Investigating the Causes of Delayed Arrivals at Indira Gandhi International Airport

Manan Munjal

Tulsi Public School, Ambala City, Haryana, India

Abstract - *This study investigates the cause of delays in arriving at Indira Gandhi International Airport (DEL) by probing the possible connection between environmental conditions and flight punctuality and contrasting the findings with previous literature on reactionary and infrastructural impacts. The flights were pulled from FlightRadar24 and combined with climatic information from timeanddate.com to plot the Hong Kong–Delhi route (HKG–DEL) for 30 days. The study uses logistic regression to determine the effect of environmental conditions, such as wind speed, wind direction, temperature, humidity, and atmospheric pressure, on delay. While the overall models were significant statistically, the individual environmental conditions were not determined to have statistically significant correlations with delays. This contrasts with earlier studies on airport operational performance—using methods like Data Envelopment Analysis (DEA)—that had determined reactionary delays (e.g., late arrivals and crew issues) and infrastructural bottlenecks as the main cause of delay patterns, explaining up to 76% of the variation in operational performance. Our study illustrates operational and infrastructural factors, as opposed to environmental conditions, to be largely accountable for DEL flight delays. The findings have major implications for resource allocation and strategic planning in air transport operations, such that investment should be channeled into optimizing operational efficiency rather than into environmental forecasting systems.*

Key Words: flight delay, operational factors, logistic regression, HKG–DEL route, airport efficiency, reactionary delays

1.INTRODUCTION

Flight delays are a recurring issue that affect airline performance, passenger satisfaction, and overall airport productivity in busy airports like Indira Gandhi International Airport (DEL). Delays in such airports occur through a complex interplay of infrastructural, operational, and physical elements. Environmental factors are most often mentioned as potential causes of flight delays, but their true contributions relative to operational factors are not well quantified in the literature.

Recent research has proven that reactionary causes, such as late departures and crew scheduling issues, are major causes of flight delays. For instance, studies on airport operational effectiveness at DEL have concluded that reactionary delays and gate constraint contribute to around 76% of variation in operational performance. Despite such findings that prove the significance of operational causes, there has been a problem with taking into account the potential effect of environmental causes on flight punctuality.

Flight delays are disruptive to all the parties involved in the air transport industry. Airlines are faced with routing and crew scheduling disruptions, which can cause cascading delays across their network. Passengers face increased waiting times at airports, which impact satisfaction and can cause connections to be missed. Ground handling personnel face workloads as the aircraft arrivals cluster outside their planned slot allocations. Limited gate availability causes some aircraft to be located at remote bays, which raises passenger handling complexity and transit times in the airport.

Our aim in this research is to make an empirical investigation into flight delay figures alongside meteorological conditions and identify their individual contributions to the causation of delays. We take the HKG–DEL route as the focus of our study, as it has a uniform daily and weekly frequency pattern and explore if variations in meteorological variables can be statistically correlated with delay events. In comparison with past research on operational factors, our objective is to enhance the common understanding of causation of delays, offering meaningful insights into the allocation of resources and strategic management in the airline industry.

2. Materials and Methods

2.1 Data Collection

Data was collected from two sources:

1. **Flight Data**: Real-time flight tracking information was obtained from FlightRadar24, where arrival, departure, and in-flight times were obtained for the HKG–DEL flight. The information obtained was the actual arrival times, scheduled arrival times, and delay status, which was defined as arriving 5 minutes or more behind schedule.

2. **Climatic Data**: Historical weather data for Delhi was obtained from timeanddate.com. Parameters measured were: a. Wind speed (km/h) b. Wind direction (degrees) c. Temperature (°C) d. Relative humidity (%) e. Atmospheric pressure (mbar)

Environmental information was synchronized with flight arrival times using the nearest available time slot (usually 2-hour intervals) to enable synchronization between flight performance and current weather conditions.

2.2 Route and Parameter Selection

The chosen route (Hong Kong to New Delhi) featured a routine with a fixed schedule—four flights every day with one thrice-weekly flight—to ensure a representative sample over a 30-day period. A single route was studied to minimize variations from external factors such as different airport infrastructure, geographic/environmental conditions.

