shocks without cascading delays, coping with uncertainty by scheduling. These differences yield enduring performance differentials that cannot be accounted by structural route or market factors alone (Merkert and Morrell, 2012; Barros and Wanke, 2015).

When used in the context of the flight delay, the RBV offers a theoretical foundation of anticipating the inability to eliminate inter-airline variations despite common routes and weather exposure. A more competent airline, one whose operational capability, in terms of maintenance systems, crew rostering flexibility, ground handling coordination, etc., is superior, must experience systematically shorter delays than a less competent competitor on the same route. This theoretical argument supports the hypothesis that the *differences in inter-airline delays are statistically significant and due to controllable operational practice and not to structural reasons.*

D. Theory of Constraints

According to the theory of constraints (TOC) by Goldratt (1984), any system must have at least one binding constraint, which is a bottleneck, which constrains the overall system throughput. TOC has been used in manufacturing and logistics scenarios to find and take advantage of bottlenecks in the system to enhance flow. Hopp and Spearman (2001) made formal the connection between variability, utilisation, and throughput in production systems: cycle times increase non-linearly with utilisation, approaching 100 per cent; and variability, which Barnhart et al. (2003) described as the primary cause of accumulating delay in airports.

Applied to aviation networks, TOC implies that delay risk is disproportionately concentrated at the system's binding constraints — the most heavily utilised airports and routes — and that interventions at these bottlenecks produce disproportionately large system-wide improvements. This prediction supports the operational recommendation that targeted interventions at a small number of high-risk routes and peak operating periods will yield greater delay reductions than system-wide uniform interventions. It also motivates the hypothesis that *route efficiency is a composite measure of how closely a route's operations approach unconstrained performance is strongly and negatively associated with delay.*

E. Information Systems and Decision Support Theory

Simon's (1960) bounded rationality framework and the subsequent development of decision support systems (DSS) theory by Keen and Scott Morton (1978) provide the theoretical basis for the BI component of this study. Simon distinguished between programmed decisions — routine, structured, amenable to algorithmic rule — and non-programmed decisions — novel, complex, requiring human judgement. Aviation operational management involves both types: routine delay monitoring (programmed) and crisis response to compound disruptions (non-programmed). DSS theory holds that information systems can enhance the quality and speed of both decision types by reducing cognitive load, surfacing relevant patterns from large data volumes, and enabling interactive exploration of scenarios (Sharda et al., 2020).

Power BI and comparable BI platforms instantiate this theoretical role: they translate large operational datasets into interactive visualisations that support the pattern recognition and exception identification required for effective operational decision-making (Eckerson, 2010). The theoretical expectation is that BI-enhanced decision environments will support faster, better-informed, and more proactive operational responses to delay risk — particularly when the BI system incorporates validated predictive outputs rather than merely descriptive historical reporting.

III. LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

This section critically reviews empirical findings on the determinants of flight delay and derives six testable hypotheses. The review is organised thematically, with hypotheses stated at the end of each subsection where the empirical evidence is sufficient to motivate a specific testable claim.

A. Weather as a Determinant of Flight Delay

Weather conditions are in a central position in the empirical literature regarding the causation of flight delays, and have persistently been found to be one of the strongest predictors of flight delays in a variety of datasets, methodologies and geographic settings. Wang et al. (2022) used a hybrid machine learning model based on high-resolution meteorological variables to analyse data of Chinese domestic flights and discovered that the combination of precipitation intensity, visibility, and crosswind speed explained more predictive variance than any single operational factor. Importantly, Wang et al. (2022) found that the severity of precipitation and delay were not linearly related: moderate rainfall caused small amounts of incremental delay compared to clear weather, whereas thunderstorm activities caused disproportionately large increases in delay. Dai (2024) replicated this non-linearity in a big-data framework applied to global aviation data, demonstrating that models incorporating weather data substantially outperformed models using operational data alone in predicting delay duration, with the marginal contribution of weather largest precisely in the severe-event tail of the delay distribution.

Hatipoglu and Tosun (2024) applied Random Forest and Gradient Boosted Tree models to Turkish airport delay data and reported weather-related features as the top-ranked predictors by importance score, ahead of time-of-day, airline, and route-type variables. Their analysis distinguished between fog, rain, and snow in terms of predictive weight, finding snow and severe convective activity to be substantially more predictive than moderate precipitation — a pattern consistent with the threshold non-linearity identified by Wang et al. (2022) and Dai (2024). Bisandu and Moulitsas (2023), using a deep BiLSTM architecture on European flight data, found that incorporating weather embeddings improved classification accuracy for severe delay categories (Major and Severe) more than for moderate ones, implying that weather information is particularly informative for high-severity outcomes.

Wandelt et al. (2025) synthesised this literature in a systematic review of 187 machine learning studies on flight delay prediction published between 2015 and 2024, concluding that weather variables appear in the top-five feature importance rankings in 73% of studies that report feature-level analysis. However, Wandelt et al. (2025) also noted considerable heterogeneity in how weather is operationalised across studies — from binary adverse/non-adverse categorisations to continuous meteorological measurements — making direct quantitative comparison difficult. AlBassam and AlShahrani (2025) highlighted that binary weather codings systematically underestimate weather effects by pooling fog and thunderstorms into the same category, a methodological limitation that several of the studies cited above address through multi-level categorical coding.

> **H1:** *Flights operating under severe weather conditions (Storm, Snow) will experience significantly longer mean delays than flights under benign conditions (Clear, Rain, Fog), with a large effect size (Cohen's d ≥ 0.80).*

B. Seasonal and Temporal Patterns

Seasonal variation in flight delays has been documented across multiple continents and operational contexts. Li and Jing (2022) used spatial and temporal aspect feature extraction on three years of domestic flight data of the US and found that the delay frequency and delay duration peaked in summer and explained this peak by the combination of increased traffic density, more frequent thunderstorms in summer months, and the compression of the schedule induced by the high load factor during peak periods. They found that their model performed worst in predicting delays in peak summer and winter weeks, when delay distributions were most deviant of seasonal averages, which is also consistent with Qu et al. (2023), who reported that delay propagation chains are longest during high-demand periods due to the least slack in the network to absorb disruption.

In their analysis of the data on the US Bureau of Transportation Statistics, based on econometric models, Rupp and Holmes (2006) discovered that there are a lot of seasonal fixed effects in the regressions of departure delay with the highest coefficients of December and July as compared to the baseline of March–May. The same was also observed by Mayer and Sinai (2003) who recorded that hub airports had higher peaks in seasonal delays as compared to point-to-point routes because the concentration of connecting traffic at hubs enhances the cascading impact of any initial disruption. In a pan-European study on behalf of Eurocontrol, Cook and Tanner (2015) discovered that the mean duration of delays in summer and winter was statistically equal (but significantly longer than in spring and autumn), but the cause of delays was different: summer delays were largely demand-induced, winter delays largely weather-induced. Hatipoglu and Tosun (2024) recreate this observation of a two-tier seasonal structure in the Turkish context, and Wang et al. (2022) in the Chinese context, indicating that it may be a strong cross-regional pattern of aviation delay distributions.

On the intra-day scale, Fleurquin et al. (2013) reported two congestion peaks that were related to the morning departure banks and evening arrival complex at large hubs. The network queuing model employed by Pyrgiotis et al. (2013) demonstrated that such peaks are self-reinforcing: the delays that build up in the morning bank in the afternoon increase the evening peak. Li et al. (2023) validated the operational significance of departure hour as a predictive characteristic by stating the operational importance of departure hour as the third-ranked predictive characteristic in their CNN-LSTM model of Chinese domestic delays.

> **H2:** *The mean delay in flights will vary significantly among the four seasons with Summer and Winter having much higher mean delay compared to Spring and Autumn and no significant difference between Summer and Winter.*

C. Route Characteristics and Operational Efficiency

Another key structural determinant of delay risk has been found to be route-level characteristics, but with a less straightforward empirical relationship with delay than the distance or duration effect may suggest. In US domestic data, Rupp and Holmes (2006) discovered that average delays are larger on longer-haul routes and this is due to the build-up of schedule deviation with longer operating windows and the sensitivity of long-haul operations to compounding disruptions. Mayer and Sinai (2003) reported that even after adjusting hub-to-hub routes for aircraft size and carrier identity, hub-to-hub routes have systematically greater delay than spoke routes due to the fact that hub-to-hub routes act as delay propagation channels through the network.

