論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	A Multifaceted Data-driven Analysis of Flight Departure Delays in the Japanese Domestic Aviation Market
著者(和文)	SADEEKSOUMIK NAFIS
Author(English)	Soumik Sadeek
出典(和文)	学位:博士(学術), 学位授与機関:東京科学大学, 報告番号:甲第512号, 授与年月日:2025年9月22日, 学位の種別:課程博士, 審査員:花岡 伸也,山下 幸彦,髙田 潤一,室町 泰徳,髙山 雄貴,杉下 佳辰
Citation(English)	Degree:Doctor (Academic), Conferring organization: Institute of Science Tokyo, Report number:甲第512号, Conferred date:2025/9/22, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

（博士課程）
Doctoral Program

論 文 要 旨

THESIS SUMMARY

系・コース： Department of, Graduate major in	融合理工学 地球環境共創	系 コース	申請学位（専攻分野）： Academic Degree Requested	博士 Doctor of （ Philosophy ）
学生氏名： Student's Name	Soumik Nafis Sadeek		審査員主査： Chief Examiner	Shinya Hanaoka

要旨（英文800語程度）
Thesis Summary (approx.800 English Words)

Flight departure delays pose significant operational and economic challenges in the aviation industry. In Japan's domestic market where two full-service carriers: All Nippon Airways (ANA) and Japan Airlines (JAL) dominate (more than 70% domestic flights), these challenges are compounded by dense networks, tightly coordinated schedules, and the need for punctuality. This thesis presents a multifaceted, data-driven analysis of departure delays by integrating network science, machine learning, conformal prediction, and topological data analysis (TDA). Using flight data from the Official Aviation Guide (OAG) between 2018 and 2021, the study focuses on the temporal, structural, and predictive dimensions of delay propagation at major Japanese airports.

The dataset includes all ANA and JAL domestic flights across 36 and 47 airports, respectively including eight key hub airports - Haneda (HND), Narita (NRT), Itami (ITM), Kansai (KIX), Chubu Centrair (NGO), Fukuoka (FUK), New Chitose (CTS), and Naha (OKA). Each day's data is used to build a dynamic delay network where airports are nodes and flights with delays of at least one minute form the edges. These networks serve as the foundation for the three chapters.

Chapter 3 investigates the association between network properties and departure delay duration. Using Prais-Winsten regression with fixed effects and panel-corrected standard errors, the study models delay duration accounting for serial correlation and heteroskedasticity in the panel data. Four network metrics are examined: in-degree, betweenness centrality (BC), eigenvector centrality (EC), and transitivity. Results show that in-degree centrality is consistently associated with longer delays suggesting that airports receiving more delayed flights are more vulnerable to accumulating further delays. BC has mixed effects-hubs may absorb delays more efficiently while high-BC spoke airports can act as bridges for delay propagation. EC generally correlates with shorter delays likely due to stronger integration within well-performing subnetworks. Transitivity exhibits differing effects across carriers: it often reduces delays for ANA but may exacerbate clustered delays for JAL. In 2020, despite fewer flights, some airports experienced longer delays highlighting how reduced resources and procedural constraints can still induce inefficiency. The findings suggest that network metrics can support proactive delay management to implement trajectory-based air traffic operations.

Chapter 4 addresses the need for uncertainty quantification in delay forecasting. While most studies focus on point predictions, this chapter introduces conformal prediction (CP) to produce prediction intervals for average daily departure delays. Three adaptive CP methods are tested: Adaptive Conformal Inference (ACI), Aggregate Adaptive Conformal Inference (AgACI) and Fully Adaptive Conformal Inference (FACI). These models are built using features such as lagged delay, network metrics, seasonal variables, and state-of-emergency indicators. Random Forest algorithm is used for feature importance consistently identifying lagged delay as the most important predictor underscoring the temporal dependence of delay patterns. At high coverage levels (80%), all models struggle to widen prediction intervals fast enough during delay surges leading to under-coverage. AgACI performs best at moderate coverage levels (50%) by aggregating learning rates and adjusting interval widths adaptively. FACI achieves a balance between precision and responsiveness through expert reweighting based on pinball loss. These findings allow

stakeholders to choose forecasting tools that match operational needs: ANA's strategy may prioritize wider intervals for reliability while JAL's narrower intervals emphasize precision. Thus, CP offers a flexible, interpretable, and adaptive framework for delay forecasting under uncertainty.

Chapter 5 introduces Topological Data Analysis (TDA) to capture higher-order patterns of delay propagation not visible through conventional network metrics. The study constructs daily undirected, weighted delay networks and applies Vietoris-Rips complexes to analyze two types of topological features: H_0 (connected components) and H_1 (delay loops). Before COVID-19, ANA's networks exhibited more persistent and broader delay loops whereas JAL's were more fragmented. During the pandemic, delay connectivity collapsed for both carriers as service was suspended leading to sharp drops in connected components and feedback loops. By 2021, particularly in winter, high-persistence loops (lasting ≥20 minutes) reemerged especially for ANA. These loops were more triangular in ANA's case while JAL featured more complex four- and five-airport cycles. Airports such as HND, OKA, NGO, FUK, and NRT consistently appeared in high-persistence loops identifying them as critical nodes for delay control. Persistence distributions also showed seasonal effects, with longer-lived loops more common in winter. Therefore, TDA provides a powerful lens to detect persistent delay structures, offering operational insights for preemptive buffer allocation and network-wide interventions.

In conclusion, this thesis provides a multidimensional perspective on departure delays in Japan's domestic aviation market. Chapter 3 highlights how network positioning influences delay duration. Chapter 4 introduces adaptive methods for quantifying forecast uncertainty using conformal prediction. Chapter 5 reveals hidden loop/cycle structures using topological tools. Together, these contributions offer practical strategies for airlines and air traffic managers to control delays, allocate resources efficiently, and respond dynamically to operational disruptions. The novel integration of CP and TDA into aviation analytics also opens new pathways for delay management in the air transport systems.