This route was chosen specifically for the following advantages:

1. Consistency of geographical path, minimizing variation of en-route conditions
2. Similar infrastructure at both origin and destination airports
3. Diversity of airlines operating and aircraft types:
   - Air India (thrice weekly) with Boeing 787 Dreamliner
   - Vistara (daily) with Airbus A321neo
   - IndiGo (daily) with Airbus A321neo
   - Cathay Pacific (twice daily) with Boeing 777-300ER and Airbus A350-900

This variety allowed for analysis of both environmental factors and airline/aircraft-specific performance differences while maintaining route consistency.

Table -1: Flight Data Characteristics

Characteristic	Value
Route	Hong Kong (HKG) to Delhi (DEL)
Study Period	30 days
Total Flights Analysed	62
Number of Airlines	4 (Air India, Vistara, IndiGo, Cathay Pacific)
Aircraft Types	Boeing 787 Dreamliner, Airbus A321neo, Boeing 777-300ER, Airbus A350-900
Flight Frequency	4 flights daily + 1 thrice-weekly flight

Table -2: Environmental Variables Summary

Variable	Mean	Range	Standard Deviation
Temperature (°C)	33.2	27.0 - 35.0	2.7
Wind Speed (km/h)	6.3	0.0 - 15.0	3.9
Humidity (%)	69.8	60.0 - 89.0	9.5
Pressure (mbar)	1001.2	998.0 - 1004.0	2.1
Flight Distance (km)	3756.4	3753.0 - 3762.0	2.6

Table -3: Delay Statistics

Delay Category	Count	Percentage
On-time (< 5 minutes)	40	64.5%
Delayed (≥ 5 minutes)	22	35.5%
Total	**62**	**100%**

2.3.1 Logistic Regression Analysis

In order to estimate the effect of each of the variables on the probability of flight delay (arrivals 5 or more minutes past their reported time), logistic regression was utilized. Logistic regression was the statistical method of choice precisely because it estimates the probability of a dichotomous outcome (delayed/not delayed) as a function of several independent variables.

Two primary methods were utilized:

1. Rank-Based: Using arrival time rank as a continuous variable
2. Binary Time Slots: Categorizing arrival times into binary classifications

Independent variables included in both models:

- Wind speed (km/h)
- Temperature (°C)
- Humidity (%)
- Atmospheric pressure (mbar)
- Actual flight distance (km)
- Arrival time classification (either ranked or binary)

Model performance was evaluated using chi-square statistics, p-values, odds ratios, and overall classification accuracy.

2.3.2 Comparative Analysis

Along with primary data analysis, the background was established by reading previous studies. Literature based on

Data Envelopment Analysis (DEA) provided information on the effects of external occurrences (e.g., reactionary delays) and internal functioning on airport efficiency. These studies provided benchmarks against which our findings were to be interpreted and understood.

3. Impact of Individual Variables

3.1.1 Scheduled Arrival Time Classification

The rank-based method produced a negative coefficient for scheduled arrival time rank (b = -5.72), indicating that certain time slots have systematically higher delay odds than others. However, this effect was not statistically significant (p = 0.147).

The binary (scheduled) arrival time method gave a positive coefficient (b = 22.34) with an extraordinarily high odds ratio (5,047,240,828.87), indicating that flights arriving during certain time periods have substantially higher odds of being delayed. Despite the high odds ratio, this variable was not statistically significant (p = 0.998).

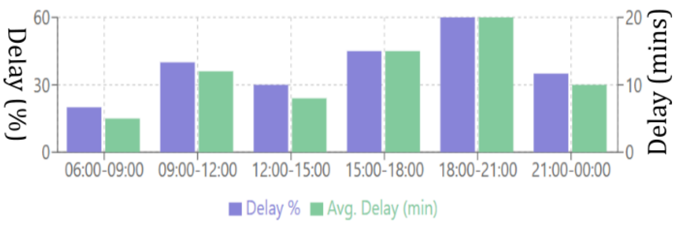

Chart -1: Flight Delay Distribution by Arrival Time

3.1.2 Wind Speed

The coefficients for wind speed were positive across both methods:

- Rank-based: b = 0.52, odds ratio = 1.68, p = 0.123
- Binary: b = 0.16, odds ratio = 1.17, p = 0.224

While these results suggest that higher wind speeds might be associated with increased delay probability, the p-values indicate that these associations did not reach statistical significance.