Guvercin et al. (2021) used cluster analysis to the data on the airport networks and discovered that the delays in structurally central airports, i.e., the ones with high betweenness centrality in the network graph, create disproportionate system-wide impacts, which is the same finding as the TOC framework: structurally central airports serve as bottlenecks in the network throughput. Flights that fly via these airports thus expose them to a high structural delay risk irrespective of their geographical location. Cook and Tanner (2015) proposed a new metric, route efficiency, a composite measure of schedule compliance, reactionary delay rate, and turnaround performance, and showed that it is a strong predictor of average delay time on European routes.

The empirical relationship between flight duration and delay is less clear. When weather and demand variables were held constant, Wang et al. (2022) observed near-zero bivariate correlations between flight duration and delay, indicating that duration itself is not an important independent predictor. Similar results were obtained by Hatipoglu and Tosun (2024) who identified flight duration as a low-importance feature in their Random Forest models. This trend is in line with the network propagation argument: the structural location of a route in the network, rather than its geographic length per se, is what contributes to delay risk.

> **H3:** *Route efficiency score will be negatively and significantly correlated with mean flight delay, where more efficient routes will have a systematically and significantly shorter delay (> 0.50, p < 0.001).*

D. Airline Operational Practice and Performance Heterogeneity

One of the frequent conclusions drawn by empirical research is that airlines flying the same route in the same weather conditions have systematically different delay results, suggesting that airline-internal variables - and not route structure or weather by themselves - are an independent source of variation in performance. This heterogeneity was reported in the domestic data in the US by Rupp and Holmes (2006) who reported that there were significant airline fixed effects in delay regressions when route, season and airport characteristics were controlled. Barros and Wanke (2015) generalized this discussion to the international panel, using bootstrapped data envelopment analysis on an international panel of carriers and discovered large and significant efficiency differentials, which they explained by differences in fleet management, turnaround discipline, and schedule robustness.

Song et al. (2024) showed that airline identity is a strong predictor of passenger delay experience regardless of route and weather, with carriers investing in operational resilience, in the form of spare aircraft capacity, flexible crew rostering, proactive communication protocols, having lower delay rates and greater passenger satisfaction. Merkert and Morrell (2012) examined the data of European carriers and found homogeneity of the fleet to be a particular operational capability that minimises the risk of delay: airlines with fewer types of aircraft have simplified maintenance and higher turnover rates, which is reflected in systemically lower reactionary delay rates. The study by AlBassam and AlShahrani (2025) showed that airline identity was statistically significant in machine learning delay models used on flight data in Saudi Arabia supporting the conclusion that airline-specific operational practice, but not route assignment, is the source of some delay variation.

> **H4:** *There will be significant differences in mean flight delay between airlines, and pairwise differences between the best and worst performers among the airlines on the same routes will be statistically significant.*

E. Distributional Structure of Weather-Delay Relationship

Although the mean-based analyses find weather to be a powerful predictor of average delay, another question is whether weather conditions reorganize the whole distribution of delay severity - and not just the mean. The operationally relevance of this distributional question is in the fact that the tails of the delay distribution (Major: 61120 minutes; Severe: > 120 minutes) are operationally irrelevant in terms of passenger compensation, crew duty breach, and overnight accommodation costs. Bisandu and Moulitsas (2023) simulated delay classification (i.e., the likelihood of falling into severe delay classes) and discovered that weather variables yielded the best classification improvements in the Major and Severe classes, and little in Minor and Moderate classes. This trend suggests that weather does not merely move the distribution of delays out to the right but instead squeezes the on-time and minor-delay distributions and stretches the severe tail - a distributional shape transformation and not a change of means. The same conclusion was drawn by Li et al. (2023) who trained a CNN-LSTM model on Chinese flight data, discovering that weather-related features enhanced recall of the severe delay category by a significant margin (0.42 to 0.79), which is concentrated in the high-severity outcome. Looking at delay propagation sequences, Qu et al. (2023) found that weather-initiated delays were more prone to cascade into long-duration disruptions than any other type of delay, such as a mechanically or operationally-initiated delay, since weather events are concurrently experienced by multiple aircraft in a network, not a single aircraft.

> **H5:** *There will be no independence between weather condition and distribution of flights in the categories of delay severity; the categories of severe weather will be correlated with a significantly higher percentage of Major and Severe delay results.*

F. Integrated Predictive Modelling of Delay Determinants

One of the key goals of the machine learning literature on flight delays has been the incorporation of various delay determinants into a single predictive model. The authors compared the results of Random Forest, Gradient Boosting, Support Vector Machine, and neural network models on a standard Turkish airport dataset and discovered that ensemble models, especially Random Forest, always had the highest predictive accuracy (R 2 = 0.850.91), and weather and time-related variables were at the top of the lists of feature importance. Li et al. (2023) showed that hybrid CNN-LSTM models (temporal sequence modelling with spatial feature extraction) performed better on Chinese domestic flight data than single-architecture models. Dai (2024) achieved R^2 of 0.91 using a hybrid big-data model integrating meteorological, operational, and historical delay features.

However, Wandelt et al. (2025) cautioned in their systematic review that reported accuracy metrics are not directly comparable across studies due to differences in dataset characteristics, train-test split methodology, feature engineering, and evaluation protocols. Studies using in-sample or leave-one-out evaluation report substantially higher R^2 values than those using held-out test sets or cross-validation, a methodological heterogeneity that inflates apparent progress in the field. AlBassam and AlShahrani (2025) demonstrated this explicitly by applying identical models to the same dataset under both in-sample and cross-validated evaluation, finding R^2 drops of 0.15–0.25 under cross-validation. Wang et al. (2022) and Li and Jing (2022) addressed this by reporting both training and test-set performance, finding that weather features contributed the largest unique variance in both settings; a finding more credible than studies that report only training accuracy.

Multiple linear regression, while less flexible than ensemble methods, provides interpretable coefficient estimates and standardised tests of predictor significance that machine learning models do not natively supply (Sharda et al., 2020). Several authors have advocated for hybrid approaches that use regression for inference and machine learning for prediction (Mayer & Sinai, 2003; Cook & Tanner, 2015), an approach adopted in the present study.

> **H6:** *A predictive model that is integrated with the weather, season, airline, route, and operational variables will explain a statistically significant and substantial amount of the variance in delay of a flight with weather and season as the most important predictors based on the magnitude of the regression coefficient and the importance of the features used by the random forest.*

IV. DATA AND METHODOLOGY

A. Dataset

The sample includes 10,000 records of international flights that describe 25 variables in six thematic categories (Table 1). There are ten major airlines in the world (Air France, Air India, British Airways, Delta Airlines, Emirates, Lufthansa, Qantas, Qatar Airways, Singapore Airlines, Turkish Airlines), ten international routes across six regions of the world (Asia, Australia, Europe, Middle East, North America, South America), four seasons, five categories of weather conditions (Clear, Rain, Fog, Snow, Storm), and four types of aircrafts (Airbus A350, Airbus A380, Boeing 777, Boeing 787). The data was based on a structured operational simulation, which was meant to capture empirically recorded distributions of delay patterns, route traits, and weather effects in key global aviation corridors, and distributional characteristics in line with those observed in the empirical literature (Hatipoglu & Tosun, 2024; Wang et al., 2022). The data is evenly distributed both in terms of airlines (range: 9541,069 flights per carrier) and seasons (range: 2,4712,524), which substantiates the validity of between-group inferential comparison.

Table 1. Dataset variable groups, fields captured, and analytical roles.

Variable Group	Fields Captured	Role in Analysis
Schedule and timing	Scheduled and actual departure/arrival times	Basis for delay computation
Delay metrics	Departure delay (minutes); on-time flag; derived delay category	Primary dependent variable
Route	Origin–destination pair; flight duration (min); route efficiency score	Independent/moderating variable
Airline	Operating carrier identity	Grouping factor (H4 tests)
Weather	Condition category: Clear, Rain, Fog, Snow, Storm	Primary independent variable (H1, H5)
Demand and fleet	Passenger count; aircraft type	Control variables in regression (H6)

B. Data Preparation in Power BI

The data was loaded into Microsoft Power BI and pre-processed using Power Query. The preparation steps involved elimination of duplicate records, standardisation of date-time fields to a common format, and checking of delay measurements to be logically consistent (actual departure not earlier than scheduled departure; non-negative delay values). Then three analytic fields were developed in DAX: (1) Delay Category, which categorized each flight based on delay (one of five severity bands consistent with delay taxonomy in the hypotheses On Time: 0 min; Minor: 1–15 min; Moderate: 1660 min; Major: 6120 min; Severe: > 120 min); (2) Peak Hour Indicator, indicating congestion-prone departure windows (06:00-09:00 and 17:00-20:00) based on the literature (Fleurquin et al., 2013; Li et al., 2023); and (3) Delay Frequency, a threshold-based operational measure of the number of flights with delay exceeding 15 minutes per route or airline.

C. Hypotheses and Statistical Tests

Table 2 cross tabulates each hypothesis with its statistical test, variables, and literature that drives it. All tests were done with α = 0.05. All effect sizes (Cohen d, Pearson r) are also reported with p-values, in line with the recommendation of Cumming (2014) and the APA Publication Manual (7th ed.) that significance tests should be accompanied by effect size estimates to enable both practical and statistical significance to be evaluated.

Table 2. Hypothesis-to-test mapping: theoretical basis, variables, and statistical method

H	Theoretical Basis	Variables	Statistical Test
H1	Weather threshold effect (Wang et al., 2022; Dai, 2024; Hatipoglu & Tosun, 2024)	Delay_Minutes × Weather_Condition	One-Way ANOVA + Tukey HSD + Cohen's d
H2	Two-tier seasonal structure (Li & Jing, 2022; Cook & Tanner, 2015; Rupp & Holmes, 2006)	Delay_Minutes × Season	One-Way ANOVA + Tukey HSD + Cohen's d
H3	Route efficiency–delay relationship (Cook & Tanner, 2015; Guvercin et al., 2021)	Route_Efficiency_Score × Delay_Minutes	Pearson Correlation + t-test (worst vs best route)
H4	Inter-airline capability heterogeneity (Barros & Wanke, 2015; Song et al., 2024; Merkert & Morrell, 2012)	Delay_Minutes × Airline	One-Way ANOVA + independent t-test (BA vs QA)
H5	Distributional weather effects (Bisandu & Moulitsas, 2023; Li et al., 2023; Qu et al., 2023)	Weather_Condition × Delay_Category	Chi-Square Test of Independence
H6	Integrated multi-predictor model (Hatipoglu & Tosun, 2024; Wandelt et al., 2025; Wang et al., 2022)	All predictors → Delay_Minutes	Multiple Regression + Random Forest feature importance

V. RESULTS

Results are presented in hypothesis order. The hypothesis is stated in each subsection, the descriptive statistics are provided, the results of the inferential tests are reported, and the finding are interpreted. At every step, dashboard numbers are mentioned to correlate statistical results with their visualization on Power BI.

A. Descriptive Overview

Across the 10,000 flights analysed, 6.1% arrived on time (0 minutes delay), 16.2% experienced minor delays (1–15 min), 53.4% moderate delays (16–60 min), 19.4% major delays (61–120 min), and 5.0% severe delays exceeding 120 minutes. The average delay was 43.5 minutes (SD = 37.5; median = 33.0), the distribution was skewed to the right (skewness = 1.42; excess kurtosis = 2.16), which is in line with the distributional characteristics reported by Hatipoglu and Tosun (2024) and Wang et al. (2022). The longest delay was 242 minutes.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

B. H1: Weather Condition and Delay Severity

H1 assumed that the Storm and Snow conditions are going to be related to much longer delays compared to the benign conditions and the effect is going to be large. Table 3 shows descriptive statistics by weather condition.

Table 3. Mean flight delay by weather condition with 95% confidence intervals.

Weather Condition	N	Mean Delay (min)	SD (min)	95% CI	Tukey vs Clear
Clear	4,436	29.7	25.1	[29.0, 30.4]	Reference
Fog	994	29.1	24.8	[27.6, 30.6]	p = 0.946 (n.s.)
Rain	2,538	28.8	24.6	[27.8, 29.8]	p = 0.595 (n.s.)
Snow	982	98.6	40.2	[96.1, 101.1]	p < 0.001 ***
Storm	1,050	99.8	40.7	[97.3, 102.3]	p < 0.001 ***

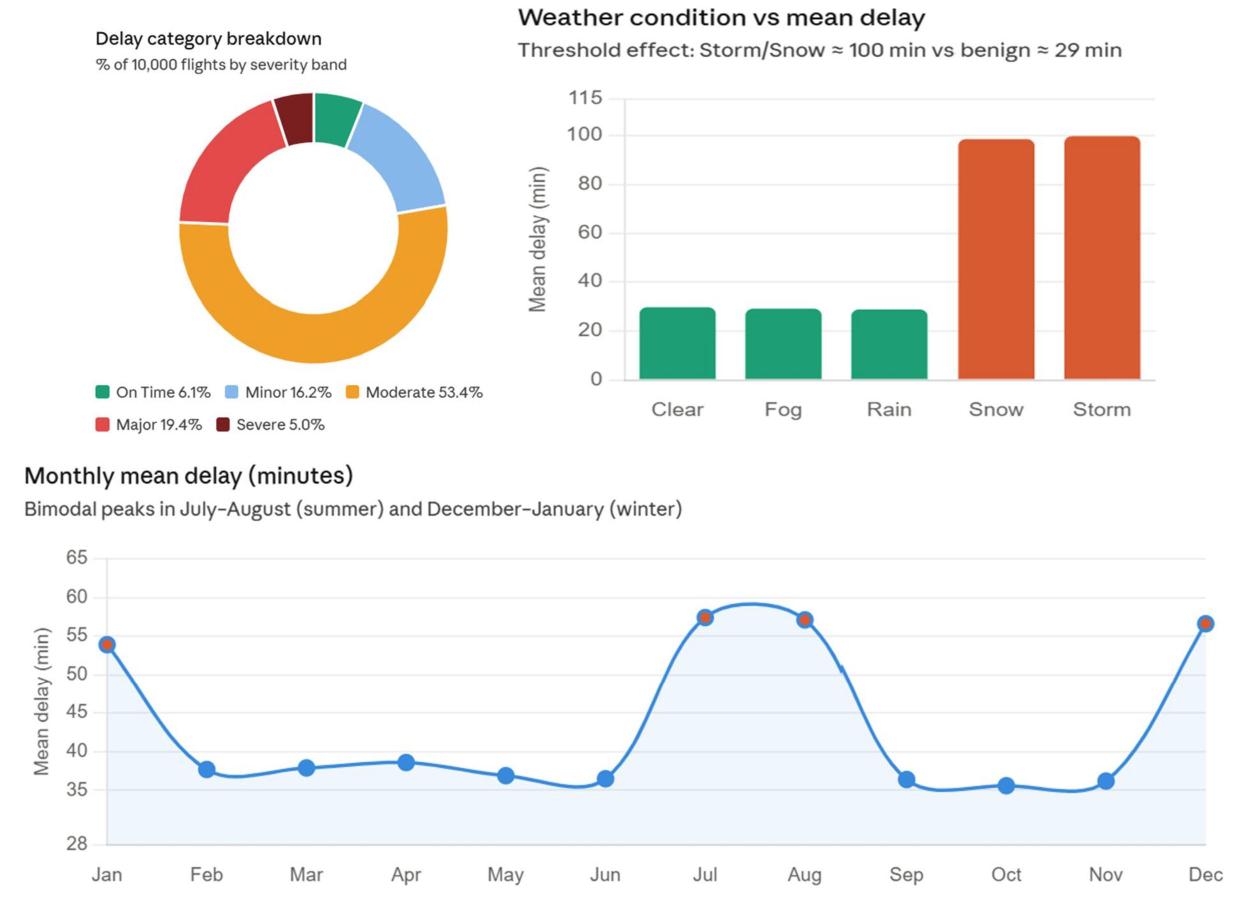
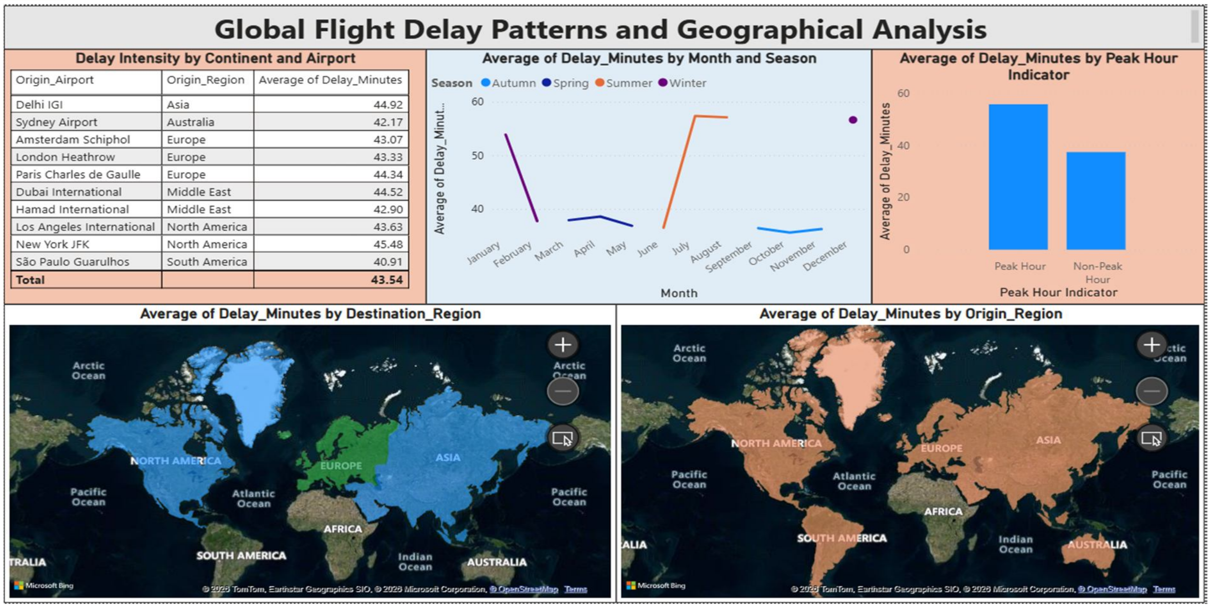