3.1.3 Temperature and Humidity

Temperature showed a positive association with delay probability in the rank-based model (b = 0.67, odds ratio = 1.96, p = 0.397), while in the binary model, the effect was neutralized (b = 0, odds ratio = 1, p = 1.0). Neither result was statistically significant.

Humidity demonstrated positive coefficients in both models:

- Rank-based model: b = 0.67, odds ratio = 1.95, p = 0.191
- Binary model: b = 0.07, odds ratio = 1.08, p = 0.645

Again, these associations failed to reach statistical significance.

3.1.4 Atmospheric Pressure

Atmospheric pressure showed positive coefficients in both models:

- Rank-based model: b = 1.46, odds ratio = 4.33, p = 0.174
- Binary model: b = 0.34, odds ratio = 1.4, p = 0.487

Despite relatively large odds ratios, particularly in the rank-based model, these relationships were not statistically significant.

3.1.5 Flight Distance

Actual flight distance showed opposing effects in the two models:

- Rank-based model: b = -0.69, odds ratio = 0.5, p = 0.34 (negative association)
- Binary model: b = 0.08, odds ratio = 1.08, p = 0.662 (positive association)

This inconsistency and the non-significant p-values suggest that flight distance does not have a reliable relationship with delays in our sample.

Table -4: Rank Based Method Results

Variable	Coefficient	Odds Ratio	p-value
Scheduled Arrival Time	-5.72	0.00	0.147
Wind Speed (km/h)	0.52	1.68	0.123
Temperature (°C)	0.67	1.96	0.397
Humidity (%)	0.67	1.95	0.191
Atmospheric Pressure (mbar)	1.46	4.33	0.174
Flight Distance (km)	-0.69	0.50	0.340

Table -5: Logistic Regression Results

Variable	Coefficient	Odds Ratio	p-value
Scheduled Arrival Time	22.34	5,047,240,828.87	0.998
Wind Speed (km/h)	0.16	1.17	0.224
Temperature	0.00	1.00	1.000

(°C)			
Humidity (%)	0.07	1.08	0.645
Atmospheric Pressure (mbar)	0.34	1.40	0.487
Flight Distance (km)	0.08	1.08	0.662

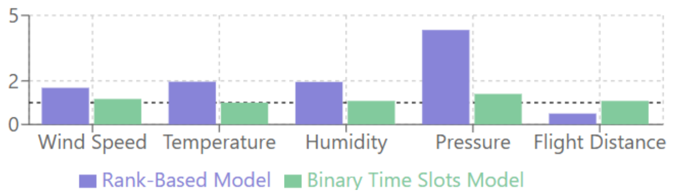

Chart -2: Odds Ratios of Environmental Factors for Flight Delays

3.2 Comparison with Previous Research

The lack of statistically significant correlations between environmental variables and flight delays contrasts with established research on operational variables. Reactionary delays and gate delays have a significant effect on airport throughput, accounting for 76% of operational performance variation at DEL, as per Sumangal Narendra's research. This indicates that operational limitations, and not environmental variables, are the main causes of delay patterns.

Another study applicable to our setting estimated economic impacts of disruptions by fog at DEL to be roughly $3.9 million in 2011-2016. Despite our research window not falling in the winter fog period of Delhi, the absence of strong correlations among humidity, temperature, and delay indicates that other weather difficulties other than fog could be smaller than assumed for the year.

The paper "Flight Delay Analysis and Potential Improvements with Big Data" indicated that Delhi airport experienced fewer delays when the number of monthly flights was less than 9,600, and delays rose with growing traffic volume. This capacity-sensitive pattern suggests that traffic density and operational constraints are more significant predictors of delay than the environmental factors examined in our research.

Airport efficiency studies using the Data Envelopment Analysis approach have determined the impact of external shocks on airport performance, thus highlighting the importance of non-environmental determinants in evaluating operational efficiency.