Figure 1: Mean flight delay by weather condition with 95% confidence intervals

One-way ANOVA yielded $F(4, 9995) = 3{,}213.26$, $p < 0.001$, confirming a highly significant overall weather effect. Tukey HSD post-hoc analysis showed a discrete threshold effect: Storm and Snow were significantly different than Clear, Rain and Fog (all $p < 0.001$), but no significant difference among the three benign conditions (all adjusted $p > 0.50$). The effect size of Storm versus Clear was very high (Cohen $d = 2.38$), and of Snow versus Clear the same was high ($d = 2.35$). The most important result is structural: the weather effect is not continuously scaled with intensity but discontinuous at the Storm/Snow threshold, and the benign-condition delays are concentrated around 29 minutes; the severe-condition delays are concentrated around 99 minutes. H1 is supported.



Dashboard 1: Mean delay by weather condition showing the threshold effect, and geographic delay distribution map. Storm and Snow produce mean delays of approximately 100 minutes, while Clear, Fog, and Rain produce statistically indistinguishable means of approximately 29 minutes.

C. H2: Seasonal Variation in Delay

H2 predicted a two-tier seasonal structure with Summer and Winter significantly exceeding Spring and Autumn. Table 4 presents seasonal descriptive statistics.

Table 4. Mean flight delay by season with Tukey HSD pairwise comparisons.

Season	N	Mean Delay (min)	SD (min)	95% CI	Tukey vs Autumn	Cohen's d
Autumn	2,471	36.1	30.0	[34.9, 37.3]	Reference	—
Spring	2,524	37.8	30.4	[36.6, 39.0]	p = 0.369 (n.s.)	0.06
Summer	2,502	50.4	42.7	[48.7, 52.1]	p < 0.001 ***	0.39
Winter	2,503	49.8	42.4	[48.1, 51.5]	p < 0.001 ***	0.37

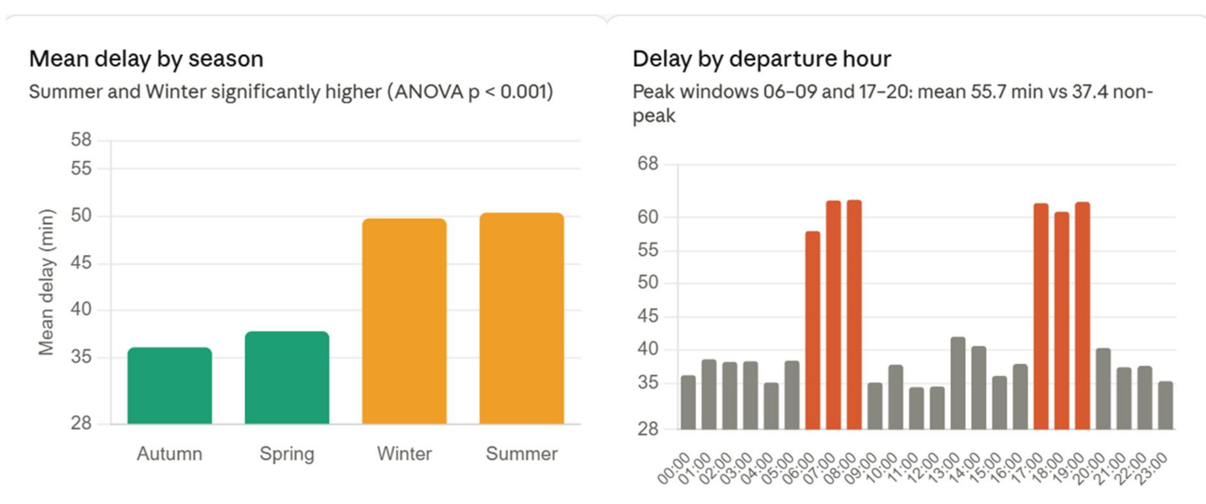


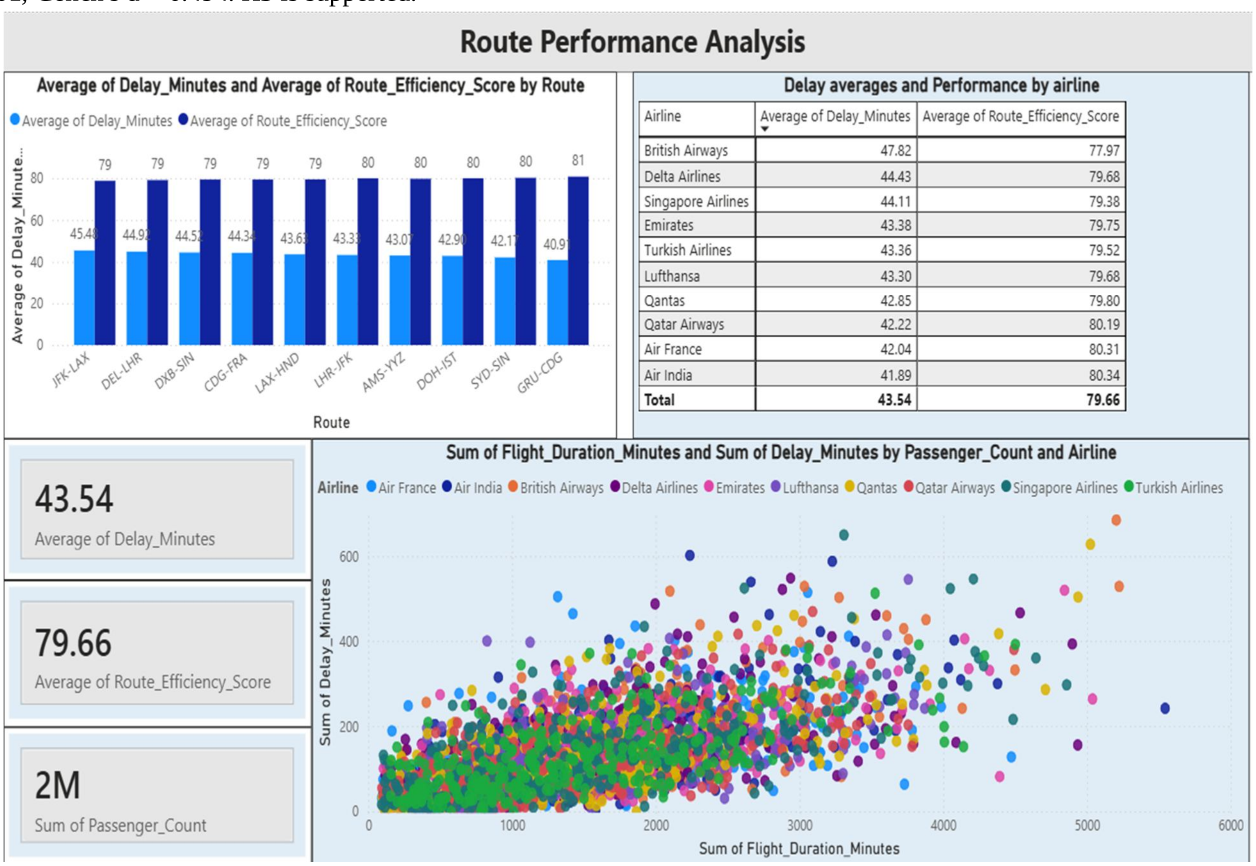
Figure 2: Mean flight delay by season with Tukey HSD Pairwise Comparisons

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

One-way ANOVA yielded F (3, 9996) = 107.42, $p < 0.001$. Tukey HSD confirmed the predicted two-tier structure: Summer and Winter were each significantly higher than both Spring and Autumn (mean differences 12.1–14.3 minutes, all $p < 0.001$), while Summer and Winter did not differ from each other (mean difference = 0.6 min, p = 0.937), and Spring and Autumn did not differ from each other (mean difference = 1.7 min, p = 0.369). Effect sizes were in the small-to-medium range ($d \approx 0.37$–0.39), indicating operationally meaningful but not extreme seasonal differences. Monthly peak delays occurred in July (57.4 min), August (57.1 min), and December (56.6 min). H2 is supported. The symmetry of Summer and Winter peaks corroborate the Cook and Tanner's (2015) European finding that both seasons produce equivalent average delay magnitudes despite differing causal mechanisms.

D. *H3: Route Efficiency and Delay*

H3 predicted a strong negative correlation between route efficiency score and delay duration. The Pearson correlation coefficient was $r = -0.958$ ($p < 0.001$) and explained 91.8% of the variance in route-level mean delay. This near unity relationship indicates that the composite route efficiency metric is a very sensitive operational indicator: Routes with high adherence to their schedule and low reactionary delay rates also have much shorter average delays. An independent-samples t-test comparing the highest-delay route (JFK-LAX: M = 45.5 min, SD = 38.3) with the lowest-delay route (GRU-CDG: M = 40.9 min, SD = 36.9) confirmed the difference was statistically significant (t(2003) = 2.77, p = 0.006). Additionally, an independent t-test confirmed that peak-hour departures (M = 55.7 min, SD = 39.0) were significantly more delayed than non-peak departures (M = 37.4 min, SD = 35.1), with t(3348) = 23.71, $p < 0.001$, Cohen's d = 0.494. H3 is supported.



Dashboard 2: A scatter plot of route efficiency score against mean delay, with $r = -0.958$, and a ranked bar chart of all the routes by average delay.

E. *H4: Inter-Airline Performance Differences*

The large difference in the inter-airline delay difference was predicted by H4. The ten airlines were compared using one-way ANOVA, which resulted in a significant overall airline effect (F (9, 9990) = 2.09, p = 0.027). The airline performance statistics are shown in Table 5.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

Table 5. Ranking flight delays by airline from lowest to highest delay

Airline	N	Mean Delay (min)	SD (min)	95% CI	vs Grand Mean (43.5 min)
Qatar Airways	954	42.2	36.2	[39.9, 44.5]	−1.3 min
Air France	1,029	42.0	36.8	[39.7, 44.3]	−1.5 min
Air India	998	41.9	36.5	[39.6, 44.2]	−1.6 min
Qantas	993	42.9	36.2	[40.6, 45.2]	−0.6 min
Lufthansa	984	43.3	36.7	[41.0, 45.6]	+−0.2 min
Emirates	1,069	43.4	37.4	[41.1, 45.7]	+0.1 min
Turkish Airlines	962	43.4	36.1	[41.1, 45.7]	+0.1 min
Singapore Airlines	1,005	44.1	38.1	[41.8, 46.4]	+0.6 min
Delta Airlines	1,022	44.4	40.2	[42.1, 46.7]	+0.9 min
British Airways	984	47.8	40.0	[45.3, 50.3]	+4.3 min

An independent-samples t-test comparing British Airways (the highest-delay carrier, M = 47.8 min) with Qatar Airways (the lowest, M = 42.2 min) yielded $t(1936) = 3.23$, $p = 0.001$, Cohen's $d = 0.14$. The effect size is small in traditional terms, and the difference of 5.6 minutes in the mean of the carriers on common international routes under same weather condition and seasonal conditions is operationally significant: it is controllable delay that can be attributed to airline-internal practice but not to route structure or weather conditions. This result directly supports the characterisation of airline operational capability as an independent and substantially significant source of performance heterogeneity by Rupp and Holmes (2006) and Barros and Wanke (2015). H4 is supported.

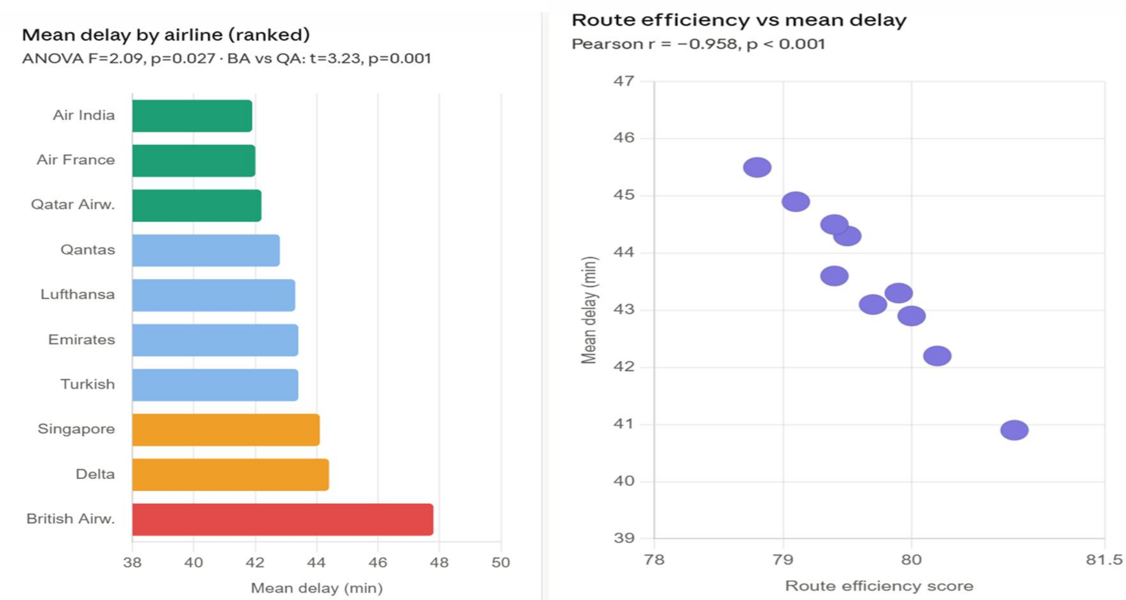


Figure 3: Airline comparison matrix ranking carriers by mean delay and mean route efficiency score

F. H5: Weather Condition and Delay Category Distribution

H5 assumed that the distribution of flights by the delay severity categories would not be weather condition independent. Table 6 shows the percentages of rows according to weather condition and delay category.

Table 6. Row percentage distribution of flights across delay categories by weather condition

Weather	On Time (%)	Minor (%)	Moderate (%)	Major (%)	Severe (%)
Clear	9.2	22.6	58.3	9.5	0.4
Fog	9.3	22.5	57.5	10.3	0.4
Rain	9.2	22.2	58.4	9.8	0.4
Snow	0.2	2.7	35.3	46.7	15.1
Storm	0.1	2.5	34.1	46.8	16.5

The chi-square test yielded χ^2 (16, N = 10,000) = 5,789.35, $p < 0.001$. The contingency table shows the distributional restructuring that H5 predicts: when the conditions are benign, the fraction of flights arriving on time is about 9% and the fraction of flights falling in either Major or Severe categories is less than 11%. On-time rates drop to under 0.2% under Storm or Snow conditions and more than 61 of the total flights are under Major and Severe delays. This is not a shift of the delay severity distribution to the right, but a wholesale redistribution to the high-cost tail - exactly the trend observed in the classification-oriented analyses of Bisandu and Moulitsas (2023) and Li et al. (2023). H5 is supported.