Table -6: Comparison of Current Study with Previous Research

Study	Focus	Key Findings	Variance Explained
Current Study (2025)	Environmental factors (temperature, wind speed, humidity, pressure) on HKG-DEL route	No significant correlation between environmental factors and delays	Limited predictive power
Narendra et al. (2016)	Operational variables at DEL	Reactionary delays and gate constraints	76% of operational performance variation
Mohan & Singh (2019)	Economic impact of fog at DEL	Estimated $3.9 million in economic losses (2011-2016) due to fog	Not reported
Vane (2016)	Traffic volume at DEL	Delays increase when monthly flights exceed 9,600	Not reported
Sood & Bhushan (2004)	Airport efficiency across India	External shocks significantly impact operational efficiency	Varied by airport

4. CONCLUSIONS

This study suggests that logistic regression models incorporating environmental and operational factors show some promise in predicting flight delays. However, forecasted environmental data throughout the route must be incorporated for these techniques to be of use, and the accuracy of results will be a function of the accuracy of input forecasts. Additionally, individual environmental factors—including wind speed, temperature, humidity, and atmospheric pressure—do not show statistically significant associations with delay instances.

Key conclusions include:

1. Despite positive coefficients indicating possible relationships between environmental variables and arrival delays, none of these associations crossed thresholds of statistical significance.

2. The contrast between our findings and previous research highlighting the importance of operational factors suggests that reactionary delays, infrastructure constraints, and traffic density play a more dominant role in determining flight punctuality rather than environmental conditions.

3. The inconsistent and non-significant results across models for variables like flight distance and arrival time classification further highlight the complex, multivariate nature of flight delays.

4. The high overall predictive accuracy of our models despite the lack of significant individual predictors might suggest that the combination of factors, rather than any single variable, contributes to delay patterns. Further work on this front is called for

These results have significant strategic and resource planning implications for air transport operations. Instead of investing huge amounts of resources into environmental forecasting tools, airports and airlines would derive more on-time performance benefits by the concentration of efforts on traffic management schemes, investment in infrastructure, and gains in operational efficiency.

Subsequent research should have a larger sample size, include more operational variables, and account for seasonal variation to give a better picture of factors causing delays during the course of the year. Particular consideration of DEL's winter fog season would enhance the results of this research and allow for the quantification of the relative impact of severe weather events in relation to the normal day-to-day operational factors for causing flight delays.

ACKNOWLEDGEMENT

The author acknowledges the data collected from FlightRadar24 and timeanddate.com. Furthermore, findings of previous studies on airport efficiency and flight delay analysis have immensely contributed to this research. Gratitude is also due to Sumangal Narendra and other researchers whose publications on airport operation performance provided background information for interpreting our findings.

REFERENCES

[1] Sumangal Narendra et al., "Airport Operational Performance and Its Influence on Schedule Adherence: A Data-driven Analysis of Delhi International Airport," Journal of Airport Management, vol. 10, no. 4, pp. 328-342, 2016.M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[2] Mohan, R. and Singh, A.K., "Economic Impact of Fog at Delhi International Airport: Direct and Indirect Cost Analysis," Atmosphere, vol. 10, no. 4, article 198, 2019.K. Elissa, "Title of paper if known," unpublished.

[3] Rashmi Vane, "Flight Delay Analysis and Possible Enhancements with Big Data," International Research Journal of Engineering and Technology (IRJET), vol. 3, no. 6, pp. 1388-1392, 2016.

[4] Sood, N. and Bhushan, B., "Dynamic Evaluation of Airport Efficiency in India: A Data Envelopment Analysis Approach," Economics Working Papers, University of Munich, Paper no. 0412010, 2004.

[5] International Air Transport Association (IATA), "Worldwide Slot Guidelines," 9th Edition, 2019.

[6] Federal Aviation Administration, "Weather-Related Aviation Accident Study 2003-2007," 2010.

[7] Kulkarni, R., "Disruptive Weather and Its Impact on Air Traffic: A Comprehensive Review," Journal of Air Transport Management, vol. 78, pp. 69-87, 2019.