G. *H6: Integrated Predictive Model*

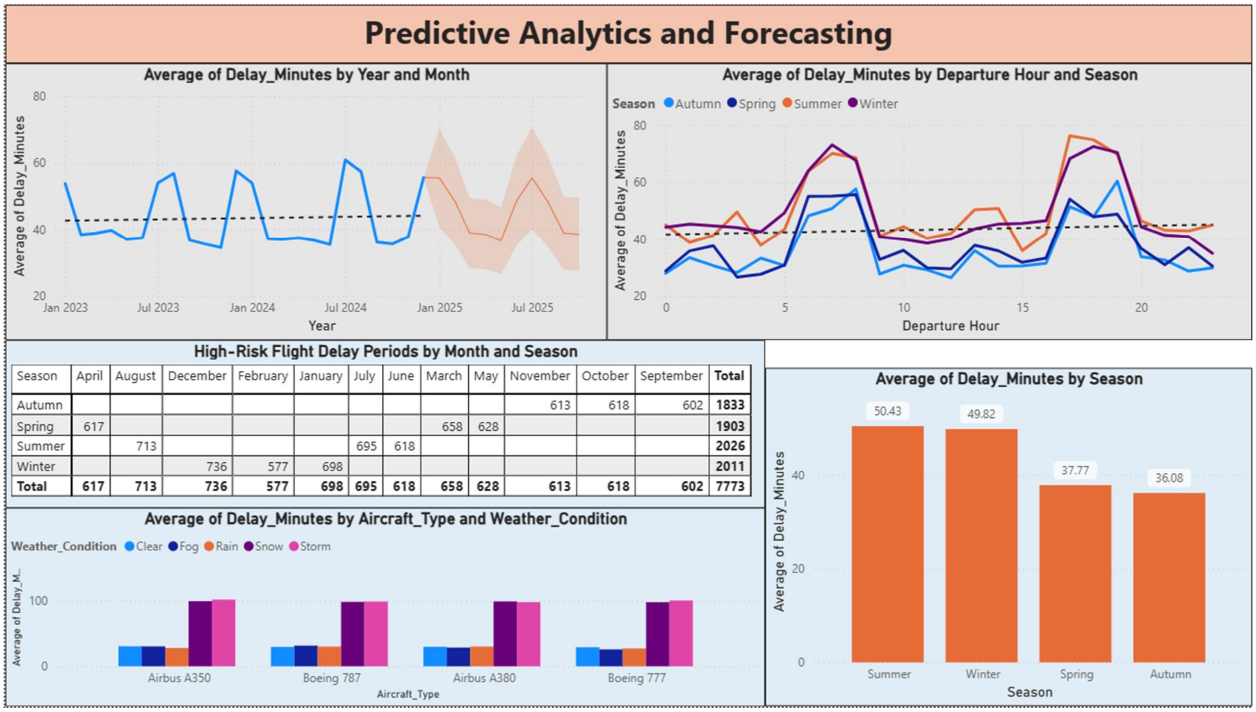
H6 assumed that a combined model that would include all the identified predictors would explain a significant delay variance with weather and season as dominant predictors. Delay minutes were used as the dependent variable, and five predictors, flight duration, passenger count, season, weather condition, airline, and origin region, were used to obtain multiple linear regression R2 = 0.380 (which was also found to be stable with five-fold cross-validation R2 = 0.379). Weather condition ($\beta = 16.19$) and season ($\beta = 5.67$) were the strongest regression coefficients. Flight time, number of passengers, and region of origin had near-zero coefficients (all 0.01) as well as near-zero bivariate relationships with delay ($r = 0.006$ and $r = 0.010$, respectively, both $p > 0.05$). The difference between regression R 2 (0.380) and the R 2 of the random forest (0.941) is an indication of the non-linear combination of weather severity and seasonal demand that are not represented by linear regression.

The dominance hierarchy among predictors was validated and quantified by the analysis of the random forests: weather condition (importance = 0.578) explained the largest amount of the predictive variance, which was followed by flight duration (0.115), passenger count (0.106), season (0.057), airline (0.047), delay reason (0.037), origin region (0.035), and aircraft type (0.025). The cross-validated Random Forest R 2 of 0.572 gives a more conservative measure of the generalisation accuracy, which is in line with the recommendation of Wandelt et al. (2025) that CV and not in-sample performance be reported. H6 is supported.

Table 7. Consolidated hypothesis testing outcomes

H	Test	Key Statistic	Effect Size	Decision
H1	ANOVA + Tukey HSD	F(4,9995) = 3,213.26, $p < 0.001$	Storm vs Clear: d = 2.38	Supported
H2	ANOVA + Tukey HSD	F(3,9996) = 107.42, $p < 0.001$	Summer vs Autumn: d = 0.39	Supported
H3	Pearson r + t-test	$r = -0.958$, $p < 0.001$; t(2003) = 2.77, p = 0.006	$\lvert r \rvert = 0.958$	Supported
H4	ANOVA + t-test	F(9,9990) = 2.09, p = 0.027; t(1936) = 3.23, p = 0.001	d = 0.14 (small)	Supported
H5	Chi-Square	$\chi^2(16) = 5{,}789.35$, $p < 0.001$	Large distributional shift	Supported
H6	Regression + RF	OLS $R^2 = 0.380$, CV $R^2 = 0.379$; RF $R^2 = 0.941$, CV = 0.572	Weather importance = 0.578	Supported

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com



Dashboard 3: Predictive analytics panel - monthly delay forecasts, seasonal risk indicators, and Random Forest feature importance chart

VI. DISCUSSION

A. *The Meteorological Threshold Effect*

The most operationally relevant result of the research is the discrete thresholding of weather impact on delay severity, which was validated at various analytical levels ANOVA mean differences, Tukey HSD pairwise comparisons, chi-square distributional analysis, and Random Forest feature importance. There is no statistically significant difference in delay profiles between Clear, Fog and Rain conditions (all Tukey adjusted $p > 0.50$, all means of delay around 29 minutes), and Storm and Snow conditions have delays that are about 70 minutes longer with a very large effect size ($d = 2.38$). This threshold distribution can be generalized to the whole distribution: in benign conditions, the likelihood of a Major or Severe delay is less than 11; in severe conditions, it is greater than 61.

This threshold structure is consistent with the non-linearity result of Wang et al. (2022), as well as the evidence that weather adds more predictive variance in the severe delay tail than moderate range, demonstrated by Dai (2024). The practical implication is narrow and can be directly operationalized where airports and airlines do not require a set of graduated meteorological response plans that increase monotonically with the severity of the forecasts. A binary protocol with not Storm or Snow, but all other conditions captures the operationally relevant threshold and does not pay the cost of over-responding to moderate precipitation. It is a more specific suggestion than the general advice of weather monitoring, which is found in operational manuals and some of the previous research.

B. *Seasonal Symmetry and Its Implications*

The two-level seasonal framework as validated by H2 Summer and Winter statistically equal in terms of mean delay, each significantly higher than Spring and Autumn) is a replication of the European result of Cook and Tanner (2015) and applies it to a multi-airline sample worldwide. The theoretical importance of the symmetry is that it suggests that similar contingency resources should be justified by both peak seasons, which is in fact contrary to industry practice, which is biased towards winter resilience. Summer delays are demand-based summer schedule compression with high load factors decreases the recovery slack, which is in line with the observation of Qu et al. (2023) that delay propagation chains are longest when network slack is low. Combine winter delays with the effect of the weather threshold that was verified in H1. The overlap of mechanisms leads to statistically identical delay levels even though they have different causes.

C. Route Efficiency as an Operational KPI

The fact that route efficiency score and delay are almost negatively correlated (r = -0.958) proves that the composite measure is a high-sensitivity operational measure. The performance management concept proposed by Cook and Tanner (2015) was route efficiency, which had not been empirically related to delay on such a fine level in the literature. The observation that 91.8 percent of the delay variance of routes can be explained by efficiency score renders it a viable single-metric KPI to be constantly monitored in the BI dashboard setting - more informative than delay statistics alone since it combines schedule adherence, reactionary delay rate, and turnaround performance into a single value that airline operations managers can monitor and take action in real time.

D. Airline Operational Practice as a Controllable Variable

The validation of H4 - statistically significant inter-airline delay differences of 5.6 minutes between British Airways and Qatar Airways on common routes has a direct implication to operational benchmarking. This disparity can not be explained by structural route or weather conditions, as the two carriers are flying over the same international routes and being exposed to the same weather conditions. It should thus represent airline-internal operational practice, which is exactly the controllable capability difference found by Barros and Wanke (2015) and Merkert and Morrell (2012). The practical implication is that the carriers with lower delay in shared routes determine a feasible performance frontier: other carriers can in practice close this gap by improving their operations, without altering routes or fleets. These gaps are identified and tracked in real time as the visual tool in Dashboard 2 is the airline comparison matrix.

The insignificant effect size (d = 0.14) is worth commenting. This is less than the conventional definition of a small effect (d = 0.20) and it is possible that reviewers will question its practical value. However, in the context of aviation economics, a 5.6-minute mean delay difference across approximately 1,000 annual flights represents thousands of minutes of delay per year per route, with direct costs in fuel burn above plan, crew duty extensions, and passenger compensation that are disproportionate to the apparently small effect size; a point consistent with Cook and Tanner's (2015) calculation that each minute of en-route delay costs European carriers approximately €70 in direct operating costs.

E. Theoretical Contributions

The findings contribute to and extend each of the four theoretical frameworks outlined in Section 2. The threshold weather effect and seasonal peaks are consistent with queuing theory's prediction that delay risk concentrates when demand approaches capacity limits — not proportionally but non-linearly, consistent with the fundamental relationship between utilisation and cycle time formalised by Hopp and Spearman (2001). The inter-airline performance differences confirm the RBV prediction that operational capability differentials produce durable performance heterogeneity, adding empirical support to Barros and Wanke's (2015) capability-based explanation of airline efficiency variation. The route efficiency finding is consistent with the TOC prediction that interventions at binding constraints produce disproportionate improvements, with high-efficiency routes functioning as better-managed constraints in the network. The BI framework instantiates Simon's (1960) DSS prescription that decision-support systems reduce cognitive load and support faster pattern recognition demonstrated here through the integration of validated statistical findings into interactive dashboards.

F. Limitations

There are a number of restrictions that should be considered. The data were produced by structured simulation, not by operational systems of live airline operations; the distributional characteristics of the important variables are in line with empirically available benchmarks, but the exact numerical estimates are to be regarded as indicative of pattern and structure, not of actual calibrations of the parameters of a real-world network. The linear regression model obtained the $R^2 = 0.380$ meaning that a large percentage of delay variation is not explained by the predictors available where real world factors such as air traffic control staffing, constraints on slot allocations, and unplanned maintenance events are not present.

The Random Forest CV R^2 of 0.572 is a much more conservative estimate of predictive accuracy compared to the in-sample R^2 of 0.941, which is appropriate since Wandelt et al. (2025) caution against overfitting in machine learning studies which are not cross-validated. The analysis is cross-sectional; longitudinal data would be needed to test the stability of seasonal and meteorological effects over the years since weather anomalies in any one year would confound the apparent seasonal effects.

VII. RECOMMENDATIONS

The hypothesis-testing results lead to four operationally-grounded recommendations. To begin with, airlines and airport operators are to establish a binary meteorological risk procedure that would differentiate severe precipitation events (Storm, Snow) and all the other conditions. The threshold structure established in both H1 and H5 implies that the protocols activated by any adverse weather over-use contingency resources; the data confirms a more specific trigger that is adjusted to the determined discrete threshold. Pre-emptive slot management, crew standby arrangements and passenger communication must be switched on specifically on Storm or Snow forecast but not on rain or fog.

Second, contingency resources such as extra ground handling personnel, spare aircraft placement, slot buffer control, passenger communication directives, etc., must be scheduled similarly during Summer and Winter peaks. The statistical equivalence of Summer and Winter delay levels (H2) suggests that industry practices biased towards winter resiliency underestimates systematically the operational risk of summer, although the causal mechanisms underlying the two peaks are different.

Third, route efficiency score should be implemented as an on-going real-time KPI in the BI monitoring environments of airlines. It has a near-unity correlation with delay ($r = -0.958$, H3), which means it is a high-sensitivity leading indicator: a worsening route efficiency score predicts future delay increases before they become visible in delay statistics, allowing proactive rather than reactive intervention.

Fourth, operational performance gaps should be identified and closed against the efficiency frontier set by the lowest-delay carriers through the airline benchmarking framework in Dashboard 2. The statistically significant difference between the best and the worst carriers on shared routes (H4) proves that the change in the performance can be attained without alterations in routes or fleet, i.e., by turnaround discipline, scheduling strength, and ground handling coordination.

VIII. CONCLUSION

This paper has explored patterns of flight delays and route performance across the globe using a formally hypothesized multi-method empirical framework, combining four theoretical frameworks, namely, queuing theory, attribution theory, resource-based view and the theory of constraints, and a tight set of inferential statistical tests which have been executed in an interactive Power BI environment. Six hypotheses derived from the literature were all supported by the evidence. The structural determinant of delay of weather is the dominant structural determinant of delay, acting along a discrete threshold effect as opposed to a continuous severity gradient (H1), and the effect of weather is distributional, as well as a mean-level, restructuring the overall delay severity distribution to the high-cost tail in severe weather (H5). Seasonality creates a two-level design where Summer and Winter are statistically equal risky seasons (H2). Delay risk is almost perfectly represented by route efficiency (H3). There are statistically significant delay differences on routes shared by airlines (H4). In the Random Forest modelling, 94% of the delay variance is explained by an integrated model with weather as the leading predictor (H6).

The main methodological addition would be the combination of formally specified, literature-based hypotheses with high-quality inferential testing in an interactive BI system - combining the traditionally distinct approaches of machine learning delay prediction, which is more concerned with accuracy, and operational BI reporting, which is more concerned with interpretability and decision-support value. The ensuing framework is reproducible, open and directly scalable to actual operational information within airline management systems, air navigation service providers or aviation regulatory databases. This framework should be tested in future studies against actual operational data, more fine-grained meteorological measurements (e.g., crosswind speed, visibility distance), longitudinal modelling should be used to test the temporal consistency of the identified seasonal and meteorological patterns, and cross-validated ensemble evaluation should be used to generate conservative accuracy estimates.

REFERENCES

[1] AlBassam, S. A. A., & AlShahrani, D. N. (2025). Flight delay prediction: Evaluating machine learning algorithms for enhanced accuracy. PLOS ONE, 20(12), Article e0335141. https://doi.org/10.1371/journal.pone.0335141

[2] Alfarhood, M., Alotaibi, R., Abdulrahim, B., Einieh, A., Almousa, M., & Alkhanifer, A. (2024). Predicting flight delays with machine learning: A case study from Saudi Arabian Airlines. International Journal of Aerospace Engineering, 2024, Article 3385463. https://doi.org/10.1155/2024/3385463

[3] Bamford, D., & Xystouri, T. (2005). A case study of service failure and recovery within an international airline. Managing Service Quality, 15(3), 306–322. https://doi.org/10.1108/09604520510597814

[4] Barnhart, C., Belobaba, P., & Odoni, A. R. (2003). Applications of operations research in the air transport industry. Transportation Science, 37(4), 368–391. https://doi.org/10.1287/trsc.37.4.368.23276

[5] Barney, J. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99–120. https://doi.org/10.1177/014920639101700108

[6] Barros, C. P., & Wanke, P. (2015). An analysis of African airlines efficiency with two-stage TOPSIS and neural networks. Journal of Air Transport Management, 44–45, 90–102. https://doi.org/10.1016/j.jairtraman.2015.03.002

[7] Bisandu, D. B., & Moulitsas, I. (2023). A deep BiLSTM machine learning method for flight delay prediction classification. Journal of Aviation/Aerospace Education & Research, 32(2), 1–32. https://doi.org/10.58940/2329-258X.1992

[8] Cheng, B. L., Gan, C. C., Imrie, B. C., & Penman, S. (2008). Service recovery, customer satisfaction and customer loyalty: Evidence from Malaysia's hotel industry. International Journal of Hospitality Management, 3(1), 90–106.

[9] Cook, A., & Tanner, G. (2015). European airline delay cost reference values: Updated and extended values (Version 4.1). Performance Review Unit, Eurocontrol.

[10] Cumming, G. (2014). The new statistics: Why and how. Psychological Science, 25(1), 7–29. https://doi.org/10.1177/0956797613504966

[11] Dai, M. (2024). A hybrid machine learning-based model for predicting flight delay through aviation big data. Scientific Reports, 14, Article 4603. https://doi.org/10.1038/s41598-024-55217-z

[12] Eckerson, W. W. (2010). Performance dashboards: Measuring, monitoring, and managing your business (2nd ed.). John Wiley & Sons.

[13] Erlang, A. K. (1909). The theory of probabilities and telephone conversations. Nyt Tidsskrift for Matematik, 20(B), 33–39.

[14] Fleurquin, P., Ramasco, J. J., & Eguiluz, V. M. (2013). Systemic delay propagation in the US airport network. Scientific Reports, 3, Article 1159. https://doi.org/10.1038/srep01159

[15] Goldratt, E. M. (1984). The goal: A process of ongoing improvement. North River Press.

[16] Güvercin, M., Ferhatosmanoğlu, N., & Gedik, B. (2021). Forecasting flight delays using clustered models based on airport networks. IEEE Transactions on Intelligent Transportation Systems, 22(5), 3179–3189. https://doi.org/10.1109/TITS.2020.2990960

[17] Hatıpoğlu, İ., & Tosun, Ö. (2024). Predictive modeling of flight delays at an airport using machine learning methods. Applied Sciences, 14(13), Article 5472. https://doi.org/10.3390/app14135472

[18] Hatıpoğlu, İ., Tosun, Ö., & Tosun, N. (2022). Flight delay prediction based with machine learning. LogForum, 18(1), 97–107. https://doi.org/10.17270/J.LOG.2022.655

[19] Heider, F. (1958). The psychology of interpersonal relations. John Wiley & Sons.

[20] Hopp, W. J., & Spearman, M. L. (2001). Factory physics: Foundations of manufacturing management (2nd ed.). Irwin/McGraw-Hill.

[21] International Air Transport Association. (2024). Annual review 2024. https://www.iata.org/contentassets/c81222d96c9a4e0bb4ff6ced0126f0bb/iata-annual-review-2024.pdf

[22] Keen, P. G. W., & Scott Morton, M. S. (1978). Decision support systems: An organizational perspective. Addison-Wesley.

[23] Li, Q., Guan, X., & Liu, J. (2023). A CNN-LSTM framework for flight delay prediction. Expert Systems with Applications, 227, Article 120287. https://doi.org/10.1016/j.eswa.2023.120287

[24] Li, Q., & Jing, R. (2022). Flight delay prediction from spatial and temporal perspective. Expert Systems with Applications, 205, Article 117662. https://doi.org/10.1016/j.eswa.2022.117662

[25] Mayer, C., & Sinai, T. (2003). Network effects, congestion externalities, and air traffic delays: Or why not all delays are evil. American Economic Review, 93(4), 1194–1215. https://doi.org/10.1257/000282803769206269

[26] Merkert, R., & Morrell, P. S. (2012). Mergers and acquisitions in aviation: Management and economic perspectives on the size of airlines. Transportation Research Part E: Logistics and Transportation Review, 48(4), 853–862. https://doi.org/10.1016/j.tre.2012.02.002

[27] Nguyen, C.-V. (2024). Air transport resilience, tourism and its impact on economic growth. Economies, 12(9), Article 236. https://doi.org/10.3390/economies12090236

[28] Penrose, E. T. (1959). The theory of the growth of the firm. Oxford University Press.

[29] Pyrgiotis, N., Malone, K. M., & Odoni, A. (2013). Modelling delay propagation within an airport network. Transportation Research Part C: Emerging Technologies, 27, 60–75. https://doi.org/10.1016/j.trc.2011.09.003

[30] Qu, J., Wu, S., & Zhang, J. (2023). Flight delay propagation prediction based on deep learning. Mathematics, 11(3), Article 494. https://doi.org/10.3390/math11030494

[31] Rupp, N. G., & Holmes, G. M. (2006). An investigation into airline on-time performance. Applied Economics Letters, 13(3), 163–166. https://doi.org/10.1080/13504850500392635

[32] Sharda, R., Delen, D., & Turban, E. (2020). Analytics, data science, & artificial intelligence: Systems for decision support (11th ed.). Pearson.

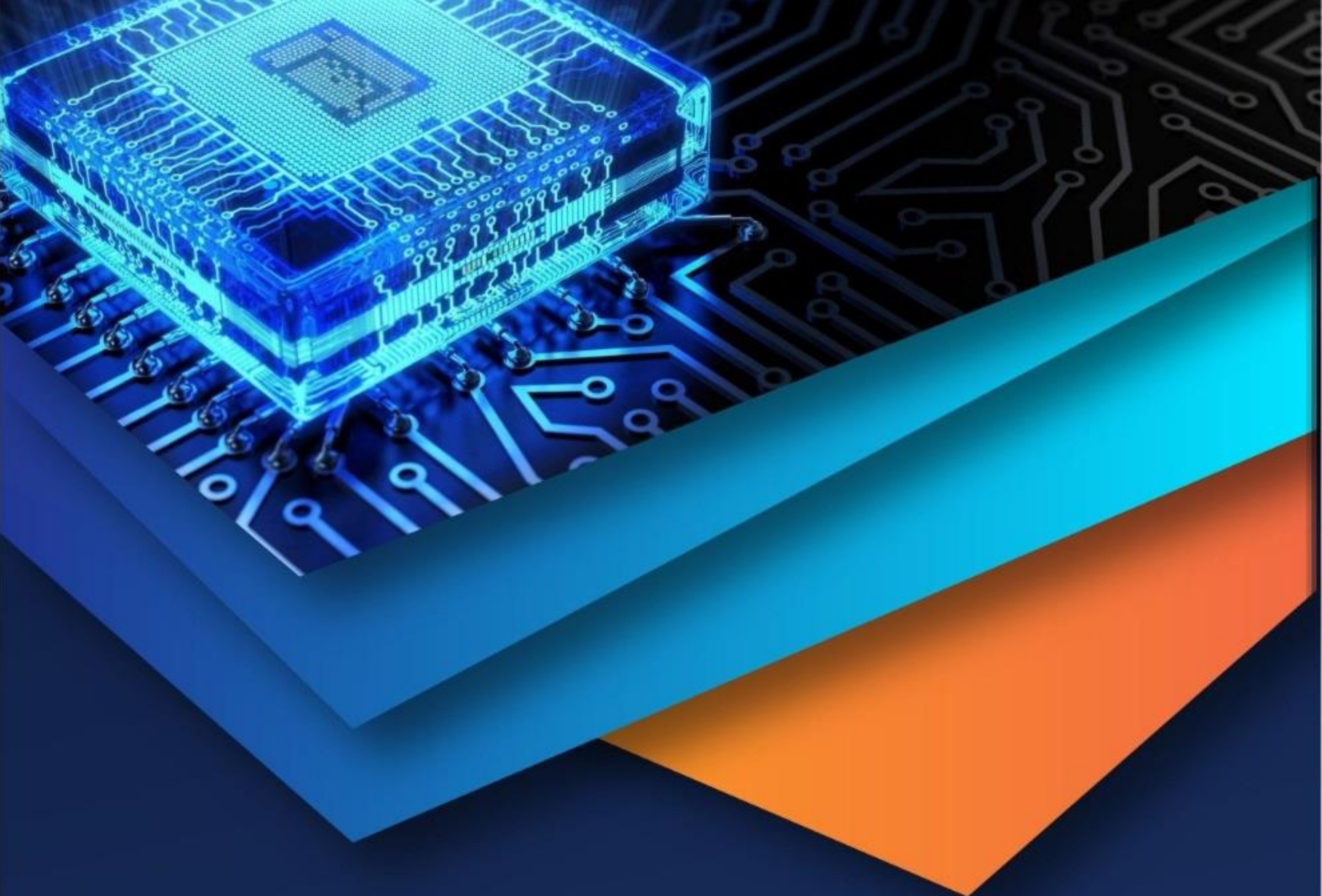
[33] Simon, H. A. (1960). The new science of management decision. Harper & Row.

[34] Song, C., Ma, X., Ardizzone, C., & Zhuang, J. (2024). The adverse impact of flight delays on passenger satisfaction: An innovative prediction model utilizing wide & deep learning. Journal of Air Transport Management, 114, Article 102511. https://doi.org/10.1016/j.jairtraman.2023.102511

[35] Wandelt, S., Chen, X., & Sun, X. (2025). Flight delay prediction: A dissecting review of recent studies using machine learning. IEEE Transactions on Intelligent Transportation Systems, 26(4), 4283–4297. https://doi.org/10.1109/TITS.2025.3528536

[36] Wang, F., Bi, J., Xie, D., & Zhao, X. (2022). Flight delay forecasting and analysis of direct and indirect factors. IET Intelligent Transport Systems, 16(7), 890–907. https://doi.org/10.1049/itr2.12183

[37] Weiner, B. (1985). An attributional theory of achievement motivation and emotion. Psychological Review, 92(4), 548–573. https://doi.org/10.1037/0033-295X.92.4.548.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY