A DATA-DRIVEN METHODOLOGY TO ANALYZE AIR TRAFFIC MANAGEMENT SYSTEM OPERATIONS WITHIN THE TERMINAL AIRSPACE

A Dissertation
Presented to
The Academic Faculty

By

Samantha J. Corrado

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Aerospace Engineering
Department of Aerospace Engineering

Georgia Institute of Technology

December 2021

A DATA-DRIVEN METHODOLOGY TO ANALYZE AIR TRAFFIC MANAGEMENT SYSTEM OPERATIONS WITHIN THE TERMINAL AIRSPACE

Thesis committee:

Prof. Dimitri N. Mavris
School of Aerospace Engineering
Georgia Institute of Technology

Prof. Daniel Schrage
School of Aerospace Engineering
Georgia Institute of Technology

Prof. Graeme Kennedy
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Tejas G. Puranik
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Xavier Olive
Department for Information Processing and Systems
ONERA, the French Aerospace Lab

Date approved: December 08, 2021

To my family.

ACKNOWLEDGMENTS

As my journey through graduate school concludes, I would like to acknowledge those who made the journey such a great experience. First, I would like to express my deep gratitude to my adviser, Prof. Dimitri Mavris for the opportunity to join his lab and facilitating of my diverse research experiences. I am certain the skills I have learned in your lab will continue to be extremely valuable throughout my career. Additionally, I would like to express my gratitude for the guidance I have received from Dr. Tejas Puranik. I have thoroughly enjoyed our many meetings, and will miss working with you very much. Further, I extend my gratitude to the other members of my thesis committee: Prof. Daniel Schrage, Prof. Graeme Kennedy, and Dr. Xavier Olive for the time taken to review my work and provide valuable insights and feedback.

I am so grateful to have made so many wonderful, lifelong friends at Georgia Tech. Thank you to my friends from my very first days of graduate school: Cynthia, Rebecca, Maya, Meron, and Erol. Thank you to Lea for all of the positive energy you have brought to my experience. And, a huge thank you to my favorite group of people to play board games with: Chloe & Jon, Julie & Colby, and Domitille & Alex.

In addition to lifelong friends, I am so incredibly grateful to have met my lifelong partner, Alex Markov, during my time at Georgia Tech. Thank you so very much for all that you do for me. I cannot wait to be Dr. & Dr. Markov!

Lastly, I would like to express my overwhelming gratitude to my family, whose love, support, and encouragement means the world to me. Thank you to my brother, Ben, for always checking in on me. Thank you to my brother, Matt, who inspired my Aerospace Engineering graduate degrees and with whom I have so many great memories with at Georgia Tech. Finally, thank you to my biggest supporters, my mom and dad, without whom my graduate school journey would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AAM Advanced Air Mobility

ADS-B Automatic Dependent Surveillance-Broadcast

ANSP Air Navigation Service Provider

ARMD Aeronautics Research Mission Directorate

ASM Airspace Management

ASOS Automated Surface Observing System

ASPM Aviation System Performance Metrics

ATC Air Traffic Control

ATFM Air Traffic Flow Management

ATM Air Traffic Management

ATS Air Traffic Services

DBSCAN Density-Based Spatial Clustering of Applications with Noise

ED Euclidean Distance

FAA Federal Aviation Administration

FDM Flight Data Monitoring

FDR Flight Data Recorder

FOQA Flight Operational Quality Assurance

GMM Gaussian Mixture Model

HDBSCAN Hierarchical Density-Based Spatial Clustering of Applications with Noise

ICAO International Civil Aviation Organization

KOAK Oakland International Airport

KSFO San Francisco International Airport

KSJC San Jose International Airport

NAS National Airspace System

NASA National Aeronautics and Space Administration

NextGen Next Generation Air Transportation System

PBN Performance Based Navigation

PCA Principal Component Analysis

SED Standardized Euclidean Distance

SESAR Single European Sky ATM Research

SKE Specific Kinetic Energy

SKER Specific Kinetic Energy Rate

SME Subject Matter Expert

SPE Specific Potential Energy

SPER Specific Potential Energy Rate

SSPD Symmetrized Segment-Path Distance

STE Specific Total Energy

STER Specific Total Energy Rate

SVM support vector machine

t-SNE t-Distributed Stochastic Neighbor Embedding

TBM Time Based Management

TBO Trajectory Based Operations

TMI Traffic Management Initiative

UMAP Uniform Manifold Approximation and Projection

UTM Universal Transverse Mercator

WED Weighted Euclidean Distance

SUMMARY

This research is motivated by the overarching emphasis on increasing Air Traffic Management (ATM) system efficiency and capacity, while maintaining a high level of safety as the current systems undergo a transformation in concept of operations under global ATM system modernization plans. Considering the global modernization efforts' current state and future milestones, gaining a comprehensive understanding of both flight-level and airspace-level operations is required. Though, the complexity of ATM systems restricts the use of classical physics-based methods for the analysis of new operational concepts. However, an increased availability of operational data provides several new analysis opportunities. Specifically, Automatic Dependent Surveillance-Broadcast (ADS-B) data provides the basis for offline data-driven methods to be applied to analyze ATM system operations. Further, the analysis of ATM system arriving aircraft operations has the highest potential to impact system safety, capacity, and efficiency levels. Therefore, this research presents an offline data-driven methodology to be applied to ADS-B data to analyze ATM system arriving aircraft operations at both the flight level and the airspace level. The proposed methodology requires three steps: (i) air traffic flow identification, (ii) anomaly detection, and (iii) airspace-level analysis.

The air traffic flow identification step required a reliable method of clustering trajectories considering the converging nature of trajectories within the terminal airspace, where the trajectory clustering algorithm must be able to: (i) determine the optimal number of clusters within the data set, (ii) identify outliers, and (iii) identify clusters of varying densities. Therefore, the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) clustering algorithm was selected. To improve the existing applications of HDBSCAN, the novel use of the Weighted Euclidean Distance (WED) with HDBSCAN was demonstrated to more reliably identify air traffic flows.

The anomaly detection step first required the novel distinction between spatial anoma-

lies and energy anomalies detected in ADS-B data. Next, the anomaly detection step required the detection of spatial anomalies, the subsequent detection of energy anomalies, and the computation of an anomaly score to provide a measure of the "degree of anomalousness" of a trajectory with respect to its energy metrics. Spatial anomaly detection was performed by the more reliable method of identifying air traffic flows considering **Hypothesis 1**. A clustering algorithm was similarly identified to be most appropriate for energy anomaly detection. However, the requirements of the clustering algorithm differ compared with the requirements for spatial anomaly detection. Namely, it is not necessary for an algorithm to have the capability to identify clusters of varying densities; therefore, Density-Based Spatial Clustering of Applications with Noise (DBSCAN) was selected to perform energy anomaly detection. Once spatial and energy anomalies have been detected, it is relevant to statistically investigate whether: (i) if trajectories detected as spatial anomalies are more likely to be detected as energy anomalies, (ii) when considering only energy-nominal trajectories, if spatially anomalous trajectories are "relatively more anomalous" than spatially nominal trajectories, and (iii) when considering only energy-anomalous trajectories, if there are distinct differences in the statistical properties of a quantitative measure of the "degree of anomalousness" of trajectories that are spatially anomalous versus spatially nominal.

Finally, related to the development of the outline of the airspace-level analysis step in the proposed methodology required a method to identify operational patterns, characterize whether an operational state is nominal, anomalous, or transition, and predict operational patterns. It was determined that a clustering algorithm capable of identifying outliers and requiring no a priori specification of the number of clusters is most appropriate to simultaneously identify operational patterns and characterize operational states. Due to meeting the aforementioned criteria and considering the effectiveness of DBSCAN in other aviation applications, DBSCAN was selected. However, the presence of transitional operational states within the data set may result in skewed results. Therefore, a recursive DBSCAN

procedure was proposed such that transitional operational states are removed from the data set and the clustering repeats until no new transitional operational states are characterized. Considering the need for the capability to predict operational patterns and that the few existing methods related to prediction of what may be considered operational patterns have leveraged features derived from recorded weather measurements, a set of classification algorithms is trained to predict operational pattern provided derived weather metric features. An artificial neural network, gradient-boosted decision tree (XGBoost), and support vector machine (SVM) were trained, where XGBoost was found to provide the best prediction accuracy.

CHAPTER 1
BACKGROUND AND MOTIVATION

This research is motivated by an overarching emphasis on increasing Air Traffic Management (ATM) system efficiency and capacity, while maintaining a high level of safety as current systems undergo a transformation in concept of operations. An overview of the anticipated trends in air traffic volume is presented. Next, a description of the functions of ATM systems is provided. Additionally, a review of the global modernization efforts that place the emphasis on increasing ATM system safety, capacity, and efficiency is provided. Finally, the operations of aviation systems in the context of the global modernization efforts are described.

1.1 Air Traffic Volume

Aviation systems are considered fundamental for the development of modern societies as they have become enablers of global business [1]. While, currently, the aviation industry is emerging from the most severe aviation crisis in its history: the COVID-19 pandemic, the industry had previously been experiencing large increases in the volume of operations. Figure 1.1 displays the decline in number of passengers carried since 1945, providing context related to the declines experienced as a result of other aviation and world crises [2]. An unprecedented decline is observed as COVID-19 is declared a pandemic in March of 2020 [2]. The pandemic resulted in a 60% reduction in total world passengers in 2020 compared to the same time period in 2019 and an approximate 371 billion USD loss of gross passenger operating revenues of airlines [2]. Thus far, in 2021, a 49-50% decline in world total passengers has been observed [2]. Though, despite other crises in history, the aviation industry has experienced recovery and subsequent growth.

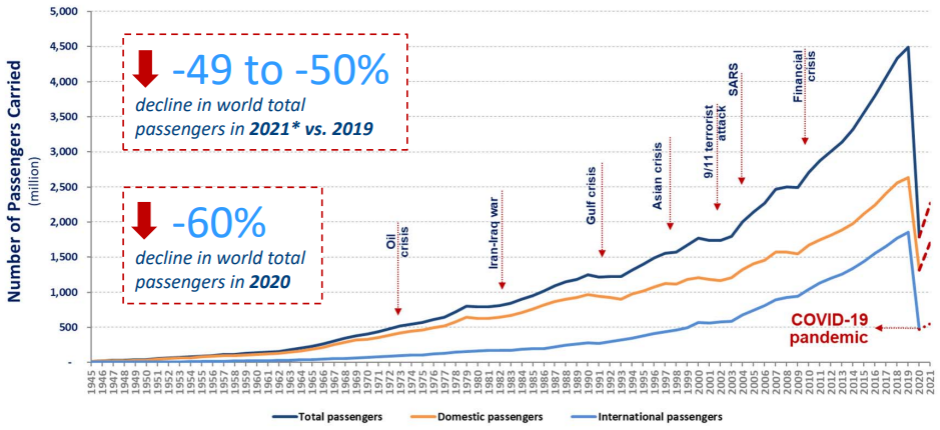

Figure 1.1: World Passenger Traffic Evolution, 1945-2021 [2]

Air traffic demand is driven by economic activity, where the growing U.S. and world economies provide the foundation for long term growth [3]. Further, long-term drivers for expanded air traffic demand remain in existence as populous emerging markets are still anticipated to desire more air services [4]. For instance, the global economic growth provides support for growing air cargo demand, with current levels above that observed in 2019 [5]. While recovery in air traffic demand has been slow in 2021 due to international travel restrictions, there is an expectation that vaccines will enable governments to relax restrictions, which supports global travel reaching 61% of 2019 levels in 2022 [6]. It is anticipated that domestic travel will continue to be strong [5]. Figure 1.2 displays the recovery growth that is expected to occur, including both international and domestic, through quarter four in 2021 [5]. According to Figure 1.2, by the end of 2022 it is expected that domestic travel demand will rise above pre-COVID-19 levels.

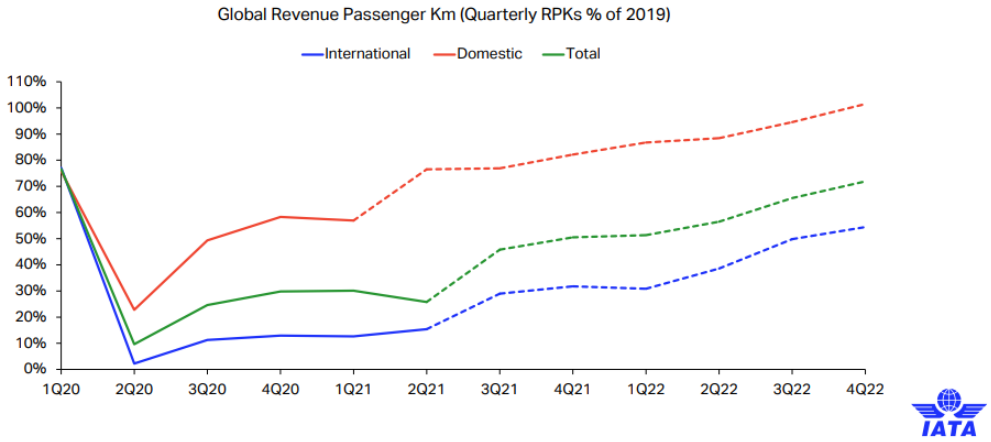

Figure 1.2: Return to Growth Through Quarter Four of 2021 [4]

While much is still uncertain regarding the impacts of the pandemic as the impacts depend on variables such as duration and magnitude of the outbreak, containment measures, the degree of consumer confidence for air travel, economic conditions, etc., there is high confidence air traffic growth will ultimately resume [2]. Further, the Federal Aviation Administration (FAA) had predicts U.S. airline passenger volume would increase by 400 million passengers to carry just under 1.3 billion passengers in 2038 [7]. The FAA predicts even faster growth for U.S. airline cargo traffic [7]. In contrast to the sharp decrease in passenger air travel, a surge in cargo flights was observed in 2020, with an increase in cargo-only operations utilizing passenger aircraft [2], as displayed in Figure 1.3.

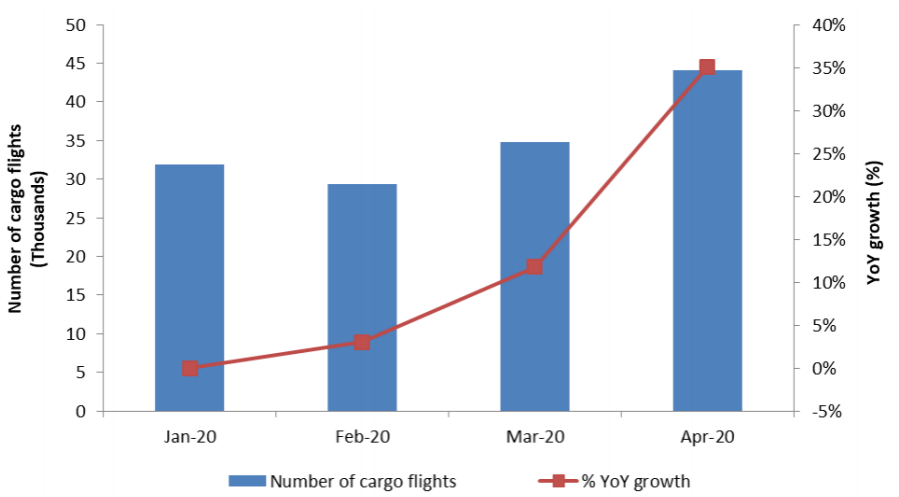

Figure 1.3: Cargo Flights in Early 2020 Months [2]

As the economy recovers from the impact of the COVID-19 pandemic, the economic growth coupled with technological advances, market liberalization, oil prices, and other trends will continue to affect worldwide commercial aviation [8]. For instance, the FAA projects that system traffic in revenue passenger miles will increase by 5.5% a year between 2021 and 2041 [3]. Further, the FAA anticipates robust travel demand growth between 2022 and 2026 due to U.S. economic recovery, which has the potential increase controller workload; thus, operations at FAA and contract towers are projected to grow 1.9% a year between 2021 and 2041 [3]. Considering the historical growth in air traffic volume, and in preparation for forecast growth and an increase in complexity of aviation systems, changes to the existing concept of operations are necessary.

1.2 Air Traffic Management Systems

ATM systems are tasked with managing the operations of some of the world's most complex system-of-systems as they are responsible for the operations of all aircraft. ATM systems consist of numerous interacting systems and subsystems that are further complicated by human-in-the-loop decisions. The vast complexity of ATM systems will continue to increase as air traffic volume increases and Advanced Air Mobility (AAM) concepts are

introduced in the airspace.

Formally, ATM systems are those generally responsible for managing aviation airspace operations, where an ATM system is defined by International Civil Aviation Organization (ICAO) as "*a system that provides ATM through the collaborative integration of humans, information, technology, facilities and services, supported by air and ground- and/or space-based communications, navigation and surveillance*" [9]. The term ATM is defined as "*the dynamic, integrated management of air traffic and airspace including Air Traffic Services (ATS), Airspace Management (ASM) and Air Traffic Flow Management (ATFM) - safely, economically and efficiently - through the provision of facilities and seamless services in collaboration with all parties and involving airborne and ground-based functions*" [9]. Figure 1.4 provides a visual breakdown of the primary services within the ATM system.

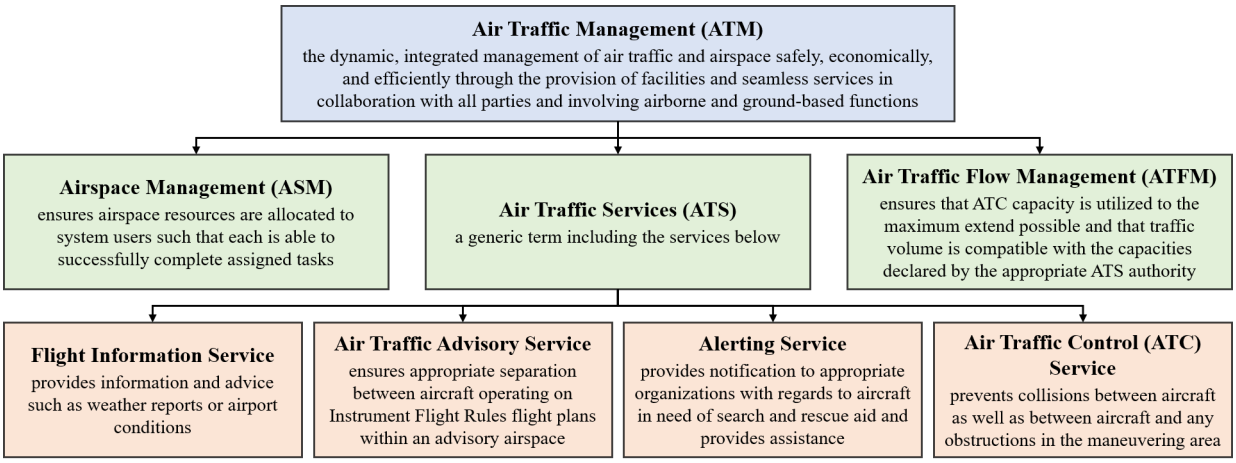

Figure 1.4: Air Traffic Management Services Breakdown [9]

The three primary services, or functions, within the ATM system are described in more detail, as defined by [9]:

- **Air Traffic Services (ATS)**: ATS is "*a generic term meaning variously, flight information service, alerting service, air traffic advisory service, Air Traffic Control (ATC) service (area control service, approach control service or aerodrome control*

service)" [9]. Specifically, the flight information service enables the safe and efficient conduct of flights by providing advice and information, such as weather reports or airport conditions. The alerting service exists to provide notification to the appropriate organizations with regards to aircraft in need of search and rescue aid and to assist these organizations in such efforts. The air traffic advisory service exists to ensure appropriate separation between aircraft operating on Instrument Flight Rules flight plans within an advisory airspace. Finally, the ATC service is critical to prevent collisions between aircraft as well as between aircraft and any obstructions on the manoeuvring area. Additionally, the ATC service exists to expedite and maintain an orderly flow of air traffic. Ground-based air traffic controllers direct aircraft both on the ground and through controlled airspace.

- **Airspace Management (ASM)**: ASM "*means a planning function with the primary objective of maximising the utilisation of available airspace by dynamic time-sharing and, at times, the segregation of airspace among various categories of airspace users on the basis of short-term needs*" [10]. The overall objective of ASM is to ensure airspace resources are allocated to system users such that each is able to successfully complete assigned tasks. ASM is generally accomplished by Air Navigation Service Providers (ANSPs). ANSPs strategically design the airspace and operational procedures.

- **Air Traffic Flow Management ATFM**: ATFM is "*a service established with the objective of contributing to a safe, orderly and expeditious flow of air traffic by ensuring that ATC capacity is utilized to the maximum extent possible, and that the traffic volume is compatible with the capacities declared by the appropriate ATS authority*" [9]. At times during which traffic demand exceeds defined ATC capacity, an ATFM service is implemented. Specifically, in the National Airspace System (NAS), ATFM is accomplished through Traffic Management Initiatives (TMIs), where TMIs

are implemented within strategic time frames (planning two to eight hours out) and tactical time frames (procedures carried out in less than two hours). ATFM enables demand to be matched with available capacity.

It is evident the three primary ATM services operate interdependently to ensure the safe and efficient operation of all aircraft within an airspace. For instance, ATFM focuses more on the broader systems approach (airspace-level) to managing air traffic and supports ATC, which provides flight-level instructions, in achieving the most efficient utilization of airspace/airport capacity, considering imposed safety constraints. Therefore, ATM systems undoubtedly impact operations at both the *flight level* and the *airspace level*. In this context, the flight level refers to the operations of a single flight within an airspace, whereas the airspace level refers to the operations of multiple flights within an airspace.

Considering the flight level, ATC may instruct an aircraft to follow a specific arrival procedure, perform a go-around, or enter into a holding pattern. On the other hand, considering the airspace level, an ATFM program, such as a Ground Delay Program, may be implemented through a TMI that results in the alteration of operations within an entire airspace. However, ATFM programs similarly impact several flight-level operations within the affected airspace. Therefore, this underscores the importance of an assessment of the relationship between flight-level and airspace-level operations.

Though, existing ATM systems continue to be limited by outdated technologies and operational procedures [11]. Specifically, ATM challenges include: (i) high volume and heterogeneous demand, (ii) dynamic system capacity, and (iii) fragmented, multi-stakeholder and human-based decision-making [11]. Of important consideration in improving the operations of ATM systems is the challenge of high volume and heterogeneous demand, which is anticipated to worsen as the forecast growth in air traffic volume is realized and AAM are introduced.

1.3 Global Modernization Efforts

In the past two decades, in response to current ATM challenges, as well as future challenges, global efforts have been underway to modernize ATM systems. The global modernization efforts include the FAA's Next Generation Air Transportation System (NextGen) [12] portfolio in the U.S., the Single European Sky ATM Research (SESAR)'s [13] program in Europe, and other country-specific programs [14, 15, 16]. NextGen is focused on modernization of the U.S. NAS, while SESAR is focused on the European ATM system. Currently, the U.S. operates the safest, largest, and most complex ATM system in the world [17], where the FAA's NextGen program aims to retain this leadership position in global aviation [18]. All global modernization efforts are long-term plans motivated by concerns that current ATM systems will not be able to accommodate forecast air traffic flow. NextGen, and other global modernization efforts, are not simply comprised of one technology, product, or goal. Instead, the modernization efforts are a composition of several diverse and interlinked portfolios of technologies, systems, policies, and procedures.

The long-term objectives of the global modernization efforts are formulated based upon the ICAO Global Air Traffic Management Operational Concept, which has the vision "*to achieve an inter-operable global air traffic management system, for all users during all phases of flight, that meets agreed levels of safety, provides for optimum economic operations, is environmentally sustainable and meets national security requirements*" [19]. Further, ICAO has established five Strategic Objectives: (i) Safety, (ii) Air Navigation Capacity and Efficiency, (iii) Security and Facilitation, (iv) Economic Development of Air Transport, and (v) Environmental Protection [20]. The FAA describes NextGen as an "evolution of air traffic management". As displayed in Figure 1.5, the ATM system has evolved from a procedural-based system, to a surveillance-based system, to, finally, a trajectory-based system, which is enabled by the implementation of NextGen concepts [17]. The ultimate objective of the global modernization efforts is a transition to Trajectory Based Operations

(TBO), as implementation of new technologies and concepts enables a transition from a traffic-controlled ground-based ATM system to a traffic-managed satellite-based ATM system. This transition to TBO is anticipated to enable management of four-dimensional (4D) aircraft trajectories (latitude, longitude, altitude, and time) such that aircraft operations become more efficient and predictable, while maintaining operational flexibility [21].

Figure 1.5: Air Traffic Management System Evolution [17]

In the context of the successful introduction of TBO, much emphasis has been placed on the "Safety" and "Air Navigation Capacity and Efficiency" Strategic Objectives. Thus, common among all of the global modernization efforts are long-term overarching objectives of increasing ATM system efficiency and capacity, while also maintaining safety, where development and implementation of new technologies are necessary to meet these objectives. ICAO's "Safety" and "Air Navigation Capacity and Efficiency" Strategic Objectives are necessarily functionally and organizationally interdependent [20], where ICAO identifies the attainment of a safe system as the highest priority in managing aviation operations [19]. Therefore, to successfully transform air traffic management systems worldwide, holistic consideration of both safety and capacity and efficiency is required.

An brief overview of both safety and capacity and efficiency in the context of the ATM system is provided. Additionally, due to implementation of TBO concepts being a large component of objectives to increase ATM system safety, capacity, and efficiency, a discussion on TBO is provided.

1.3.1 Safety

ICAO defines *safety* as "*the state in which risks associated with aviation activities, related to, or in direct support of the operation of aircraft, are reduced and controlled to an acceptable level*", where aviation *safety management* is defined as the activity that "*seeks to proactively mitigate safety risks before they result in aviation accidents and incidents*" [22]. However, aviation safety management did not always emphasize mitigating safety risks before they result in aviation accidents or incidents.

The first aviation accident with causalities occurred in 1908, and since this time many countless efforts have been made to improve aviation safety [23]. Prior to 1995, aviation safety was *reactive* in that after an accident or incident occurred, a mitigation strategy was then developed and implemented [24]. A reactive approach requires an accident or incident to have previously occurred such that a "working-backwards" approach may subsequently be taken to identify and address the underlying problem. A reactive approach is not capable of maintaining safety at adequate levels as air traffic increases and aviation systems are modernized. Thus, the aviation industry realized novel approaches were required [24].

For instance, in the past 25 years, the industry began a shift toward *proactive* and *predictive* approaches to safety. A proactive approach to safety involves identifying potential unsafe events in advance of their manifestation as accidents or incidents. This approach aims to enable mitigation strategies to be developed to prevent the occurrence of accidents or incidents related to any unsafe events [25]. However, to successfully implement a proactive approach to safety, records must exist of the current operational states. Taking safety analysis a step further, a *predictive* approach to safety involves monitoring data obtained from routine operations in addition to accidents and incident data and reports to detect potential negative future outcomes. Figure 1.6 summarizes the three approaches to safety.

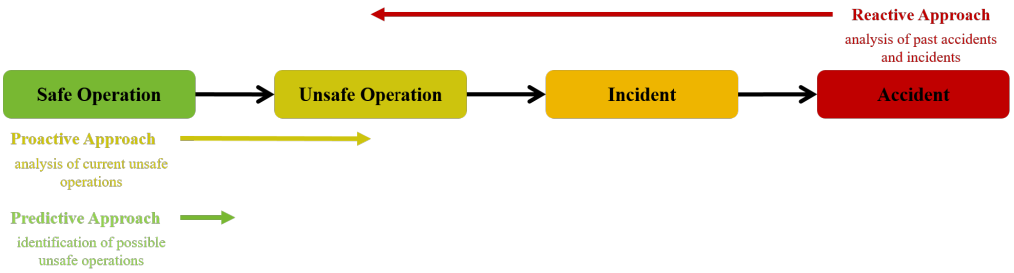

Figure 1.6: Different Approaches to Safety Management (adapted from [25])

A reactive approach to safety focuses on prevention of accident or incident *recurrence*, while proactive and predictive approaches to safety focus on prevention of accident or incident *occurrence*. The shift toward proactive and predictive safety approaches has enabled air transportation to remain the safest mode of transportation. Further, this shift is key to maintaining aviation safety in the future as aviation systems modernize and become more complex. Thus, today's aviation safety programs place a strong emphasis on proactive and predictive safety approaches. In the U.S., these safety programs include:

- **Aviation Safety Action Program (ASAP)**, which encourages voluntary reporting of safety issues and events that come to the attention of employees and certain certificate holders [26].
- **Aviation Safety Reporting System (ASRS)**, which enables anonymous reporting of aviation incidents from pilots, air traffic controllers, flight attendants, maintenance staff, and eye witnesses [27].
- **Flight Operational Quality Assurance (FOQA)**, also known as Flight Data Monitoring (FDM), which involves collection and analysis of routine flight data, where it is widely encouraged and utilized in commercial operations [28].
- **Aviation Safety Information Analysis and Sharing (ASIAS)**, which connects approximately 185 data and information sources across government and industry, including voluntarily provided safety data [29].

- **System Approach for Safety Oversight (SASO)**, which was established to develop and implement a comprehensive system safety approach to the oversight of aviation entities [30].

- **Safety Assurance System (SAS)**, which is the FAA's oversight tool to perform certification, surveillance, and continued operational safety [31].

Additionally, to facilitate proactive and predictive approaches to safety, one of National Aeronautics and Space Administration (NASA) Aeronautics Research Mission Directorate (ARMD)'s Strategic Thrusts for research is In-time System-wide Safety Assurance [32, 33, 34]. The vision of this Strategic Thrust is "*the ability to predict, detect, and mitigate emergent safety risks throughout aviation systems and operations*" [32]. "System-wide" alludes to encompassing multiple sources for NAS data, including a wide array of systems and system levels [34].

1.3.2 Capacity and Efficiency

Capacity and efficiency are necessarily intertwined. From an individual-flight perspective, efficiency addresses the operational effectiveness of gate-to-gate flight operations [17]. As exceeding capacity often results in air traffic flow restrictions, there is often an impact to a flight's efficiency. On the other hand, increasing the efficiency of the flights within an airspace system results in capacity increases. Capacity may be measured at the airport level (arrival rate), within the terminal airspace (maximum number of flights operating within the terminal airspace), en-route level (maximum number of flights in a sector), or at the system level (maximum number of flights that can be handled from gate to gate) [35].

Capacity is often uncertain, as it is dependent on several factors, where these factors include the ATC route structure, navigation accuracy of aircraft using the airspace, weather, and air traffic controller workload [9]. The most significant factor impacting airspace capacity is often identified as being weather [11]. The specific ATM configuration is often

heavily influenced by the weather conditions. Further, weather conditions directly impact the throughput performance of the selected ATM configuration [11].

In the context of the strategic planning of an airspace system, capacity to meet airspace demands during peak times and at peak locations is of the utmost importance [17]. Additionally, it is desirable to minimize restrictions on overall traffic flow such that the efficiency of individual flights is not negatively impacted [17]. Current FAA concepts and initiatives to increase ATM system capacity and efficiency include:

- **Performance Based Navigation (PBN)**, which represents a shift from sensor-based to performance-based navigation [36]. PBN enables aircraft to fly shorter, and, therefore, more efficient, flight paths [17].

- **Standard Terminal Automation Replacement System (STARS) and En Route Automation Modernization (ERAM)**, which enable air traffic controllers to work more effectively [17].

- **Automatic Dependent Surveillance-Broadcast (ADS-B)**, where of specific relevance is ADS-B Out technology, which allows messages including position, ground speed, and other surveillance information to be broadcast by aircraft at a rate of 1 Hz. As of January 1^{st}, 2020, all aircraft flying within most controlled airspace are required to be equipped with ADS-B Out. This transition to the use of satellite-based technology, which provides more accurate position measurements, enables controllers to separate aircraft closer to minimum separation standard, increasing capacity [37].

- **Decision Support System (DSS) Automation**, which includes three decision support hardware and software systems (Traffic Flow Management System, Time Based Flow Management System, and Terminal Flight Data Manager) that enable ATM system operators to quickly and efficiently respond to dynamic traffic and weather conditions [17].

- **System Wide Information Management (SWIM)**, which aggregates ATM system data from multiple sources. This aggregation of data enables faster information-sharing as multiple computer interfaces are no longer required to access data from various sources [17].

Finally, improvements in ATM system efficiency, and therefore capacity, are strongly related to the global strides towards implementation of TBO. For instance, TBO is expected to enable decisions to be made more strategically, which improves both flight and airspace level efficiency [38]. In addition, more accurate trajectories result in observed demand being more closely matched to capacity [38].

1.3.3 Trajectory Based Operations

As introduced, TBO is intended to efficiently manage aircraft trajectories in 4D (latitude, longitude, altitude, and time). TBO enables strategic planning, management, and optimization of trajectories operating within and across various airspace by leveraging time-based management, information exchange between air and ground systems, and an aircraft's ability to fly precise paths in time and space [21]. The FAA describes TBO as "a collection of systems, capabilities, processes, and people working together to achieve operational objectives" [39].

The amalgamation of Time Based Management (TBM), PBN, and technologies that enable the distribution and sharing of trajectory information support efforts to introduce TBO into the NAS [39]. TBO is intended to enable airspace users to "negotiate" trajectories with the ANSPs such that a trajectory that is a result of collaboration and information-sharing is agreed upon. This trajectory, known as the "agreed trajectory", includes a route between an origin and destination pair with predicted cross time estimates at key points [21]. TBO is expected to enable collaborative decisions that consider the entire airspace system, providing routes optimized at both the*flight level* and the *airspace level*. In this context, flight-level refers to the operations of a single flight, while airspace-level refers to

the aggregated operations of all flights within an airspace.

As a new concept of operations is deployed with TBO, it is of paramount importance to develop methods to assess a flight's trajectory in the context of the other flights' trajectories operating within an airspace. Further, that the envisioned TBO is anticipated to result in a less structured airspace as a result of more flexible trajectory planning, which further underscores the importance of assessment of airspace-level operations. Additionally, acquiring an understanding of how operations at the two levels are related is relevant to support ATM system modernization efforts to increase safety, capacity, and efficiency. To enable successful introduction of new TBO concepts, robust ATM system analyses are required that consider operations at both the flight and airspace levels, where the relationship between operations at the two levels is likewise considered and assessed. This leads to **Observation 1**:

Observation 1

Gaining a comprehensive understanding of both flight-level and airspace-level operations enables Air Traffic Management system operators, planners, and decision-makers to make better-informed and more robust decisions related to the implementation of future operational concepts.

1.4 Aviation System Operations

Aviation systems are complex systems-of-systems, where these systems will continue to experience an increase in complexity as a result of the concurrent impact of the forecast rise in volume of air traffic and the transformation of ATM systems in response to the global modernization efforts. A brief overview of the significant challenges posed by the complexity of aviation systems in the context of the global transformation of ATM systems is presented. Additionally, an introduction to the aviation system operational data recorded and its importance in the transformation of ATM systems is provided. Finally, while all

aviation systems are incredibly complex systems of systems, the aviation system that often experiences one of the highest levels of complexity in the context of ATM system operations is the terminal airspace system. Therefore, the complexity of terminal airspace ATM system operations and their relevance in the context of broader ATM system operations are discussed.

1.4.1 System Complexity

As introduced, aviation systems are complex systems-of-systems that will continue to evolve and transform as air traffic volume increases and initiatives and programs related to the global modernization efforts are implemented. Moreover, the complexity of aviation systems is further anticipated to increase with the introduction of AAM concepts. The operations, or *behaviors*, of a "high-level" aviation system, such as an airspace, are the product of an aggregation of the operations, or behaviors, of the flights ("lower-level" aviation systems) within the airspace of interest. In this context, *behaviors* refer to the operations of a system that are *visible* to other systems, including other systems at the same level and the higher-level system-of-interest. Related to an airspace and the flights operating within it, the behavior of a flight may refer to its trajectory. Thus, the higher-level system behavior of entire airspace is dependent upon the lower-level system behavior of flights, or trajectories, observed to be operating within the airspace bounds. Both flight-level and airspace-level behaviors/operations are impacted by an increase in aviation system complexity and are relevant in the context of the holistic consideration of both safety and capacity and efficiency as ATM systems are transformed. Further, as introduced, the relationship between behaviors observed at the flight level and the airspace level must be assessed and understood. Analysis of operations at both the flight level and the airspace level as well as the relationship between operations at the two levels is paramount to support ATM system planners and decision-makers.

Specifically, in complex systems-of-systems theory, a property of *emergence* originates

from the idea that when a physical, higher-level system reaches some level of complexity in its organization, it may begin to exhibit genuinely novel properties or behaviors not possessed by its simpler, lower-level components, or systems [40]. Emergence is said to "*occur when an entity is observed to have properties that its parts do not have on their own such that these properties or behaviors emerge only when the parts interact as a wider whole*" [41]. As defined by the INCOSE Systems Engineering Handbook, an *emergent behavior* is "*a behavior of the system that cannot be understood exclusively in terms of the individual system elements*" [42].

To improve levels of ATM system capacity and efficiency, considering increases in complexity, a comprehensive understanding of current ATM system operations is necessary in addition to the ability to recognize and perform timely mitigation of potentially undesirable emergent behavior. Moreover, it is of paramount importance to discover new, or emergent, safety risks in existing ATM systems, including safety risks that did not exist in ATM systems previously, i.e. those safety risks that are a result of the implementation of new a concept of operations related to the modernization efforts [43]. There is increased difficulty in developing the ability to analyze and predict threats posed by emergent safety risks as the threats could increase in frequency and severity as ATM systems evolve [33].

ATM approaches to assess the behavior of flights operating within an airspace most commonly rely on monitoring trajectory data for exceedance of a priori Subject Matter Expert (SME)-set thresholds in a specified set of parameters [44, 45, 46]. In this context, an exceedance is the deviation of a metric or multiple metrics beyond an established threshold within a specified time interval [44]. This approach has a number of limitations, with a primary limitation being reliance on a predefined criteria to detect anomalies, which leaves emergent, or previously unknown, risks or issues undetected [47]. Further, current methods to assess airspace behavior rely on human experience, where air traffic controllers form mental models and create abstractions of the behavior in a local region to inform their decisions [11]. The reliance on air traffic controller experience/intuition may introduce

inefficiencies as the decisions are based more on observations of a local region rather than an analysis from a systems perspective [11].

Thus, to assess the impact of new technologies and a new concept of operations at both the flight and airspace levels, novel, more automated, analysis methods are required that enable detection of emergent behavior in aviation systems [12, 13, 14, 15, 16, 32, 43]. However, the large scale and complexity of aviation systems make it challenging to use classical dynamical or physics-based methods for the necessary analyses [46, 48, 49, 50]. Fortunately, as a product of the modernization efforts, new data-generating technologies and programs have been introduced in aviation systems, which has resulted in more sources, volume, and availability of operational data than ever before [12]. This leads to **Observation 2**:

Observation 2

The complexity of Air Traffic Management systems restricts the use of classical physics-based methods for the analysis of new operational concepts, yet increased availability of operational data provides new analysis opportunities.

1.4.2 Operational Data

The efficient use and management of ATM system operational data is key to facilitate the modernization efforts [11]. The types of operational data recorded and stored include numerical (continuous and discrete numbers), textual (characters and strings), and multimedia (audio and video) data. The multitude of operational data sources may include data sources related to environmental conditions (e.g. weather reports), traffic conditions (e.g. airline schedules, flight plans, delays), or flight tracks (e.g. surveillance data). Arguably the highest-fidelity operational data source, the Flight Data Recorder (FDR) on-board an aircraft, records massive amounts of operational data for each flight. This data is collected as part of the structured FOQA program. The amount of operational data recorded includes

anywhere between 80 to 2,000 metrics at a sampling rate of 0.25 to 8 Hz [51]. As operational data collection technologies continue to advance, it is predicted that the global fleet of aircraft could generate upwards of 98 million terabytes of data by 2026 [52]. Though, access to FOQA data is highly restricted.

On the other hand, the expansion of ADS-B technology has enabled open-source trajectory data to become publicly available and much more accessible, which makes large-scale analyses more feasible [53]. As mentioned, ADS-B technology has been deployed to enhance ATM system capacity, efficiency, and safety by enabling operators to more precisely determine aircraft position, as well as enabling aircraft to determine their position with respect to similarly-equipped aircraft, using satellite, inertial, and radio navigation [54]. ADS-B Out periodically emits (at approximately 1 Hz) an aircraft's position (latitude and longitude), altitude, heading, horizontal (ground) and vertical speed, along with other relevant metrics to ground stations and other equipped aircraft [54]. A decent ADS-B receiver may receive messages from cruising aircraft located up to 200 miles away. The trajectory data available in ADS-B messages is the core information that is used by ATM systems as a basis for activities such as:

- Distributing flight information to relevant airlines and ATC units
- Facilitating timely coordination between sectors and units
- Correlating flight data with tracks
- Monitoring the adherence of an aircraft to its assigned route
- Detecting and resolving conflicts

The explosion in volume and availability of operational data recorded may be exploited to provide many new opportunities for analysis of ATM systems [46, 48, 50, 55]. Specifically, as the volume of operational data recorded grows, data-driven analysis methods are fitting to address the challenge of improving capacity and efficiency, while maintaining

safety, in increasingly complex systems. The vast amount of operational data recorded provides unprecedented opportunities regarding understanding the intricacies of ATM systems. Additionally, a need has been identified to understand and monitor higher system-level phenomena in the context of leveraging the vast amount of recorded aviation data [48]. If managed well, the operational data recorded may provide novel insights into the complex dynamic problems related to ATM systems [48]. These novel insights may be extracted related to operations at both the flight and airspace levels. Moreover, the importance of a *data analyst perspective* is emphasized, which is the ability to shift attention from the focus of the dynamics of single events to the emerging statistical characteristics of large data sets [48].

The discovery of previously unknown safety risks or efficiency issues in aviation operational data falls under the scope of knowledge discovery and information extraction. Efforts to mine the operational data support the aviation safety life cycle and strategies for safety improvement, as identified by NASA, [56]. *Data mining* is defined, formally, as "*the process of applying computational methods to large amount of data in order to reveal new non-trivial and relevant information*" [57]. Data mining methods are generally split into two categories: statistical models and machine learning models. Data mining methods have been identified to play a significant role in understanding the NAS [46]. Their significance is related to developing capabilities to both mitigate current inefficient and/or unsafe operations and anticipate future inefficient and/or unsafe operations.

Human operators and analysts are simply not able to properly analyze ATM system operational data for emergent safety risks or efficiency issues due to the high volume and dimensionality of the data they are presented with. Additionally, as introduced, current methods in place to narrow the scope of operations to review, such as exceedance-based methods, have a number of limitations. Thus, data mining methods are better equipped to provide operational insights related to interdependencies in high-dimensional trajectory time series data and detect potentially undesirable emergent behavior. Recently, advances

in modern machine learning techniques have significantly contributed to the use of data-driven techniques to gain insights from the operational data, specifically ADS-B trajectory data. There is consensus that data analytic methods applied to ATM syste, operational data can significantly contribute to improving future operations [58]. Specifically, monitoring ATM system behavior to identify successful practices and inefficiencies or safety risks to provide guidance in offline system adjustments is necessary [58]. In this context, *offline* refers to analysis of stored historical data as opposed to online analysis of streamed data in real time or near-real time. Offline analysis provides the opportunity to discover something novel in the data as well as to perform data exploration, which enables an analyst to obtain a more comprehensive understanding of a system's operations [46]. This leads to **Observation 3**:

Observation 3

ADS-B data provides the basis for offline data-driven methods to be applied to analyze Air Traffic Management system operations at both the flight level and the airspace level.

1.4.3 Terminal Airspace Systems

The terminal airspace is *a general term describing the airspace surrounding an airport*. The terminal airspace system refers to *the complex system of systems related to all operations that occur within the terminal airspace area*. The terminal airspace system experiences a high density of converging and diverging aircraft, the interdependent utilization of both airspace and airport resources, and is, generally, a more constrained airspace featuring complex traffic dynamics [53]. Overall, the terminal airspace system is a decidedly complex system that is highly sensitive to traffic conditions and environmental conditions [59].

The size and shape of a terminal airspace system depends on the number of airports

contained within it, the airway configurations, and the number and length of arrival and departure trajectories [59]. A terminal airspace system does not necessarily correspond to only one airport location as there may exist a number of airport locations within one terminal airspace system, such as in metroplex terminal airspace systems. Metroplex terminal airspace systems serve large metropolitan regions and encompass two or more airports [11, 60]. Metroplex terminal airspace systems containing more than one major commercial airport feature more complex traffic dynamics and significant interactions and interdependencies, which further complicates already very complex terminal airspace systems [11].

In the context of system-wide efficiency and safety, the terminal airspace is of great significance. Operations within the terminal airspace, in particular, are known to greatly impact both flight-level and airspace-level capacity and efficiency [53]. Further, most accidents and incidents occur within the terminal airspace, so it is a safety-critical airspace. Figure 1.7 displays the relative frequency of accidents between 2001 and 2020 with respect to the flight phase in which they occur. All phases, with the exception of the *en-route* phase, occur when aircraft are operating within the terminal airspace. The en-route phase makes up less than 10% of total both fatal and non-fatal hull loss accidents; hence, over 90% of all accidents and incidents occur within the terminal airspace [61]. In addition, the majority of both fatal and non-fatal hull loss incidents and accidents occur during the operations of arriving aircraft, during the *approach* or *landing* phases of flight.

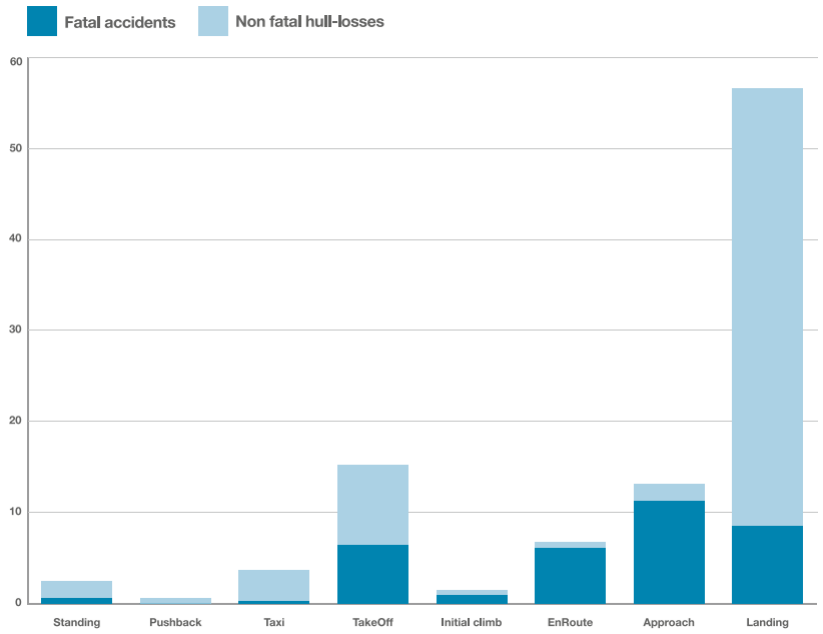

Figure 1.7: Distribution of Accidents per Flight Phase, 2001-2020 [61]

Therefore, considering the global ATM system modernization objectives of increasing safety, capacity, and efficiency, it is important to focus on operations within the terminal airspace. Specifically, due to the higher rate of incidents and accidents during the operations of arriving aircraft, a focus on arriving aircraft operations within the terminal airspace is of paramount importance. This leads to **Observation 4**:

Observation 4

Analysis of Air Traffic Management system operations of arriving aircraft has the highest potential to impact system safety, capacity, and efficiency levels.

1.5 Summary

The background concepts and observations are summarized. Prior to the onset of the COVID-19 Pandemic, air traffic was experiencing steady, long-term growth with significant anticipated growth over the next 20 years. Despite the drastic reduction in air travel, the aviation industry is expected to make a recovery and continue the trend of significant

growth. Thus, in response to the past and predicted air traffic growth, within the past two decades, global efforts, such as NextGen in the U.S., have emerged to support the modernization of ATM systems with the objectives of introducing new technologies and concepts of operation to improve efficiency and capacity, while also maintaining high levels of ATM safety.

Current aviation systems are decidedly complex systems-of-systems, where both the forecast rise in air traffic volume and implementation of a new concept of operations related to ATM transformation efforts will increase this complexity. Specifically, this complexity will result in modifications to the concepts of operations at the flight-level and, consequently, the airspace-level, where analysis of the current state of operations and proposed future state of operations is paramount to inform planning and decision-making as ATM systems are transformed. Thus, novel analysis methods are required to assess the impact of the new technologies and concept of operations. Due to system complexity, classical dynamical or physics-based models are difficult to utilize. However, new data-generating technologies, such as ADS-B, have been introduced, which has made development of offline data-driven methods for ATM system analysis more feasible and prevalent. Operations within the terminal airspace, specifically operations of arriving flights, are the most safety-/capacity-/efficiency-critical. Considering all observations leads to the following **Overarching Research Question**:

Overarching Research Question

How can the offline application of data-driven methods to ADS-B data be leveraged to analyze Air Traffic Management system arriving aircraft operations at both the flight level and the airspace level?

The remainder of this thesis adheres to the following structure:

- chapter 2 reviews the literature on current data-driven approaches that leverage ADS-B data to analyze ATM system arriving aircraft operations, both at the flight and

airspace levels, such that gaps are identified.

- chapter 3 details the formulation of pertinent research questions and associated hypotheses.
- chapter 4 provides an in-depth overview of the design of the experimental approaches required to test the hypotheses formulated in chapter 3.
- chapter 5 details the implementation of each of the experimental approaches designed in chapter 4 and presents and discusses the results of each implementation.
- chapter 6 concludes the thesis, presenting the contributions and offering recommendations for future work.

CHAPTER 2
LITERATURE REVIEW

As introduced in chapter 1, this research is motivated by developing capabilities to support the global modernization efforts to improve ATM system safety, capacity, and efficiency. Specifically, this research seeks to determine how offline application of data-driven methods to ADS-B data extracted for arriving aircraft may be leveraged to obtain actionable insights pertaining to the analysis ATM system operations at both the flight and airspace levels. The literature review is subsequently motivated by gaining an understanding of the existing data-driven methods and approaches to analyzing ATM system arriving aircraft operations.

As data-driven approaches to analyze operational data, specifically ADS-B data, have gained traction among researchers, several methods have been developed motivated by gaining increased understanding of ATM systems as well as the global modernization effort objectives of increasing safety, capacity, and efficiency. The existing aviation literature related to applying data-driven methods to ADS-B data is often concerned with two primary, sometimes inter-related tasks: *air traffic flow identification* and *anomaly detection*. Generally, air traffic flow identification and anomaly detection are leveraged to support a flight-level analysis. Air traffic flow identification and anomaly detection may be considered two relevant *research areas* related to the analysis of ATM system arriving aircraft operations at the flight level. With respect to an airspace-level analysis, the aviation literature is fairly limited with no clearly defined tasks. However, the need for an airspace-level analysis component in the analysis of ATM system arriving aircraft operations has been established. Therefore, a broader third research area may be identified as overall airspace-level analysis. Considering the three research areas identified the following **Overarching Hypothesis** is formulated:

Overarching Hypothesis

If an offline data-driven methodology to be applied to ADS-B data for arriving aircraft is developed that performs (i) **Air Traffic Flow Identification**, (ii) **Anomaly Detection**, and (iii) an **Airspace-Level Analysis**, then Air Traffic Management system arriving aircraft operations are analyzed at both the flight level and the airspace level.

An outline of the three steps in the proposed methodology are displayed in Figure 2.1.

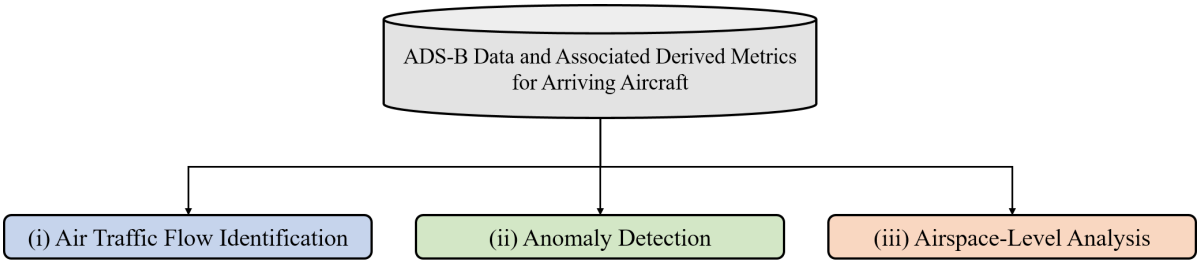

Figure 2.1: Outline of Proposed Offline Data-Driven Methodology to be Applied to ADS-B data for Arriving Aircraft

Air traffic flow identification is sometimes completed as a data pre-processing step for flight-level anomaly detection, resulting in the inter-related nature of these two tasks. However, air traffic flow identification may also be the primary method leveraged for the flight-level anomaly detection task. Methods for anomaly detection are undoubtedly diverse and no investigation into any interdependencies or relationship that may exist between different *types* of anomalies detected in ADS-B data, such as those detected leveraging air traffic flow identification methods or those detected in additional distinct sets of ADS-B data and associated derived metrics. On the other hand, related to airspace-level analysis, there has been some adoption of the output from the air traffic flow identification task to support airspace-level operational pattern analysis. However, overall, the aviation literature related to the development of methods to support airspace-level analyses is sparse.

The objective of this literature review is to identify any gaps related to the three out-

lined steps in the proposed offline data-driven methodology to be applied to arriving aircraft ADS-B trajectory data. The literature review begins with a review of air traffic flow identification methods. Next, a review of and distinction between the various anomaly detection methods available in the literature is presented. Finally, an overview of the limited methods that support airspace-level analysis is provided.

2.1 Air Traffic Flow Identification

Trajectories following a similar spatial path, or belonging to the same *air traffic flow* tend to correspond to a specific structured operation within an airspace, as defined by ATC. Formally, a trajectory may be defined as "*a mathematical object used to describe the evolution of a moving object*" [62], where a time-ordered list of *state vectors* may be used to represent this evolution. Typically, the states of interest related to aircraft trajectories include: longitude, latitude, altitude, speed (typically ground speed), and track angle or heading. While the precise definition of an air traffic flow depends on the application, it is generally considered a pattern of air traffic in the spatial and, in some applications, temporal dimensions. A large portion of aviation research focuses on the spatial dimension of air traffic flows, considering the longitude and latitude trajectory states [50, 63, 64, 65, 66, 67, 68, 69]. To provide a visual reference for the appearance of air traffic flows, Figure 2.2 displays the air traffic flows identified at San Francisco International Airport (KSFO) considering four months of operations in 2019. Air traffic flow identification is often performed to support a more in-depth analysis of the efficiency of airspace operations. Specifically, air traffic flow identification techniques have been developed to support performance assessments, airspace monitoring efforts, more effective ATFM, and ATM system decision-making [11]. Moreover, air traffic flow identification algorithms may be leveraged to support airspace redesign [58], conflict detection [70], and environmental impact assessment [71, 72]. For instance, to determine optimal sector partitioning with respect to workload measures, it is useful to be aware of the actual location of major traffic patterns [58]. Further, several ana-

lytical models for tactical ATFM rely on crude airspace capacity estimates and abstractions that do not correspond to the air traffic flows observed [58]. Thus, knowledge of the actual locations of air traffic flows supports the development of data-driven airspace capacity models that are able to provide more accurate estimates [58]. Additionally, techniques to identify and understand air traffic flows become increasingly important as the global modernization effort concepts move away from fixed routes that structure the air traffic [73]. Finally, air traffic flow identification may be implemented as a pre-processing step prior to applying anomaly detection methods, discussed in greater detail in section 2.2.

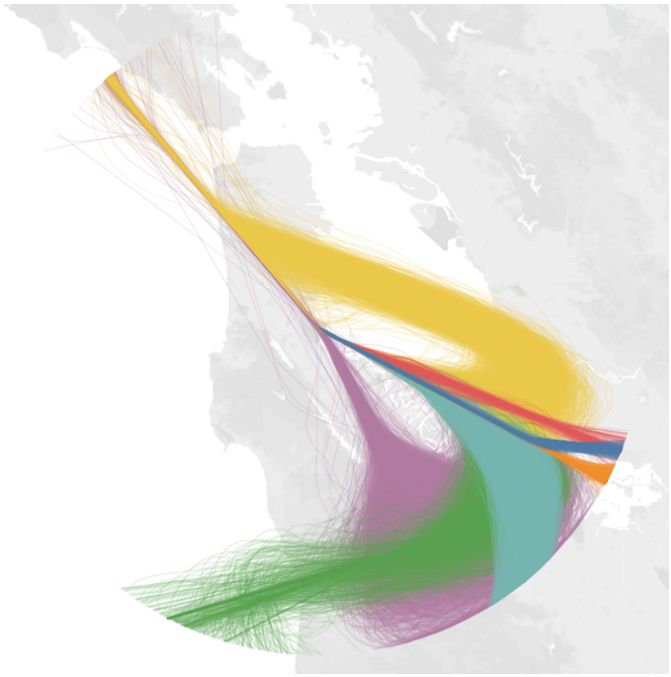

Figure 2.2: Air Traffic Flows Identified at San Francisco International Airport within Four Months of Operations in 2019

Typically, air traffic flows are identified through application of trajectory clustering methods. *Clustering* may be defined as "*the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar*" [74]. Customarily, the *trajectory clustering* task is formulated as an unsupervised machine learning problem, where the task involves partitioning a data set containing individual flight trajectories into similar groups to define air traffic flows [75]. The key distinction between

supervised and unsupervised machine learning problems is the presence of ground-truth *labels* within the data set. The input to these methods typically includes a feature vector matrix, where each row represents a single flight and each column represents a trajectory state (such as longitude and latitude coordinates) at a specified trajectory point.

Numerous diverse trajectory clustering methods have been applied to identify air traffic flows in previous works. Eckstein [76] presents an automated flight track taxonomy to identify air traffic flows within a terminal airspace by combining Principal Component Analysis (PCA) [77] and k-Means clustering [78]. PCA is a technique commonly leveraged for dimensionality reduction, where information loss is minimized. However, a potential limitation of the usage of PCA to reduce the dimension of a complex, high-dimensional data set is that it is a linear method, and may not perform well on a data set in which correlations are non-linear. Further, a limitation of the k-Means clustering algorithm leveraged by Eckstein [76] is that the algorithm does not identify nor exclude any outliers during the clustering process, which may skew results. Rehm [79] leverages a single linkage hierarchical clustering technique to identify prevalent airport arrivals routes automatically. To deal with the "outlier issue", a grid-based approach to outlier detection is used prior to applying the clustering algorithm [79]. However, implementation of this method requires setting, a priori, the number of air traffic flows to detect [79], which may not always be feasible, or desirable.

Gariel et al. [63] propose two methods for air traffic flow identication: (1) trajectory clustering based on way points and (2) clustering based on trajectory states via PCA. Both of these proposed methods leverage the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. DBSCAN is robust to noise such that outliers that do not appear to belong to an air traffic flow are officially designated as outliers by the algorithm [80]. Additionally, DBSCAN does not require any a priori specification of the number of clusters to form [80]. Simply put, density-based clustering algorithms operate on the assumption that clusters exist in high-density regions in the data space, with low-density

regions existing between the clusters. While Gariel et al. [63] do augment the feature matrix with additional derived features, usage of the PCA dimensionality technique is an inherent limitation of this method, as explained previously, due to potentially significant non-linear data relationships not being able to be preserved. Finally, Gariel et al. [63] find that way-point-based trajectory clustering performs sub-optimally compared with the application of the trajectory state-based clustering method.

Enriquez & Kurcz [81] aim to avoid the shortcomings associated with the application of PCA to trajectory state data by developing an air traffic flow identification algorithm based on spectral clustering. This algorithm is demonstrated on both en-route and terminal airspace trajectories [81]. However, the parameters specified related to the spectral clustering algorithm developed include a local scale parameter used to construct the similarity matrix, which effectively pre-determines the "width" of each cluster [81]. Due to differing specifications for the diverse procedures within a terminal airspace, the definition of a uniform "width" across all clusters may be considered a limitation of this method.

Marzuoli et al. [82] build upon the method introduced by Gariel et al. [63] and present clustering of en-route trajectories within an airspace. In the same procedure as [63], the data set is augmented with additional features, PCA is applied, and then DBSCAN is applied to identify the air traffic flows. Murça et al. [11, 58, 83] present a framework for air traffic flow characterization, which begins with the identification of air traffic flows leveraging DBSCAN. Delahaye et al. [84] present the usage of a novel distance metric formulated to measure the similarity of trajectories to one another considering entire airspace operations. This distance metric is utilized with a hierarchical clustering algorithm to cluster trajectories operating within the French airspace during a single day [84].

Basora et al. [73] introduce a trajectory clustering framework utilizing the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm [85] to identify air traffic flows. HDBSCAN is essentially an improved hierarchical version of the popular DBSCAN algorithm. The primary benefit of applying HDBSCAN is that it is

capable of managing clusters of varying densities [85]. Additionally, HDBSCAN requires only a single input parameter, which indicates the minimum number of data samples that are required to form a cluster [85]. Basora et al. [73] demonstrate the use of two different distance functions for the application of HDBSCAN: the Euclidean Distance (ED) and Symmetrized Segment-Path Distance (SSPD). Olive & Morio [54] propose a method intended specifically to identify air traffic flows within the terminal airspace. This method recursively applies DBSCAN to cluster "significant trajectory points", and then builds a dependency tree that is ultimately used to label a trajectory as belonging to an air traffic flow or as an outlier [54]. The significant trajectory points identified are those returned after the application of the Douglas-Peucker [86] algorithm. The Douglas-Peucker algorithm reduces a curve composed of line segments into fewer line segments, based on a sensitivity parameter [86]. Effectively, the Douglas-Peucker algorithm identifies trajectory "turning points", or way points. Though, Gariel et al. [63] explored a way-point-based clustering method and did not observe good performance.

Olive & Basora [87] demonstrate the use of DBSCAN and a Gaussian Mixture Model (GMM) for trajectory clustering. In this application, DBSCAN is applied to both the full data set as well as a data set that results from the application of the t-Distributed Stochastic Neighbor Embedding (t-SNE) [88] dimensionality reduction technique [87]. The GMM clustering does not allow for consideration or detection of outliers [87]. In other work, Olive & Barsora [66] introduce a recursive implementation of DBSCAN to identify air traffic flows and outliers for en-route trajectories. Recently, Olive et al. [67] demonstrated the use of autoencoder-based deep clustering techniques to identify and characterize air traffic flows. Specifically, Olive et al. [67] present two deep clustering techniques, both of which embed trajectories into latent spaces to facilitate the air traffic flow identification task. Finally, as mentioned, air traffic flow identification is often implemented as a data pre-processing step prior to the application of anomaly detection methods to prevent overly-conservative anomaly detection models from being developed [50]. For instance, Puranik

[25], Deshmukh [50], and Olive & Basora [68] applied DBSCAN to identify air traffic flows and outliers prior to applying anomaly detection methods.

To perform any of the methods of trajectory clustering presented, similar data instances are grouped into clusters according to a defined *similarity function* or *distance function*. A distance function is a function defined to compute the similarity, or distance, between two data instances. It has previously been noted that one of the most critical components of the trajectory clustering task is the definition of an appropriate distance function [54]. Trajectories are functional in nature, not independent data points, so the selection of an appropriate distance function is often not straightforward [67]. The ED is the most commonly implemented to compute the distance between two n-dimensional trajectories [68].

Due to the convergence/divergence of trajectories within the terminal airspace, the ED computed between trajectory points closest to the airport is relatively small, regardless of the air traffic flows to the two trajectories of interest belong. This is primarily due to airport runway configuration and the runways assigned for arrivals and departures. On the other hand, the ED computed between trajectory points at the terminal airspace's defined analysis border may be relatively large, even for two trajectories-of-interest belonging to the same air traffic flow. The ED between these trajectory points depends on the density of the air traffic flows and the defined terminal airspace radius around the airport. The uneven distribution of distances as aircraft arrive at or depart from the airport may skew the classification of trajectories. This skew may result in the misclassification of trajectories and, therefore, inadequate identification of air traffic flows. Therefore, the use of the ED to cluster trajectories may not be adequate [54, 89].

Other distance functions have been proposed to address the issues related to the limitations of the ED, such as the SSPD [89], as implemented by Basora et al. [73]. Basora et al. [73] compares the use of the ED versus the SSPD to cluster en-route trajectories and determined that the use of the SSPD results in more precise air traffic flow identification. However, there exists a significant trade-off in terms of computation time, where a large

computation time is likely to prevent subsequent applications of an algorithm in real-time. Gariel et al. [63] and Olive & Morio [54] present air traffic flow identification methods that are iterative and point-based. Although, the clustering of entire trajectory records or multi-point segments at once is often a requirement to enable the extension of methods to real-time applications. While the real-time application of methods is not the focus of this work, aviation organizations have set goals to ultimately deploy methods enabling the characterization of the ATM systems in real-time or near-real-time [33]. This leads to the identification of **Gap 1**:

Gap 1

A reliable method to identify air traffic flows that considers the converging nature of arrival trajectories and may ultimately be extensible to real-time applications.

2.2 Anomaly Detection

Anomaly detection may be defined broadly as *the process of detecting rare instances or sets of instances in a data set that may be of concern due to behaviors or characteristics that differ from the rest of the data set*. An anomaly is defined, formally, by the Merriam-Webster dictionary, as "*something different, abnormal, peculiar, or not easily classified*" [90]. However, more in the context of data mining, an anomaly may be defined as "*data deviating from or not being in agreement with what is considered normal, expected, or likely in terms of data probability distributions or the shape and amplitude of a signal in a time-series*" [49] or, more simply, "*patterns in data that do not conform to a well-defined notion of normal behavior*" [91]. In general, the terms anomaly and outlier are interchangeable when referring to the use of anomaly detection methods. Anomalies are often classified into three categories: *point*, *contextual* (also referred to as *conditional*), and *collective* anomalies, where a brief overview of each is presented [49, 91]:

- A **point anomaly** is a data instance that differs by a measure of significance from

the rest of the data instances in a data set [49]. The detection of point anomalies has been the focus of most research related to anomaly detection as it is considered the simplest type of anomaly. A visual representation of point anomalies in a data set of two-dimensional data points is displayed in Figure 2.3.

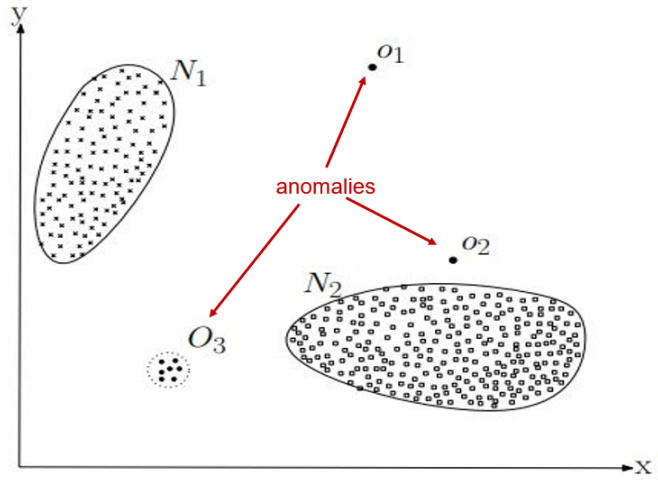

Figure 2.3: Point Anomaly [91]

- A **contextual, or conditional, anomaly** is a data instance that is an anomaly only in a particular context within a data set [49]. Context is ascertained by the structure of the data set and typically must be specified as part of the problem formulation [91]. Each data instance is said to be defined using *contextual attributes* and *behavioral attributes* [91]. Context attributes determine the context, i.e. neighborhood, for the data instance [91]. Behavioral attributes define the non-contextual characteristics of the data instance [91]. A visual representation of a contextual anomaly in a time-series data set of monthly temperature is displayed in Figure 2.4. In this monthly temperature plot, a contextual attribute would be the month the temperature is recorded in and a behavioral attribute would be the recorded temperature. Though, recorded temperature t_2 is not a temperature that differs significantly from the rest of the data set, as recorded temperature t_1 is the same value. However, the recorded temperature t_2 is

an anomaly, as it is much lower than the previous June recorded temperature, where the recorded temperature t_1 is associated with December, a usually low-temperature month.

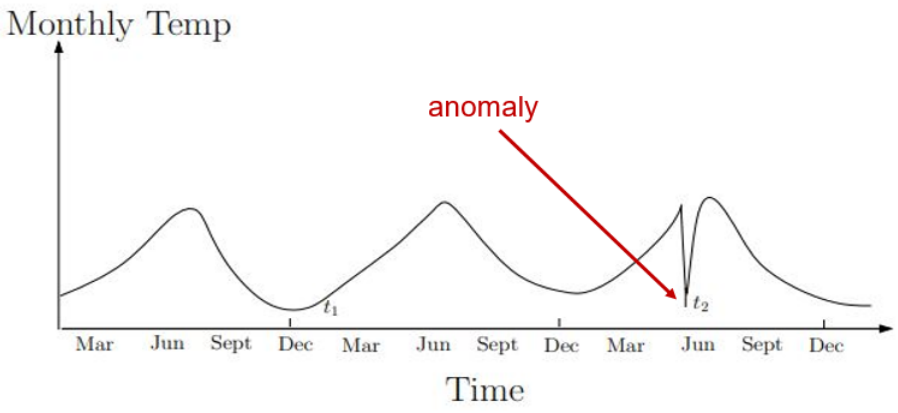

Figure 2.4: Contextual (Conditional) Anomaly [91]

- A **collective anomaly** is a data instance that may be a set of data points, or data object, such as a time-series sub-sequence, that is an anomaly as a whole, yet subsets of the data instance may not be anomalies on their own [49], i.e. the "collective" of the data instance, or data sub-sequence, is an anomaly. Collective anomalies are only able to be detected in data sets in which data instances are related in some way, such as sequential time-series data, spatial data, or graph data [49]. A representation of a collective anomaly in time-series data is displayed in Figure 2.5. In this time-series plot, there is an identifiable pattern, and it is evident that the anomalous sequence does not exceed the data set typical range in values; however, the repeated mid-range values that do not follow the typical pattern of the time-series data set cause the data instance to be identified as an anomaly. A point anomaly or a collective anomaly may also be a contextual anomaly if contextual information is specified [91].

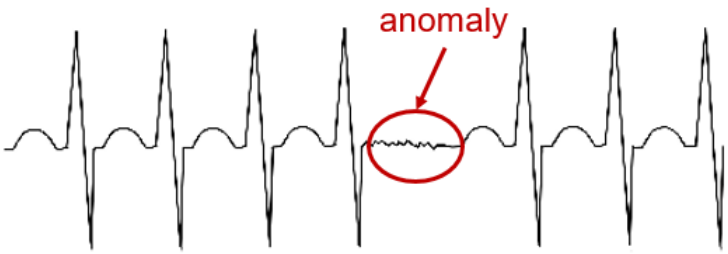

Figure 2.5: Collective Anomaly [92]

Development of methods to detect all three types of anomalies have been presented within the aviation literature related to detecting anomalies leveraging the ADS-B trajectory data. Related to aviation, anomalies have been defined as "*uncommon events with a potential for a meaningful operational or safety-related risk*" [64, 93]. While an anomaly does not inherently indicate an issue, it may point to a potential issue or emergent behavior that may be of interest to an analyst for further evaluation [25, 94, 95]. Due to less knowledge regarding potential consequences of anomalies, there is more risk, in the context of both efficiency and safety, associated with the non-standard operations. There exists much data and experience regarding the evolution of the standard, or normal, operations, yet there is little information regarding the evolution of non-standard operations.

In the context of ATM systems, model-based anomaly detection methods are generally difficult to implement due to the use of models with insufficient fidelity [49]. The insufficient fidelity results from over-simplified assumptions that are often made due to the complexity of the system. Generally, the purpose of the application of anomaly detection methods to ATM system operational data is to support operators, planners, and decision-makers in facilitating an efficient and safe airspace. Currently, exceedance detection is one of the most common methods of analyzing ATM system operational data to detect anomalies [44, 45, 46]. An exceedance is the deviation of a metric or multiple metrics beyond a pre-established threshold within a specific time interval [44]. The exceedance detection method has a number of limitations, with a primary limitation being reliance on a pre-defined criteria to detect anomalies, which leaves emergent behavior undetected

[47]. Therefore, leveraging machine learning techniques to detect anomalies in ATM system operational data has become common in the aviation literature. Utilization of machine learning techniques to detect anomalies has many benefits over methods that rely on detecting exceedances in a specified set of metrics. Thus, in recent years, the development of machine learning-based methods to detect anomalies in ATM system operational data, specifically ADS-B data, has significantly increased, as evidenced by a recent increase in the literature on this topic. Basora & Olive [49] publish a review of the recent advanced anomaly detection methods applied to aviation.

Basora & Olive's [49] review includes an overview of five identified anomaly detection method categories, which include: *distance-based*, *ensemble-based*, *statistics-based*, *domain-based*, and *reconstruction-based* [49]. Ultimately, the choice of an anomaly detection method is dependent on the type, dimensionality, and heterogeneity of the data set. Further, the desired output of the anomaly detection method is an important consideration in selecting a specific method to apply. Typically, the output of an anomaly detection method is either an *anomaly score*, which indicates how anomalous a data instance is, or the output is a label indicating whether a data instance is classified as an anomaly or nominal. For techniques in which only a label is output, there may exist additional computational steps that may be taken to obtain anomaly scores to indicate a level of *anomalousness* for the data instances. An anomaly score may be taken as a measure of *abnormality* of a data instance. Often anomaly detection methods that output an anomaly score require some threshold to be set to ultimately detect the anomalies. As one of the primary objectives in applying anomaly detection methods to ATM system operational data sets is discovering non-standard operations to be further analyzed by SMEs, often a label is a sufficient output.

Moreover, there exist *supervised*, *semi-supervised*, and *unsupervised* machine learning algorithms that may be leveraged for anomaly detection. The key distinction between supervised and unsupervised machine learning methods is the presence of *labels* within the data set. *Supervised anomaly detection* methods are utilized when the data set is la-

beled such that the anomalous data instances have been previously identified. Supervised anomaly detection techniques are often leveraged in the case where an online anomaly detection algorithm implementation is desired. In this case, a classifier may be trained with the labeled data set, then the trained model is applied to streaming data to detect anomalous data instances in real time. *Semi-supervised anomaly detection* methods are utilized in cases in which it is known that the entire training data set contains data instances that all have the same label (often nominal). The premise is that the model is trained on a nominally-labeled data set such that when the model is fed new, or previously unseen, data instances, such as in the case of an online application, then the model is able to detect anomalies in the new instances. *Unsupervised anomaly detection* methods are utilized when no labels are present in the data set, i.e. it is not known which data instances are nominal or anomalous. Unsupervised learning methods are leveraged to "draw interference from data sets consisting of input data without labeled responses" [96]. There is typically some measure of *similarity* or *distance* that is evaluated to detect data instances that are outlying or anomalous. ATM system operational data is typically *unlabeled* with respect to anomalies, i.e. there does not typically exist a parameter in the data sets that explicitly indicates whether a data instance is an anomaly or not. Accordingly, the ATM system anomaly detection task is often formulated as an unsupervised anomaly detection problem. Thus, in this research, the term *anomaly detection* is intended to refer to *unsupervised anomaly detection*.

Applying any type of anomaly detection method relies on the selection and specification of relevant metrics of interest. A wide variety of anomaly detection methods exist within the aviation literature, where each may select differing metrics on which to perform the anomaly detection. These metrics may either be those directly extracted from the data set or associated derived metrics. A prominent issue arriving aircraft face is the management of energy states [97]. For instance, an aircraft may possess excess energy due to it being too high on glide slope (high potential energy) or in instances in which a high tail wind

is experienced (high kinetic energy) [97]. Thus, as proper energy management plays a significant role in the success and safety of aviation approaches, energy metrics are often derived from trajectory data to support analysis. Puranik et al. [98] complete a survey of the existing literature on energy management in aviation operations and identifies energy metrics that are most relevant. These energy metrics include Specific Potential Energy (SPE), Specific Kinetic Energy (SKE), Specific Total Energy (STE), and their respective rates - Specific Potential Energy Rate (SPER), Specific Kinetic Energy Rate (SKER), and Specific Total Energy Rate (STER).

Due to the importance of aircraft position (spatial states) and aircraft energy state(s), metrics indicating an aircraft's position and/or energy state are most often selected to be included in an analysis, where this research refers to these metrics as *spatial metrics* and *energy metrics*, respectively. Spatial metrics are those that describe an aircraft's position, which may include longitude, latitude, altitude, or heading/track angle. Energy metrics are those that may be derived from the position and speed data. These metrics indicate the energy state of the aircraft, such as SPE, SKE, STE, and their respective rates. However, the anomaly detection methods presented within the aviation literature to detect anomalies in ADS-B data and associated derived metrics typically detect anomalies in either the spatial metrics *or* energy metrics. Therefore, this research (including Corrado et al. [99]), makes a distinction between anomalies detected in spatial metrics and those detected in energy metrics.

The detection of anomalies in trajectories' spatial metrics typically occurs through the identification of air traffic flows. This detection occurs as it is occassionally observed (depending on method employed) that there exist trajectories that do not appear to entirely align with an air traffic flow. Much of the air traffic flow identification literature report a certain percentage of the trajectories that have been detected as outliers, where these trajectories may be considered *spatial anomalies*. As defined by Corrado et al. [99], the concept of a *spatial anomaly* is defined as "*a trajectory whose spatial metrics do not conform to*

an identified set of air traffic flows representing standard spatial operations". On the other hand, as defined by Corrado et al. [99], the concept of an *energy anomaly* is defined as "*a trajectory within an air traffic flow whose energy metrics do not conform to standard energy operations*". Figure Figure 2.6 displays a notional representation of spatial and energy anomalies per the definitions provided. The notional spatial anomaly does not conform to any of the identified air traffic flows (indicated by the different colors). Similarly, the notional energy anomaly presents an unusual STE profile compared to the other trajectories within the flow, according to both the 25^{th} to 75^{th} and 5^{th} to 95^{th} percentile bounds.

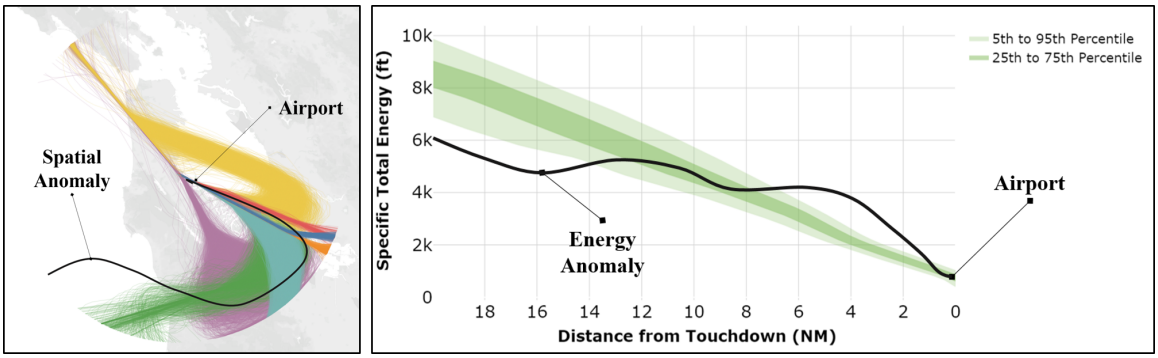

Figure 2.6: Notional Depiction of Spatial and Energy Anomalies Among ADS-B Trajectory Data for Arriving Aircraft at San Francisco International Airport within Four Months of 2019 [99]

As communicated, spatial anomalies are typically detected as a byproduct of the air traffic flow identification exercise, where section 2.1 details these methods. Though, not all methods presented in section 2.1 provide the ability to detect anomalies, such as those leveraging the k-Means clustering algorithm. However, nearly all recent air traffic flow identification methods provide the capability to identify outliers.

The majority of methods reviewed in Basora & Olive's [49] up-to-date review of the recent advances in anomaly detection methods applied to aviation may be classified as energy anomaly detection methods. Generally, when *anomaly detection* is referred to in aviation literature, the implication is that the anomaly detection methods are applied to either energy metrics derived from trajectory data or non-trajectory metrics that exist within FOQA data sets. While anomaly detection in FOQA data sets is not the focus of this

work, it is important to introduce a few relevant and prominent methods as ADS-B data energy anomaly detection methods are inspired by these methods. Subsequently, anomaly detection methods applied to ADS-B trajectory data are reviewed.

Das et al. [100] present one of the first and most prominent anomaly detection methods, multiple kernal anomaly detection (MKAD). MKAD detects anomalies in heterogeneous sequences of both continue and discrete FOQA data metrics. The novelty of MKAD is the ability to successfully and simultaneously examine both continue and discrete metrics [100]. Li et al. [101] introduce another prominent anomaly detection method, ClusterAD, which is a clustering-based method that leverages the DBSCAN algorithm to detect anomalies in FOQA data metrics. The fraction of outliers (anomalie) returned by ClusterAD is specified a priori such that the DBSCAN input parameters are systematically varied such that the specified fraction of outliers is attained [101]. Li et al. [102] present an extension of ClusterAD, ClusterAD-DataSample, which aims to detect instantaneous anomalies in FOQA data leveraging a GMM. Sheridan et al. [47] present a DBSCAN-based method for anomaly detection in the approach phase.

As introduced, Puranik et al. [25] define energy metrics to be computed from trajectory data. Puranik et al. [103], Puranik & Mavris [44, 94], and Puranik [25] derive the defined energy metrics and implement procedures to detect flight-level and instantaneous anomalies in general aviation operational trajectory data. Specifically, Puranik [25] leverages DBSCAN and support vector machines (SVMs) to detect anomalies in both departing and arriving general aviation aircraft trajectories. Considering commercial aviation operations, Kim & Hwang [104], Deshmukh & Hwang [64], and Deshmukh [50] propose TempAD, which designed to provide human-readable and easily interpreted formulas related to the bounds of normality. TempAD leverages DBSCAN as a data processing step to identify air traffic flows [50]. TempAD is applied to detect anomalies in the vertical dimension (altitude), the speed dimension (ground speed), and energy metrics such as STE and SPER [50, 64].

Jarry et al. [105] detect energy anomalies through application of a Functional Principal Components Analysis (FPCA)-based approach which is combined with a sliding window and an outlier scoring capability. Jarry et al. [97] expand on [105] by leveraging FPCA combined with HDBSCAN and Global-Local Outlier Score from Hierarchies (GLOSH) to detect energy anomalies within the terminal airspace. The usage of DBSCAN-based methods is commonplace within the aviation anomaly detection literature. Finally, Corrado et al. [106] present a deep learning-based framework to detect anomalies within the terminal airspace considering a plethora of metrics in ADS-B data and associated derived metrics as well as including weather and traffic/congestion data. Specifically, Corrado et al. [106] leverage autoencoders to detect anomalies in the fused data set.

Similar to the air traffic flow identification/spatial anomaly detection methods that leverage clustering-based methods, the clustering-based energy anomaly detection methods reviewed similarly rely on re-sampling variable-length trajectories to be n-dimensional vectors of trajectory points. However, the scope of the data considered in energy anomaly detection applications often differs. The data processing typically depends on a distance- or time-based threshold for cutting off the full trajectory data before re-sampling. For instance, Das et al. [100] used the FOQA data recorded below 10,000 ft mean sea level and Li et al. [101] used the FOQA data recorded during the final six nautical miles of the approach. Additionally, Deshmukh & Hwang considered trajectory data 20 nautical miles from touchdown [64] and Jarry et al. [105] considered trajectory data 25 nautical miles from touchdown, both establishing distance-based thresholds to process the data set.

In reviewing the literature, it is observed that there is a relative shortage of approaches that consider both spatial and energy metrics in the same analysis. As mentioned, air traffic flow identification is sometimes performed as a data processing step to identify groups of trajectories following a similar spatial path to create smaller data sets on which to perform energy anomaly detection. Trajectories belonging to the same air traffic flow tend to correspond to a specified structured operation defined by ATC, i.e. specified headings, altitudes,

speeds, and/or waypoints. For each structured operation, the approach and climb paths mandated by ATC often have distinct energy profiles. To illustrate this concept, Figure 2.7 displays the STE profiles (25^{th} to 75^{th} percentile values) of three different air traffic flows identified at KSFO in 2019. While there exists some overlap, it is evident these profiles generally differ. Accordingly, the additional spatial context (i.e. air traffic flow to which a trajectory belongs) enables an energy anomaly detection algorithm to yield less "conservative" results. For instance, as noted by Deshmukh [50], applying energy anomaly detection methods to an entire arriving aircraft data set without contextual spatial information may result in anomalous trajectories remaining undetected as their states are embedded within the entire data set. Referring to Figure 2.7, while the energy states of Flow 1 and Flow 2 overlap, there are regions where they, distinctly, do not intersect, whereas Flow 3 tends to exist more in the overlap region between Flow 1 and Flow 2. Thus, an energy anomaly for a trajectory associated with Flow 1 in the lower STE range of Flow 2 would likely remain undetected if the entire data set were considered in the application of an anomaly detection method.

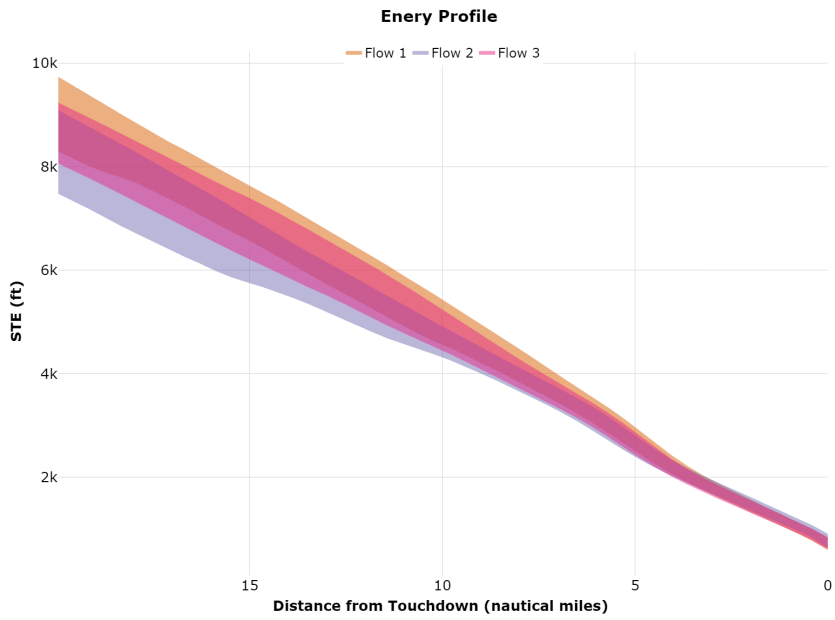

Figure 2.7: Distinct Energy Profiles of Different Air Traffic Flows Considering Four Months of Operations at San Francisco International Airport in 2019

The literature does not make it clear whether whether trajectories detected as spatial anomalies (in the air traffic flow identification data processing step) are of any further consequence with respect to their energy management. Similarly, there is no consensus on the spatial trajectories of aircraft experiencing poor energy management. It is relevant to analyze the energy states of all trajectories, even those that have previously been detected as being spatially anomalous, due to the paramount importance of proper energy management. Moreover, if spatial anomalies are, in fact, considered with respect to energy anomaly detection, it is unclear how they are considered, which is relevant with respect to evaluating the energy anomalies ultimately detected. For instance, it is of interest to ATC and other ATM system operators to know whether a trajectory not conforming to the identified standard spatial operations has a higher likelihood of experiencing energy states that are off-nominal, as this could pose a safety risk. An exploration of the relationship between spatial and energy anomalies is currently lacking in the literature, yet offers opportunities to obtain actionable ATM system insights that may be useful in planning and decision-making efforts. It is ideal for ATM system operators to understand the context in which spatial anomalies should be placed such that the potential risks or impacts of certain instructions that may result in an aircraft deviating from a standard spatial path may be assessed.

Additionally, in the context of the ATFM, the relationship and any interdependencies between spatial and energy anomalies may be significant. Further, knowledge of any relationships and/pr interdependencies is especially relevant as the ATM system is modernized, and a shift toward the more efficient TBO occurs. Therefore, a quantitative analysis of the relationship between spatial and energy anomalies is required to support current and future ATM system planning and decision-making However, current methods to energy detect anomalies in metrics derived from ADS-B data either do not consider or do not elaborate on their consideration of spatial anomalies and, therefore, no knowledge of the relationship between spatial and energy anomalies exists. Considering these observations leads to the

identification of **Gap 2**:

Gap 2

A quantitative analysis of the relationship between spatial and energy anomalies detected in arriving aircraft ADS-B data.

2.3 Airspace-Level Analysis

As mentioned, the aviation literature related to airspace-level analysis, considering that airspace-level refers to the aggregation of the operations of all aircraft within an airspace, is limited. However, it is asserted that the development of a method to perform an airspace-level analysis is of equal importance to the development of flight-level analysis methods. While analysis at the flight level (i.e. analysis of single trajectories) provides crucial knowledge to ATM system operators, planners, and decision-makers, it is important to put into context that each individual trajectory is a product of a larger airspace system, i.e. the individual trajectories are not independent. As discussed in chapter 1, it is not possible to improve the safety, capacity, and efficiency of ATM systems without analyzing ATM systems at the airspace level.

To adequately identify opportunities for ATM system safety, capacity, and efficiency improvement requires an understanding of the current airspace-level operational patterns. An airspace-level operational pattern, or, simply, *operational pattern*, may be a combination of air traffic flows, or an airspace configuration, that is observed regularly. Therefore, the *identification of operational patterns* is an important task in an airspace-level analysis. Specifically, due to the constrained nature of the terminal airspace, identification of operational patterns is of paramount importance.

Unlike an analysis at the flight level, where there exists time-series trajectory data associated with each flight that may be directly analyzed, an analysis of airspace-level operations requires that some *time interval* be specified over which to aggregate the time-series

trajectory data of all flights operating within the airspace. The aggregation of the time-series trajectory data of all aircraft operating within an airspace during a specified time interval comprises what may be referred to as an *airspace-level operational state*, or, simply, *operational state*. It is important to identify which operational pattern an operational state belongs to. However, there are inevitability time intervals in which the operational state either is in transition between operational patterns or is not in transition between operational patterns, yet does not conform to any of the identified operational patterns, i.e. may be considered to be anomalous. In the context of performing an airspace-level analysis to inform the improvement of safety, capacity, and efficiency of ATM systems, these operational states are likely the ones of most interest for further analysis by SMEs, planners, and decision-makers. Therefore, the **characterization of operational states** is an additional important task in an airspace-level analysis.

Thus, identification of the operational patterns for arriving aircraft and characterization of operational states during specified time intervals for arriving aircraft is required to obtain actionable insights useful in the context of broader ATM system planning and decision-making. The aviation literature is reviewed regarding the current state of airspace-level analysis in the context of airspace safety, capacity, and efficiency. Additionally, a discussion on the aviation literature related to the identification of operational patterns and characterization of operational states is presented.

In the context of capacity and efficiency related to airspace-level analysis, there exist some data sources containing metrics that attempt to quantify/characterize some "operational states" within the entire terminal airspace system. For instance, the FAA's Aviation System Performance Metrics (ASPM) database includes traffic condition information for terminal airspace systems such as the number of arrivals, percentage of on-time airport arrivals, etc. for a specified time interval [107]. However, these metrics are quite simple. The ASPM metrics are more so aggregations of flight-level metadata than aggregations of aircraft trajectories' time-series metrics, hence the previous placement of "operational

states" in quotations. In the context of safety analysis, there similarly exist very limited data sources containing aggregated metrics. While is often possible to ascertain information related to accident or incident rates for a specified time interval, these metrics are not especially useful in the context of assessing the impact of flight-level operations on the operational patterns or characterization of operational states at the airspace level as incidents and accidents are typically rare events. Some airlines, airports, or ANSPs gather statistics regarding less severe events on one fleet of one airline [108], yet these statistics are ,similarly, more-so aggregated flight-level metadata. Several system "safety metrics", such as Safety Performance Indicators, Safety Indexes, the Aerospace Performance Factor, etc., have been defined for various airspace systems [109, 110, 111, 112]. Though, most of these metrics lack a comprehensive method to aggregate the time-series data of all aircraft trajectories' within an airspace and rely heavily on the input of SMEs. Therefore, existing metrics related to capacity, efficiency, and safety of an airspace system are not comprehensive enough to identify operational patterns nor characterize operational states that are an aggregation of all flight-level operations within an airspace.

Mangortey et al. [113] present a method to cluster daily operations at an airport, where metrics such as the number of diversions, Ground Stops, departure delays, etc. are leveraged to characterize terminal airspace "operational states" as belonging to a specific category on a daily basis. However, the metrics used are not an aggregation of the operations of all aircraft within a terminal airspace during a specified time interval, i.e. an aggregation of all aircraft trajectories' time-series metrics. Rather, these metrics are more-so an aggregation of flight-level metadata. Hence, again, the placement of "operational states" in quotations. Moreover, the time scale (entire day) of this characterization is quite large, as operations may vary significantly throughout a single day, and it is important to be able to identify specific time intervals through a day that may be abnormal for further analysis. To conclude Mangortey et al.'s [113] systematic approach, a prediction model is presented. This prediction model is leveraged to predict the category (cluster), or "operational state",

that a daily operation belongs to. Despite the limitations noted, it is important to note the potential benefits the development of a prediction models have, namely, the ability to improve tactical decision-making.

On the other hand, a few methods do exist that attempt to aggregate the time-series behavior of multiple aircraft within a terminal airspace into operational states from which to draw airspace-level conclusions. For instance, Enriquez [75] builds on the work of Enriquez & Kurcz [81] and presents a method to identify temporally persistent flows in the terminal airspace using spectral clustering methods. For each time period in a set of time intervals, the spectral clustering procedure presented in [81] is leveraged to cluster trajectories and identify air traffic flows, where a "nominal line" is computed for each air traffic flow as the point-wise median of all trajectories assigned to the air traffic flow [75]. Across all time intervals, the nominal lines for all air traffic flows are added to a feature vector matrix for the spectral clustering algorithm presented in [81] to again be applied. Thus, air traffic flows that are persistent in time are able to be identified [75]. However, despite being concerned with time intervals initially, the time intervals have little meaning in the output of the method as demonstrated applications include identification of irregular terminal airspace traffic (groups of trajectories) and identification of the required number of RNAV procedures in the NAS [75]. Further, while this work does implement a computation of the nominal lines, the nominal lines do not represent an aggregation of the behavior of *all* aircraft operating within a terminal airspace during a specified time interval, nor all aircraft either arriving or departing. Rather, nominal lines represent an aggregation of multiple aircraft operations within only a single air traffic flow during some time interval. Thus, this work does not present the identification of operational patterns nor the characterization of operational states.

Murça et al. [58] introduce the concept of a *Resource Use Matrix*, which is a compact representation of the air traffic flows throughout a day of operations intended to enable comprehensive statistical analyses of airspace use for very large data sets. The Resource

Use Matrix's rows correspond to the air traffic flow identified and columns correspond to a time interval, where the entries of the matrix indicate how many trajectories were associated with an air traffic flow during a time interval [58]. The Resource Use Matrix is leveraged to identify "Resource Use Patterns", or recurrent modes of operations (operational patterns), by clustering the columns of the Resource Use Matrix with the k-Means clustering algorithm [58]. However, as indicated previously, the use of the k-Means clustering inherently has limitations. Specifically, this algorithm is unable to identify outliers, which may be operationally significant (in the context of characterizing operational states), and requires the number of clusters to be set a priori. Further, there may be those time intervals in which the terminal airspace system is transitioning between two distinct operational patterns (a transition operational state) that may be of interest to ATM system operators, planners, and decision-makers.

Murça [11] and Murça & Hansman [83] build on [58]. Murça [11] and Murça & Hansman [83] apply a hierarchical clustering algorithm to a set of "flow vectors", which are the columns of a "flow matrix", which is similar to the Resource Use Matrix in that it indicates which air traffic flows are active during specified time intervals. In these efforts, a hierarchical clustering algorithm is applied to reveal primary operational patterns (analogous to the recurrent modes of operation in [58]) in metroplex terminal airspace systems [11, 83]. However, time intervals in which a terminal airspace operational state does not align with a primary pattern, such as a transition or anomalous operational state, are not considered.

In addition to the identification of the primary operational patterns, the methods presented in [83] and [11] further enable prediction of the these operational patterns. Specifically, Murça & Hansman [83] assess the prediction accuracy of three different machine learning classification algorithms: random forests, support vectors machines, and multinominal logistic regression. These predictions were made at various points within an eight-hour planning horizon [83], where the one-hour planning horizon resulted in the highest average prediction accuracy. Due to slightly better predictive performance, Murça & Hans-

man [83] selected the random forest algorithm to predict operational patterns.

Additionally, Murça [114] presents a method to identify air traffic flows and subsequently predict which air traffic flows within a terminal airspace are active using weather data. Murça [114] leverages a GMM for the prediction of active air traffic flows. However, this method is motivated by identifying and predicting sections of the airspace that may be available for AAM operations by assessing the activity of individual air traffic flows [114], rather than identifying the overall operational pattern of the entire terminal airspace system and/or characterizing operational states.

Additionally, an overarching limitation of the common operational pattern identification methods presented by Murça [11], Murça & Hansman [83], and Murça [114] is that the threshold for declaring an air traffic flow as being "active" is set at one flight. Therefore, these works do not consider the *distribution* or *density* of aircraft trajectories, which may be relevant in ATM system operation, planning, and decision-making. However, the relevance of the ultimate capability to predict operational patterns is noted.

Finally, Zhong et al. [115] present a clustering-based method to analyze terminal airspace "operation status" based on spatial trajectory "characteristic points". However, this method does not provide insight into distinct operational patterns nor operational states of the entire terminal airspace system during specified time intervals. Further, unlike the methods presented by Mangortey et al. [113], Murça & Hansman [83], Murça [11], and Murça [114], no ability to predict the operation status is included [115]. Therefore, this leads to the identification of **Gap 3**:

Gap 3

An airspace-level method to identify operational patterns, characterize operational states (such as those that are nominal, transitional, or anomalous in nature), and predict operational patterns considering the operations of arriving aircraft.

2.4 Summary

The literature review was motivated by gaining an understanding of the existing data-driven methods and approaches to analyze ATM system operations at both the flight and airspace levels, related to the three steps of the proposed methodology, such that gaps may be identified. First, a review of methods to identify air traffic flows was completed. It was identified that there exists a gap related to the existence of a reliable method to identify air traffic flows within the terminal airspace that takes the converging nature of air traffic flows into account and is extensible to real-time applications (**Gap 1**). Next, a review of anomaly detection methods was presented, where an important distinction was made between spatial anomalies and energy anomalies detected in ADS-B data. Additionally, a gap was identified related to the existence of a quantitative analysis of the relationship between spatial and energy anomalies (**Gap 2**). Finally, a review of methods to analyze airspace-level operational patterns and operational states was completed. It was noted that much less aviation literature exists related to this topic, where the methods that are presented have several limitations. Therefore, a gap was identified related to the existence of an airspace-level method to identify operational patterns, characterize operational states (as nominal, transitional, or anomalous in nature), and predict operational patterns considering arriving aircraft (**Gap 3**). Consideration of the **Overarching Research Question** and the review of the relevant existing aviation literature leads to the formulation of the following **Overarching Research Objective**:

Overarching Research Objective

Contribute to the development of an offline data-driven methodology to be applied to ADS-B data for arriving aircraft that enables and performs:

- The reliable identification of air traffic flows
- A quantitative investigation of the relationship between spatial and energy anomalies
- At the airspace level, the identification of operational patterns, characterization of operational states, and prediction of operational patterns

The specific contributions to each step in the proposed methodology with respect to the existing approaches related to each step are specified in Figure 2.8 as well as the overall collective contribution of implementation of the three steps.

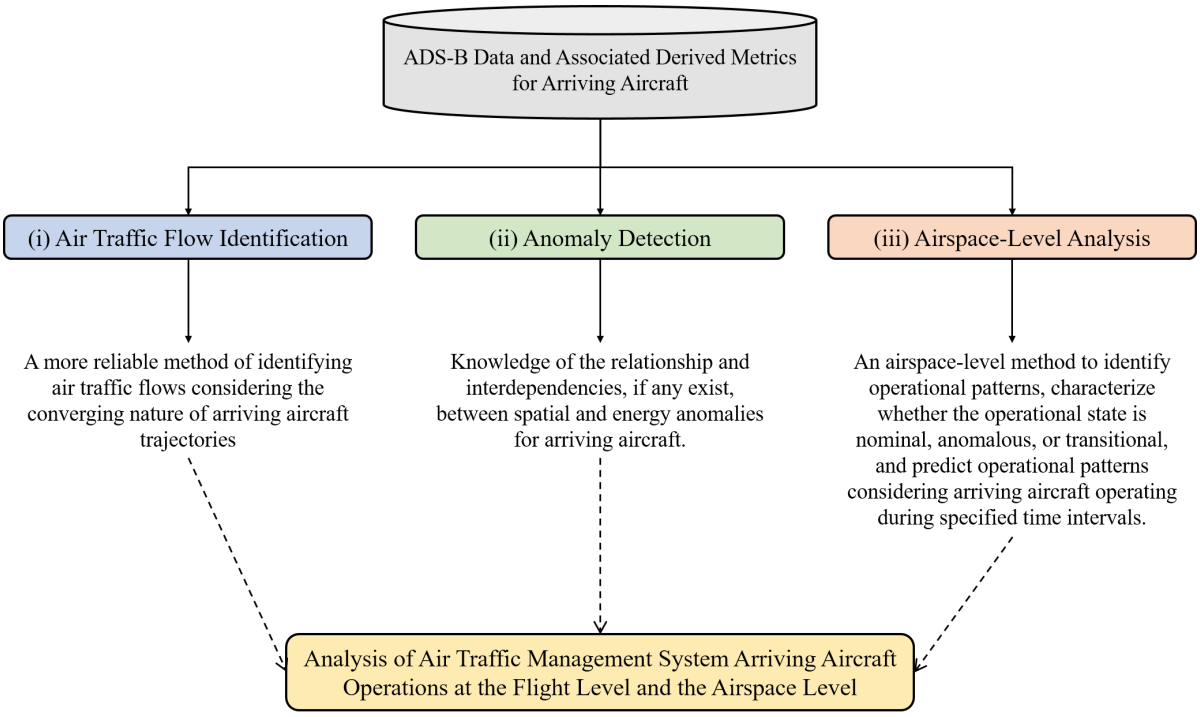

Figure 2.8: Contributions to the Proposed Offline Data-Driven Methodology to be Applied to ADS-B data for Arriving Aircraft with Respect to the Existing Approaches

An overview of the structure of this thesis presented thus far is displayed in Figure 2.9. The four observations made in chapter 1 lead to the formulation of the **Motivating Research Question**, where an **Overarching Hypothesis** is formulated. A literature review inspired by reviewing the state-of-the-art with respect to the three outlined steps in the proposed methodology leads to the identification of three gaps, and, ultimately, the formulation of the **Overarching Research Objective** of this thesis.

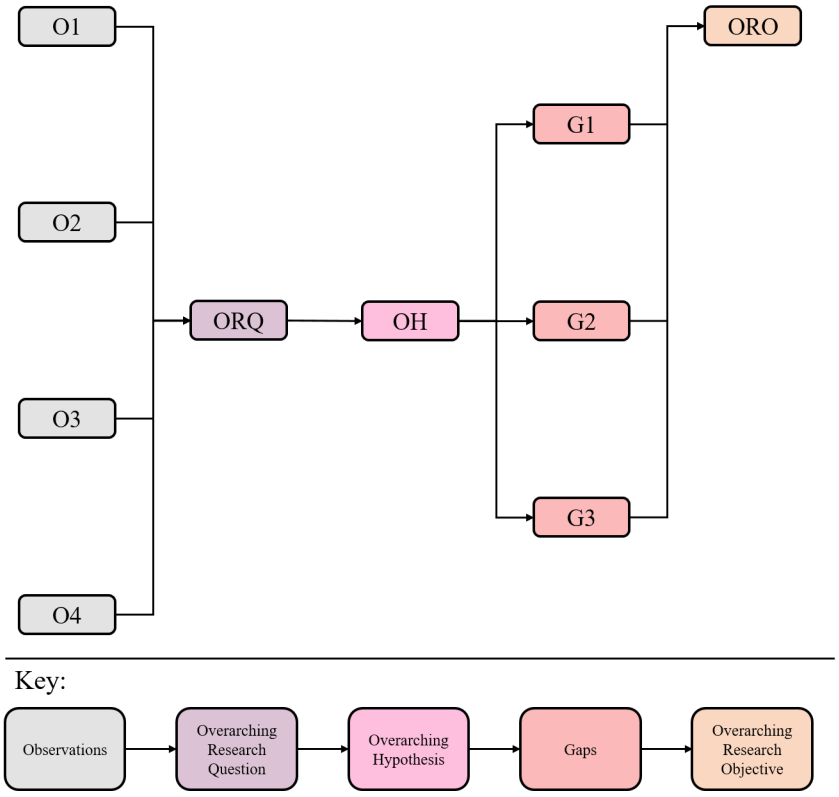

Figure 2.9: Thesis Structure: Literature Review

CHAPTER 3

FORMULATION

An offline data-driven methodology to be applied to ADS-B data is proposed that is comprised of three steps, where a specific gap to be filled is identified related to each of the three steps. The first step in the proposed methodology is **Air Traffic Flow Identification**, where the primary contribution of filling the gap associated with this step is a more reliable method of identifying air traffic flows considering the converging nature of arriving aircraft trajectories. Next, is the **Anomaly Detection** step, where filling the gap associated with this step contributes knowledge of the relationship and interdependencies, if any exist, between spatial and energy anomalies detected in arriving aircraft ADS-B data. Finally, the **Airspace-Level Analysis** step concludes the proposed methodology, where filling the gap associated with this step contributes an airspace-level method to identify operational patterns, characterize whether an operational state is nominal, anomalous, or transitional, and predict operational patterns considering arriving aircraft operating during a specified time interval. This chapter details the formulation of the appropriate research questions and hypotheses related to filling the gaps identified.

3.1 Air Traffic Flow Identification

The objective to fill **Gap 1** leads directly to the formulation of **Research Question 1**:

Research Question 1

How can existing air traffic flow identification methods be modified to consider the converging nature of arriving aircraft trajectories and the requirement of being extensible to real-time applications such that air traffic flows are more reliably identified?

First, the objective of being extensible to real-time applications is addressed. Perform-

ing a recursive clustering of full trajectories or trajectory segments or points is not conducive to an extension to a real-time application. Therefore, methods that require recursive clustering (for instance, [54, 63, 66]) are not considered. Additionally, applying a clustering algorithm and implementing a computationally expensive distance function is not conducive to an extension to a real-time application. Thus, distance functions such as the SSPD, used by Basora et al. in [73], are not considered.

In the aviation literature, the application of DBSCAN or HDBSCAN with the ED function is most commonly presented. Provided two n-dimensional trajectory vectors, T^i and T^j containing x and y features, where $T^i = [(x_1^i, y_1^i), (x_2^i, y_2^i), ..., (x_n^i, y_n^i)]$ and $T^j = [(x_1^j, y_1^j), (x_2^j, y_2^j), ..., (x_n^j, y_n^j)]$ the ED computation proceeds as:

$$D_{ED}^{i,j} = \sqrt{\sum_{k=1}^{n} [(x_k^i - x_k^j)^2 + (y_k^i - y_k^j)^2]},$$

where x_k and y_k indicate an aircraft's horizontal position relative to longitude and latitude measurements, respectively, at point k in the sequence of points comprising the aircraft's trajectory.

As a major limitation related to using the ED function is the standard length requirement, several other distance functions have been developed to overcome this limitation. Warping-based distance functions such as Dynamic Time Warping [116], Longest Common Subsequence [117], Edit Distance on Real Sequence [118], and Edit distance with Real Penalty [118], enable the comparison of trajectories of different lengths; however, they are based on a one-to-one comparison between sequences, so a choice of a particular series to be used as a reference to match all others must be selected [89]. The selection of a reference trajectory is not straightforward for the air traffic flow identification task. There also exist shape-based distance functions to enable the comparison of trajectories with different lengths such as the Hausdorff distance [119], Fréchet distance [120], and the previously-introduced SSPD [89]. The Hausdorff and Fréchet distances are limited as they

may fail to compare trajectories as a whole [89], while the SSPD is computationally expensive to evaluate, as mentioned [54]. Therefore, despite the limitation of the standard length requirement, currently, the ED is best-suited for the air traffic flow identification task.

However, the challenge of taking into account the converging nature of arriving aircraft trajectories persists with the current use of the ED. Recall, a dominant limitation in the use of the ED is the potential for an uneven distribution of distances as aircraft both enter the radius of a terminal airspace and as aircraft approach the airport in their final approach, which may result in inadequate identification of air traffic flows. Therefore, a distance function placing more emphasis on specified trajectory points (such as those in the "center" of the trajectory) in the sequence of re-sampled trajectory points could mitigate this limitation of the ED. The Weighted Euclidean Distance (WED) is an ED-based function that provides the ability to place more emphasis on specific trajectory points.

The WED slightly modifies the ED computation such that an extra multiplicative factor is added. The addition of this multiplicative factor does not significantly increase computation time nor does it prevent the extension to real-time applications. Considering the same two n-dimensional trajectory vectors introduced in section 2.1, T^i and T^j, and a weight vector, W, where $W = [w_1, w_2, ..., w_n]$, the WED computation proceeds as:

$$D_{WED}^{i,j} = \sqrt{\sum_{k=1}^{n} w_k[(x_k^i - x_k^j)^2 + (y_k^i - y_k^j)^2]}$$

The WED introduces an additional trajectory-point weighting term, w_k, which enables a weight, or "importance", to be assigned to each trajectory point. According to the relative position in the sequences of trajectory with respect to the airport location or terminal airspace radius bounds, a weight may be assigned. For instance, Figure 3.1 displays an example of how weights on trajectory points of arriving aircraft within one air traffic flow could be distributed if a WED were applied, where any number of weighting schemes may be ultimately implemented. Considering the weighting scheme displayed, generally, for all

trajectories, the trajectory points are weighted closer to zero beyond the convergence of all trajectories approaching the touchdown point. Additionally, trajectory points are weighted less as aircraft enter the bounds outlined for the terminal airspace. For comparison, Figure 3.2 indicates the "weighting" of points applying the ED, i.e. uniform weighting on all points. Thus, the modification of existing, commonly-applied clustering algorithms, such as DBSCAN or HDBSCAN, with the WED is proposed to consider the converging nature of arrival trajectories within the terminal airspace and the requirement of being extensible to real-time applications.

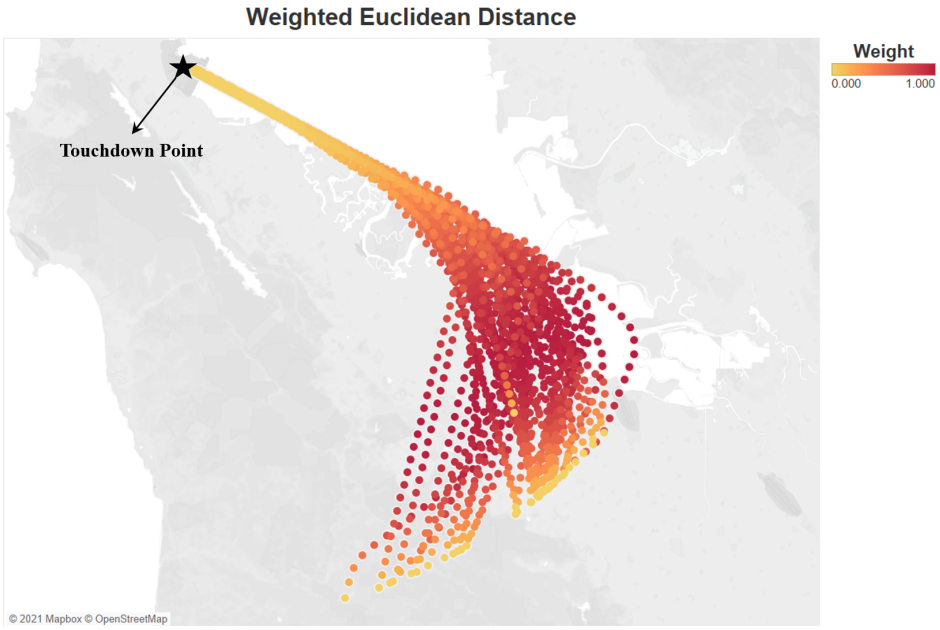

Figure 3.1: Sample Depiction of Weights Assigned to Trajectory Points If The Weighted Euclidean Distance is Applied

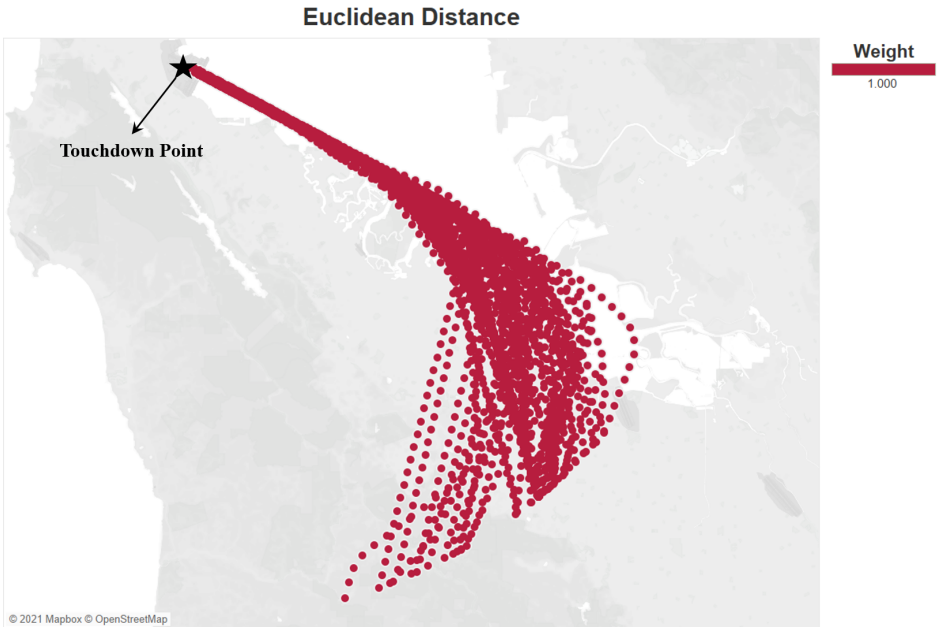

Figure 3.2: Sample Depiction of Weights Assigned to Trajectory Points If The Euclidean Distance is Applied

The selection of the DBSCAN algorithm or the HDBSCAN algorithm to perform the trajectory clustering with the WED requires a review of the core DBSCAN algorithm and how the hierarchical nature of HDBSCAN modifies the DBSCAN clustering output. Figure 3.3 displays a visualization of the key concepts related to DBSCAN implementation.

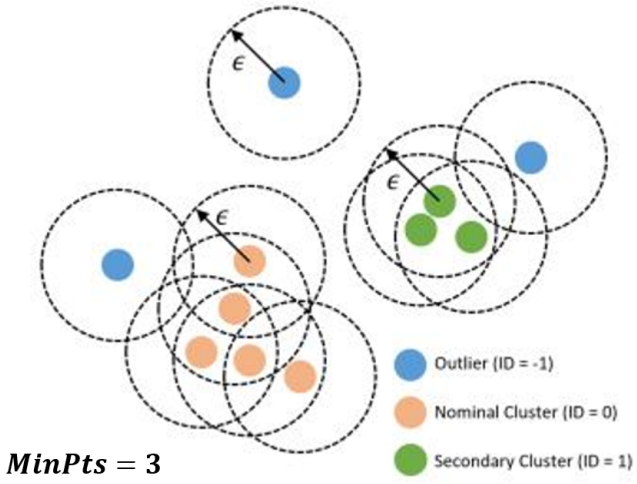

Figure 3.3: DBSCAN Key Concepts Visualization [47]

DBSCAN relies on two input parameters: 1) $MinPts$, which is a minimum number of samples, i.e. trajectories, which is dependent on the homogeneity of operations [47] and 2) ε, which is a distance threshold. The distance between data samples is computed using the specified distance function. DBSCAN locates core samples having high density and expands clusters from them. A prominent advantage of DBSCAN is that the number of clusters k in the data set does not require a priori specification, which is in contrast to algorithms such as k-Means or k-Mediods [80, 121]. DBSCAN enables identification of the air traffic flows despite the presence of outliers, or anomalies [11]. The algorithm is robust to the noise such that outliers not belonging to any of the identified clusters are "allowed" [80], and, in Figure 3.3, are labeled with a -1.

Recently, the application of HDBSCAN [85] in aviation literature in various clustering applications has become more prevalent [73, 97, 122]. Naively, HDBSCAN "extends" DBSCAN by converting it to a hierarchical clustering algorithm, where a technique is then used to extract a flat clustering (set of clusters within any explicit structure that relates the clusters to each other) based on the stability of the clusters [85]. HDBSCAN is an *agglomerative* hierarchical clustering algorithm. In this way, each data point is first considered as its own cluster, then, at each iteration, similar clusters merge with other clusters until one or k clusters are formed [123]. The primary strength of HDBSCAN is the ability to identify clusters of different densities, unlike DBSCAN in which all clusters have relatively the same density. This strength is particularly relevant for applications within the terminal airspace in which air traffic flows often have varying densities, i.e. some flows appear very "tight", while others appear more "spread out" (Figure 2.2 displays this characteristic of air traffic flows of arriving aircraft trajectories). Further, an advantage of HDBSCAN is that it requires only a single input parameter that sets the minimum cluster size (minimum number of samples within a cluster, analogous to the $MinPts$ parameter set when applying DBSCAN) [122]. However, it is also possible to set an optional "smoothing" parameter that is intended to be a measure of how "conservative" the clustering should be [124], where

conservative refers to the fraction of outliers detected. Due to the ability of HDBSCAN to identify clusters (air traffic flows) of varying densities, it is asserted that HDBSCAN is the most appropriate previously-applied clustering algorithm to select to apply with the WED.

Thus, the application of HDBSCAN with the identified current best-suited distance function, the ED, is the proposed existing method to be modified. The modification includes the use of the WED as opposed to the ED as the implemented distance function within the HDBSCAN algorithm. Therefore, **Hypothesis 1** is formulated as:

Hypothesis 1

If HDBSCAN is implemented with the Weighted Euclidean Distance, then air traffic flows are more reliably identified considering the converging nature of arriving aircraft trajectories when compared with the implementation of HDBSCAN with the Euclidean Distance and the procedure is extensible to real-time applications.

The testing and acceptance of **Hypothesis 1** fill the identified **Gap 1**.

3.2 Anomaly Detection

The objective to fill **Gap 2** leads directly to the formulation of **Research Question 2**:

Research Question 2

What relationship exists between spatial and energy anomalies detected in arriving aircraft trajectories?

To enable an analysis of the relationship between spatial and energy anomalies, an approach is required to detect the two types of anomalies. As introduced in section 2.2, specifically Figure 2.7, there generally exists observable differences in the energy profiles of different structured operations, or air traffic flows, of arriving aircraft. Further, it is mentioned in both section 2.1 and section 2.2 that air traffic flow identification is commonly

performed prior to the application of energy anomaly detection methods as a data processing step, i.e. to segment the full data set into smaller data sets on which to separately apply an energy anomaly detection method. Air traffic flow identification methods are also typically those leveraged to perform spatial anomaly detection. Therefore, in an approach to detect spatial and energy anomalies in arriving aircraft trajectories, spatial anomaly detection should be performed leveraging an air traffic flow identification method prior to energy anomaly detection, where the spatial anomaly detection step also enables the segmenting of the full data set into smaller data sets corresponding to each identified air traffic flow.

The successive detection of spatial and energy anomalies should result in trajectories being assigned one of four labels to indicate the "category" of trajectory to which they belong. These categories include:

- Nominal (N), i.e. the trajectory is not detected as anomalous in either the spatial or energy dimension
- Only-Spatial-Anomaly (S), i.e. the trajectory is detected as anomalous in the spatial dimension, yet is not detected as anomalous in the energy dimension
- Only-Energy-Anomaly (E), i.e. the trajectory is not detected as anomalous in the spatial dimension, yet is detected as anomalous in the energy dimension
- Both-Spatial-And-Energy-Anomaly (B), i.e. the trajectory is detected as anomalous in both the spatial and energy dimensions

Of interest is a comparison between rates of occurrences of the four categories of anomalies. Specifically, it is of interest whether any interdependencies exist between spatial and energy anomalies. Therefore, a quantitative investigation into the rates of occurrences of spatial and energy anomalies must be assessed, and **Hypothesis 2.1** is formulated as:

Hypothesis 2.1

If a trajectory is detected as a spatial anomaly, then it is more likely to be detected as an energy anomaly.

In addition, to further understand and continue with analyzing the relationship between spatial and energy anomalies detect in arriving aircraft trajectories an *energy anomaly score*, or simply, *anomaly score*, may be computed for each trajectory within an air traffic flow data set. Computation of an anomaly score provides more context to the "relative anomalousness" of arriving aircraft trajectories. The anomaly score may be computed as the mean ED from a given trajectory's energy metrics to all others in the data set:

$$Score^i = \frac{1}{n} \sum_{j=1, j \neq i}^{n} D_{ED}^{i,j},$$

where $D_{ED}^{i,j}$ is the ED between trajectory i's energy metrics and trajectory j's energy metrics.

Considering only energy-nominal trajectories, it is of interest whether those trajectories in the only-spatial-anomaly category have, on average, greater anomaly scores than those trajectories in the nominal category. For instance, it is important for ATC and/or air crews to understand the consequence of non-conformance to standard spatial operations with respect to energy metrics, even if a trajectory is not necessarily an energy-anomaly. The closeness to the "border" of energy-anomalousness is of interest. Therefore, **Hypothesis 2.2** is formulated as:

Hypothesis 2.2

If only energy-nominal trajectories are considered, then only-spatial-anomaly trajectories have, on average, greater anomaly scores than nominal trajectories.

Finally, if the trajectories are separated by category to which they belong, the distribution of anomaly scores corresponding to each category may be evaluated. Specifically, it

is of interest whether the distribution of the anomaly scores of only-energy-anomalies and anomaly score of both-spatial-and-energy-anomalies have different statistical properties. Investigating this relationship provides insight into whether there should be a distinction between only-energy-anomalies and both-spatial-and-energy-anomalies in the context of "severity" of the anomalies such that these situations may require different mitigation efforts. Therefore, **Hypothesis 2.3** is formulated as:

Hypothesis 2.3

If the statistical properties of the distribution of anomaly scores associated with only-energy-anomaly trajectories and both-spatial-and-energy-anomaly trajectories, then a distinct difference in statistical properties of the distribution is observed.

The testing and acceptance of **Hypothesis 2.1**, **Hypothesis 2.2**, and **Hypothesis 2.3** fill **Gap 2**.

3.3 Airspace-Level Analysis

The first consideration in developing a comprehensive method to identify operational patterns, characterize operational states, and predict operational patterns is the representation of operational states. As introduced in section 2.3, an operational state is an aggregation of the flight-level operations, or time series trajectory data, for all arriving aircraft operating during a specified time interval. Therefore, prior to any identification, characterization, or prediction, an approach to aggregating the ADS-B data for all arriving aircraft operations is required. Specifically, previous work related to airspace-level analyses considers the spatial metrics in ADS-B data [11, 75, 83, 114, 115]. However, in [75], [11], [83], and [114], to perform the analysis at the airspace level, trajectories are first aggregated at the air traffic flow-level, where an air traffic flow centroid trajectory is produced or simply an assignment of trajectories to an air traffic flow to indicate "activity" of an air traffic flow. Thus, there is a loss of information related to the distribution of trajectories within the airspace as the

trajectories that comprise the air traffic flow may be very "tight" or "spread out". Further, flight-level spatial anomalies appear to not be considered in the analysis, yet their inclusion and consideration may be valuable in terms of assessing ATM system arriving aircraft operations at the airspace level. These spatial anomalies may also include arriving aircraft that perform a go-around, where an air traffic flow centroid or assigned to an air traffic flow would not provide any information on these occurrences. Therefore, a data processing step to properly represent operational states for arriving aircraft operating during specified time intervals is required.

The use of air traffic flows as the representations of operational states does not enable consideration of the distribution of trajectories within air traffic flows or the airspace as a whole, including lack of consideration of the operational states of time intervals that may be characterized as anomalous or transitional. Moreover, a limitation of the previous works that leverage air traffic flows as a representation of the terminal airspace operational patterns is that an air traffic flow is considered to be "active" if at least one trajectory is associated with the air traffic flow during the specified time interval. This definition of "active" may result in an operational pattern being identified in which only a negligible fraction of trajectories ever are associated with that flow. Considering the threshold for an air traffic flow to be "active" as one essentially results in air traffic flows containing only one trajectory to have a "weight" equal to an air traffic flow containing ten trajectories. Therefore, it is important when evaluating the validity of the representation of the operational state as the airspace density matrix for the "weight" of different trajectory patterns to be considered as well as the distribution of the trajectories within the terminal airspace.

While it may be possible to consider the boundaries of air traffic flows as well when identifying operational patterns in an attempt to get a better picture of the distribution of the trajectories within an airspace, this approach still would neglect the occurrence of spatial anomalies. Therefore, an approach not reliant on identified air traffic flows is advantageous. Spatial metrics in ADS-B data often refer to the latitude and longitude coordinates, which

are two-dimensional (2D) pairs of coordinates. Derived spatial metrics may also be relevant depending on the application, where the derived spatial metrics may be those resulting from a projection of the latitude and longitude coordinate system onto a specified Cartesian coordinate system. Regardless of the choice of spatial metrics or derived spatial metrics, the coordinate pairs representing a trajectory in the spatial dimension are 2D.

Thus, it is possible to consider a trajectory's spatial metrics coordinate pairs as points on a 2D grid representing an airspace. For each trajectory, it is possible to "fill" in the spaces on a 2D grid that the trajectory passes through. Expressed mathematically, the 2D grid may be considered to be a 2D matrix of zeros, where the "filling" of a grid space, if a trajectory passes through it, could be represented by inserting a 1 into the matrix element corresponding to the grid space. This matrix may be referred to as a *trajectory matrix*. Figure 3.4 displays this concept for a simple sample trajectory (in the slightly darker gray), entering the airspace in the top right and directly arriving at the airport location, which is at the center of the matrix in this example.

0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Figure 3.4: Trajectory Matrix

It is possible to create trajectory matrices for each arriving aircraft trajectory within the terminal airspace. Further, it is possible to then sum the trajectory matrices for all flights

operating within a specified time interval. This matrix may be referred to as the *airspace matrix*. The airspace matrix enables the information for all flights operating within an airspace during a specified time interval to have some representation in the aggregation as well as provides more of an indication of the distribution of trajectories within the airspace as opposed to the basic air traffic flow centroids representation of the aggregation of flights. However, the number of flights arriving at an airport varies throughout the day. Therefore, some time intervals would experience higher maximum values for airspace matrix entries. This could result in the identification of operational patterns being skewed towards identifying groupings of time intervals with a similar number of flights. Thus, a normalization step is proposed in which the airspace matrix entries are divided by the total number of trajectory matrices that were summed to create the airspace matrix. This normalization effectively provides a measure of the *density* of trajectories within an airspace during a specified time interval. This normalized matrix may be referred to as an *airspace density matrix*. This data processing step in generating airspace density matrices to represent operational states is outlined in Figure 3.5.

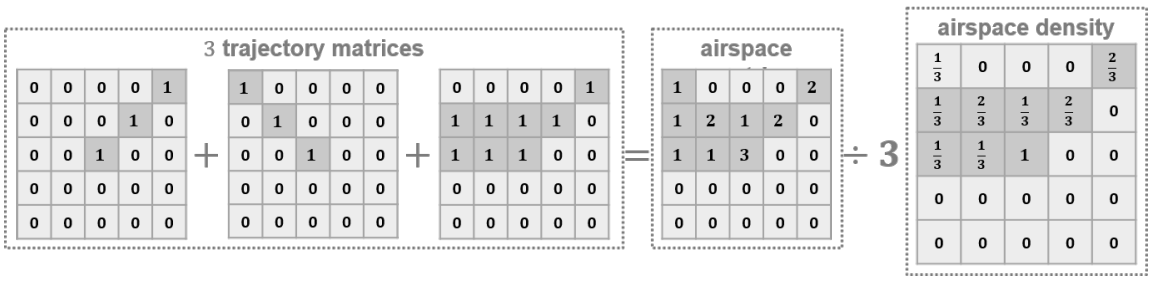

Figure 3.5: Generation of an Airspace Density Matrix for Three Trajectories within a Specified Time Interval

Once the operational states have been represented by the airspace density matrices, a procedure, or procedures, to perform the identification and/or characterization is required. Due to the relationship between operational patterns and operational states, i.e. the characterization of an operational state is dependent upon the conformance of the operational state to the operational patterns identified from the set of all operational states within some time

period comprised of multiple, smaller, time intervals, it is advantageous to develop a single procedure that can both identify operational patterns and characterize operational states. Therefore, considering the objective to fill **Gap 3**, **Research Question 3.1** is formulated as:

Research Question 3.1

What procedure can be implemented to both identify operational patterns and characterize operational states at the airspace level for arriving aircraft operations during specified time intervals?

In previous work, what may broadly be considered to be operational patterns or operational states have been identified or characterized, respectively, leveraging clustering algorithms. For instance, Mangortey et al. [113] leverage and compare the performance of various clustering algorithms, including k-Means. Enriquez [75] applies the spectral clustering method presented in [81]. Murça et al. [58] apply the k-Means clustering algorithm. Finally, Murça [11] and Murça et al. [83] apply a hierarchical clustering algorithm. The identification of operational patterns may be formulated as an unsupervised classification problem, as the objective is to identify the prominent operational patterns from a set of operational states for a defined time period, containing multiple, smaller, time intervals for which the operational states are determined. Clustering algorithms lend themselves well to unsupervised classification tasks.

In the context of characterizing operational states, specifically as nominal or anomalous, the implementation of a clustering algorithm may similarly be appropriate. For instance, leveraging those clustering algorithms that provide the capability to detect outliers, such as DBSCAN or HDBSCAN, may provide insight into characterizing operational states as nominal or anomalous. Therefore, it is assumed that application of the DBSCAN or HDBSCAN algorithm to the airspace density matrix data set of operational states enables both the identification of operational patterns, taken as the clusters identified by the algorithm,

and the characterization of operational states, taken as either nominal (strictly belonging to a cluster) or outlying (either transitional or anomalous). The selection of either DBSCAN or HDBSCAN to perform the clustering task is dependent on the expected uniformity of the density of operational states associated with the identified operational patterns. For the air traffic flow identification task, HDBSCAN was identified as being most appropriate due to the potential for trajectories to be either very "tight" or "spread out" within an air traffic flow. However, the airspace density matrix representation of the operational states inherently contains information regarding density within an airspace. Thus, it is proposed that the DBSCAN algorithm be leveraged in a procedure to both identify operational patterns and characterize operational states.

Though, the challenge of characterizing operational states as being transitional remains. An assumption is made that a robust application of DBSCAN to identify operational patterns would detect the operational states of those time intervals that are experiencing a transition of airspace-level operations from one operational pattern to another as being an outlier, or anomalous. Therefore, by comparing the operational patterns assigned to the operational states corresponding to time intervals directly before and after an operational state that has been detected as being an outlier, it is possible to determine if the time interval has been detected as being an outlier due to having an operational state that may be characterized as transitional. Explicitly, if the operational patterns identified for the operational states corresponding to the time intervals directly before and after an operational state that has been detected as being an outlier are different, then the time interval's operational state may be characterized as being transitional.

However, once transitional operational states have been discovered, it is preferable to remove them from consideration of being either nominal, i.e. strictly belonging to an operational pattern, or outlying, i.e. not conforming to an operational pattern. Thus, it is proposed that a recursive DBSCAN approach be taken to identify operational patterns and simultaneously characterize operational states. Therefore, **Hypothesis 3.1** is formulated

as:

Hypothesis 3.1

If a recursive DBSCAN procedure is implemented such that the clusters indicate identified operational patterns and outlying operational states are characterized as being either anomalous or transitional, where the recursion proceeds until no new transitional operational states are characterized, then operational patterns are identified and operational states are characterized.

Finally, it is important to consider that within the aviation literature, exploration of the ability to predict operational patterns or operational states has often occurred once the operational patterns or operational states have been identified or characterized, respectively [11, 83, 113, 114]. Mangortey et al. [113] leverage a boosting ensemble algorithm to perform the prediction task. Murça & Hansman [83] assess the prediction accuracy of three different machine learning classification algorithms (random forests, SVM, and multinominal logistic regression), where the random forest algorithm is ultimately selected for the prediction task due to slightly superior performance.

Prediction models are typically trained on features not leveraged to perform the classification of operational patterns or operational states. Moreover, while would be valuable to predict anomalous and/or transitional operational states, it is noted that anomalous and/or transitional states correspond to a small percentage of the overall data set. Therefore, prediction of the anomalous and/or transition operational states would be more appropriate potentially as a semi-supervised anomaly detection exercise, which is not the focus of this airspace-level analysis. Thus, the focus in the development of a prediction model is to provide the capability to predict operational patterns.

The ability to train a reasonably accurate prediction model with features derived from certain metric measurements provides insight into the metrics, or features, most related to the observation of a certain operational pattern. Knowledge of the importance of certain

features in the context of predicting operational pattern occurrence may provide valuable insight and/or be a useful capability for ATM system operators, planners, and decision-makers. Further, as presented by Murça [114], as AAM vehicles and concepts are introduced into the NAS, there exists value in predicting airspace availability for AAM. Therefore, considering the objective to fill **Gap 3**, **Research Question 3.2** is formulated as:

Research Question 3.2

How can the operational pattern of arriving aircraft operations during a specified time interval be predicted?

The task of predicting the operational pattern of arriving aircraft operations during a specified time interval may be considered to be a semi-supervised learning problem. A supervised learning classification algorithm is to be trained to predict the operational pattern during a specified time interval, where the labels associated with each set of input features are an output of the unsupervised operational pattern identification procedure.

Before any classification algorithm is trained, the selection of appropriate metrics (features) on which to train a prediction algorithm is critical. The metrics to be used should be derived from recorded data in the time intervals prior to the time interval for which the operational pattern is to be predicted. This is because, in a real-time scenario, the metrics for the time interval to be forecast would not have been observed yet. Traffic and/or congestion metrics, such as those recorded in the FAA's ASPM database may provide diverse airspace-level information. However, it is unlikely there is a strong relationship between metrics such as the number of arrivals/departures, the average delay time, the percent of on-time arrivals/departures, etc. in a previous time interval to the operational pattern of the time interval to be forecast. Murça et al. [53], Murça [11], and Murça [114] specify weather conditions as being observed to affect the conformance of trajectories to their assigned routes. Therefore, features derived from recorded weather measurements are proposed to be leveraged to train a model to predict operational patterns.

This prediction problem is assumed to be a multi-class problem, as there are likely to be more than two operational patterns identified for arriving aircraft within a terminal airspace. Therefore, a multi-class classification algorithm is required. Several algorithms enable multi-class classification, such as artificial neural networks, k-nearest neighbors, naive Bayes, decision trees, and SVMs, to name a few. Often, it is not immediately obvious which classification algorithm will provide the highest testing prediction accuracy. For this reason, for any given classification task, several different algorithms are often trained, and the one resulting in the highest prediction accuracy is selected. For instance, as mentioned, Murça & Hansman [83] compared the performance of three different classification algorithms (random forests, SVMs, and multinominal logistic regression) before ultimately selecting random forests as the appropriate algorithm for the problem. Therefore, **Hypothesis 3.2** is formulated as:

Hypothesis 3.2

If a set of multi-class classification algorithms are trained to predict operational patterns with input features derived from recorded weather measurements and the algorithm providing the most desirable performance is selected, then the operational patterns of a testing data set are able to be predicted with reasonable accuracy.

The testing and acceptance of **Hypothesis 3.1** and **Hypothesis 3.2** fill **Gap 3**.

3.4 Summary

The three steps in the proposed methodology include **Air Traffic Flow Identification**, **Anomaly Detection**, and **Airspace-Level Analysis**, where a gap has been identified related to each step in chapter 2. This chapter detailed the formulation of the appropriate research questions and hypotheses related to the gaps identified. An overview of the mapping of gaps, to research questions, to hypotheses is displayed in Figure 3.6.

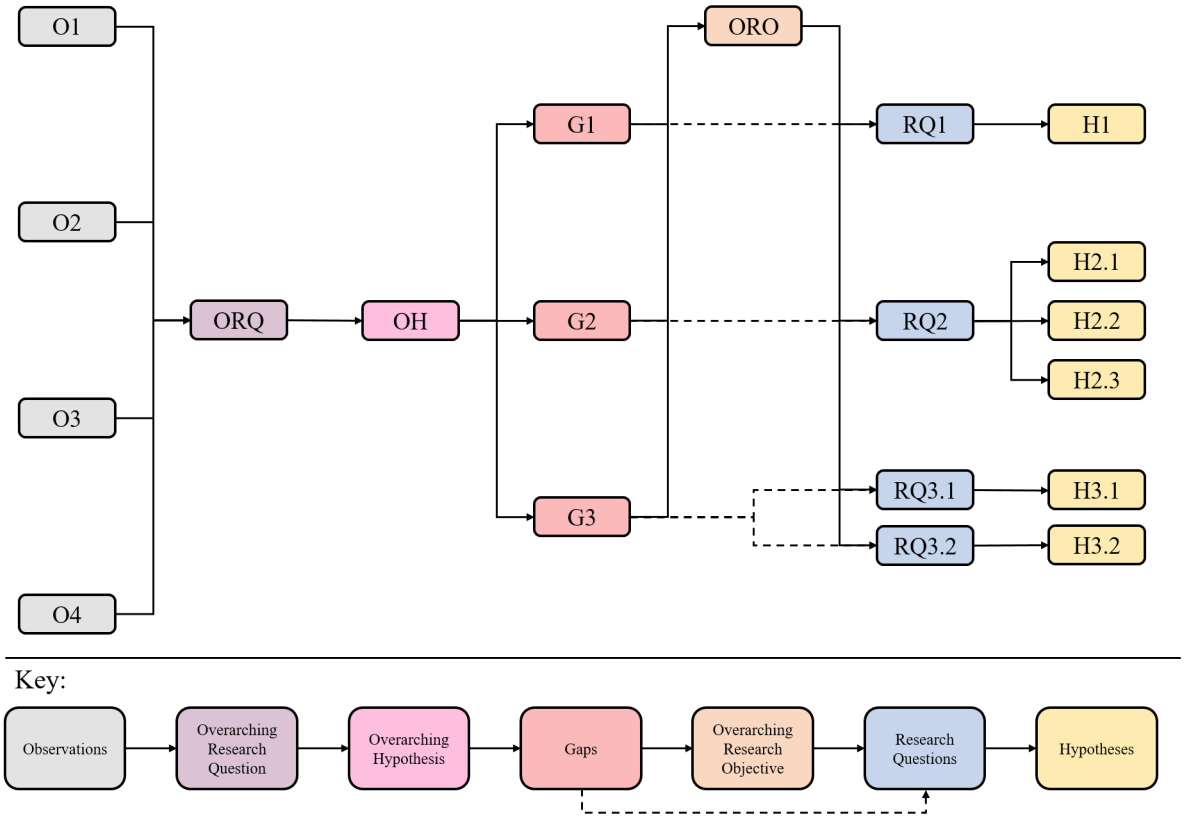

Figure 3.6: Thesis Structure: Formulation

Additionally, an extracted and cleaned ADS-B data set, augmented with appropriate derived metrics, is required to test and validate each of the formulated hypotheses.

CHAPTER 4

EXPERIMENTAL APPROACH

This chapter details the development of three experimental approaches to test and accept each of the hypotheses presented in chapter 3 such that the gaps identified in chapter 2 are filled.

4.1 Air Traffic Flow Identification

Gap 1, which is associated with the **Air Traffic Flow Identification** step in the proposed methodology, is mapped to **Research Question 1**, where **Hypothesis 1** is formulated in response.

Considering **Research Question 1** and **Hypothesis 1**, the objective of **Experiment 1** is stated as:

Experiment 1

Identify air traffic flows in arriving aircraft ADS-B data implementing both HDBSCAN with the Weighted Euclidean Distance and HDBSCAN with the Euclidean Distance.

An experimental approach to identify air traffic flows in arriving aircraft ADS-B data implementing both HDBSCAN with the WED and HDBSCAN with the ED requires two main steps: (i) data processing and (ii) trajectory clustering. Both the data processing and trajectory clustering steps are detailed. Additionally, the acceptance of **Hypothesis 1** is discussed.

4.1.1 Data Processing

To implement an experimental approach, a cleaned and augmented set of ADS-B data and associated derived metrics must be provided. The exact data cleaning procedure employed is dependent upon the source of the ADS-B data leveraged. However, regardless of the data source, specific derived metrics must be computed to augment the data set. In the context of the **Air Traffic Flow Identification** step, two derived additional metrics are relevant. These two metrics are discussed as well as the primary data processing task of re-sampling the data set.

Additional Metrics

As introduced, within the aviation literature, air traffic flow identification is typically performed considering the spatial metrics in ADS-B data. The spatial metrics relevant to the air traffic flow identification task are longitude and latitude states. However, the *scales* of longitude and latitude differ, i.e. longitude and latitude measurements belong to a spherical coordinate system rather than a Cartesian coordinate system. Because both the WED and ED are sensitive to differing magnitudes in feature dimensions, it is desirable to project the measured longitude and latitude coordinates onto an (x, y) Cartesian coordinate system. Therefore, longitude and latitude coordinates are projected onto an (x, y) Cartesian coordinate system according to some specified map projection.

Re-Sampling

It is recalled that implementation of both the WED and ED require n-dimensional trajectory vectors, T^i and T^j to compute a distance between two trajectories. In practice, n-dimensional trajectory vectors are not readily available. Rather, n-dimensional trajectory vectors are often the product of *re-sampling* trajectories of various lengths. In the context of air traffic flow identification, it is typical to consider the trajectories of all aircraft operating within some radius of the airport location. In other words, the record of a single arriving

aircraft trajectory initiates once the aircraft enters the specified radius around the airport location and ends once the aircraft touches down. Therefore, all trajectories within a data set to be used for air traffic flow identification have differing values for the maximum ground track distance from touchdown. Moreover, each trajectory will have a different temporal length.

The primary methods that exist to perform re-sampling are either distance-based or time-based. Distance-based and time-based re-sampling methods tend to generate the same set of points in the case that the aircraft maintains a relatively constant velocity. However, during the approach phase, aircraft velocity is almost never constant. Rather, aircraft velocity tends to monotonically and gradually decrease. Re-sampling based on time is likely to produce an uneven distribution of trajectory points such that many more exist closer to the airport location. Thus, distance-based re-sampling is preferred in which a uniform re-sampling of trajectory points occurs based on an aircraft's ground track distance from touchdown. Implementing a distance-based re-sampling of trajectory points produces trajectory points that are spatially distributed more evenly. Once trajectory points have been re-sampled based on the ground track distance, the trajectory clustering step may proceed.

4.1.2 Trajectory Clustering

As indicated by the objective of **Experiment 1**, air traffic flows must be identified both by implementing HDBSCAN with the WED and by implementing HDBSCAN with the ED. Therefore, this requires that the trajectory clustering step involve the clustering of the (x, y) coordinates that have been re-sampled based on the ground track distance of an aircraft from its touchdown point for all arriving aircraft by applying the HDBSCAN algorithm separately with both the WED and ED.

Additionally, the experimental approach, specifically the trajectory clustering step, must also enable comparison of the implementation of HDBSCAN with the WED versus the state-of-the-art implementation of HDBSCAN with the ED in the context of "more reli-

ably" identifying air traffic flows. However, first, it is important to define what is meant by identifying air traffic flows "more reliably". While visual inspection of plots of the air traffic flows identified by the implementation of HDBSCAN with the two different distance functions provides insight into the performance of the WED versus the ED, a metric to enable a quantitative comparison of the implementation of HDBSCAN with the WED versus the implementation of HDBSCAN with the ED is required.

Considering the aviation literature reviewed in section 2.1, a primary issue with the state-of-the-art methods that are inherently extensible to real-time applications, such as HDBSCAN with the ED, is the lack of consideration of the converging nature of arrival trajectories. Specifically, the distances computed between trajectory points closest to the airport are relatively small, regardless of the air traffic flow to which they belong and the distances between trajectory points at the terminal airspace's defined border may be relatively large, regardless of the air traffic flow to which they belong. Thus, the uneven distribution of distances as aircraft arrive at the airport may skew classification, resulting in the misclassification of trajectories and a "less reliable" identification of air traffic flows.

If a trajectory is misclassified, likely the trajectory is further (distance-wise) from the centroid of the set of trajectories that the misclassified trajectory is assigned to then from the centroid of the set of trajectories to which it should belong. Hence, the mean of the distance between a trajectory assigned to a cluster and its respective cluster centroid computed over the set of trajectories that have been assigned to a cluster is expected to be lower in the case of a "more reliable" air traffic flow identification. Therefore, the mean of the distance between a trajectory assigned to a cluster and its respective cluster centroid computed over the set of trajectories that have been assigned to a cluster, or, simply, the mean distance from the assigned cluster centroid, may be considered the quantitative metric to compare the utilization of the WED versus the ED in the air traffic flow identification task. Due to the common use of the ED in the aviation literature, the mean distance from the assigned cluster centroid is computed using the ED.

Though, it is important to note an important consideration: the fraction of outliers, or anomalies, detected by the HDBSCAN algorithm. It is intuitive that if the implementation of the HDBSCAN algorithm results in a higher fraction of outliers, the mean ED from the assigned cluster centroid will likely be lower. On the other hand, if the implementation of the HDBSCAN algorithm results in a lower fraction of outliers, the mean ED from the assigned cluster centroid will likely be higher. Therefore, when comparing the implementation of HDBSCAN with the WED to the implementation of HDBSCAN with the ED, it is important to ensure the fraction of outliers detected in each implementation is within a similar range. The setting of HDBSCAN parameters is the driver of the resulting fraction of outliers detected for a given distance function. However, implementation of the HDBSCAN algorithm with the same setting parameter settings for the two different distance functions, the WED and the ED, will likely result in differing fractions of outliers detected. Therefore, it is important to consider this in the comparison of the mean ED from the assigned cluster centroid metric.

As mentioned in section 2.1, the implementation of HDBSCAN requires one input parameter indicating the minimum cluster size. Though, there also exists the capability to specify an optional "smoothing" parameter that provides input into how "conservative" the clutsering should be [124]. The larger the setting of th smoothing parameter value, the more samples detected as outliers, and the clusters are restricted to a more dense area [124]. Therefore, setting the value of the minimum cluster size parameter to the same value for both the implementation of HDBSCAN with the WED and the implementation of HDBSCAN with the ED, the optional smoothing parameter may be leveraged to "control" the fraction of outliers detected. Specifically, the HDBSCAN clustering may be performed iteratively for both implementations of HDBSCAN such that some value of the minimum cluster size parameter is set and the smoothing parameter is varied until an acceptable fraction of outliers detected is attained.

"Controlling" the fraction of outliers detected for both the implementation of HDB-

SCAN with the WED and the implementation of HDBSCAN with the ED requires some specification of the range of the fraction of outliers detected in which the returned fraction of outliers from the HDBSCAN implementation may be deemed acceptable. Thus, an iterative clustering may be required for both the implementation of HDBSCAN with the WED and the implementation of HDBSCAN with the ED to ensure the fraction of outliers for the same minimum cluster size parameter setting is within a pre-determined range for each implementation.

An overview of the experimental approach to support the filling of **Gap 1** is presented in Figure 4.1.

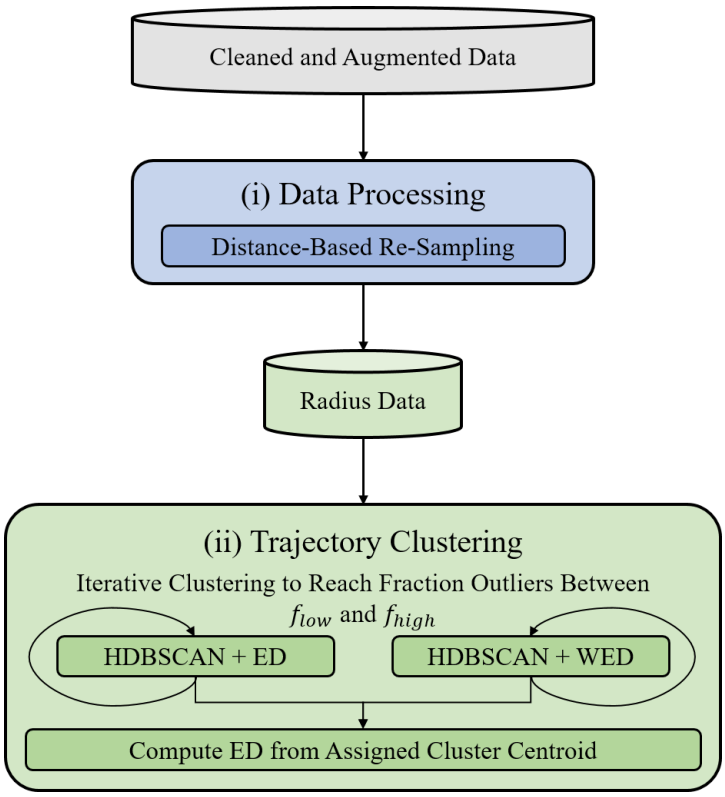

Figure 4.1: Experimental Approach: Gap 1

4.1.3 Hypothesis 1 Acceptance

Hypothesis 1 is accepted if the implementation of the experimental approach described corresponding to **Experiment 1** results in more reliable identification of air traffic flows

in the case in which HDBSCAN is implemented with the WED as opposed to the case in which HDBSCAN is implemented with the ED. As described, a direct comparison of the "reliability" of the two implementations of HDBSCAN may be made by considering the computed mean ED from the assigned cluster centroid metric in the trajectory clustering step of the experimental approach. However, this direct comparison of the mean ED from the assigned cluster centroid is only valid if the fractions of outliers detected for both the implementation of HDBSCAN with the WED and the implementation of HDBSCAN with the ED are within a specified range. Thus, the **Hypothesis 1 Acceptance** may be expressed as:

Hypothesis 1 Acceptance

If the mean Euclidean Distance from the assigned cluster centroid is lower in the case in which HDBSCAN is implemented with the Weighted Euclidean Distance than in which HDBSCAN is implemented with the Euclidean Distance, where each implementation results in the fraction of outliers detected within a specified range, then **Hypothesis 1 is accepted**.

4.2 Anomaly Detection

Gap 2, which is associated with the **Anomaly Detection** step in the proposed methodology, is mapped to **Research Question 2**, where a set of hypotheses (**Hypothesis 2.1**, **Hypothesis 2.2**, and **Hypothesis 2.3**) are formulated in response.

Considering **Research Question 2** and the associated set of hypotheses, to investigate and quantify the relationship between spatial and energy anomalies in arriving aircraft ADS-B data and associated derived metrics, spatial and energy anomalies must first be detected. Therefore, an experimental approach enable acceptance of the set of hypotheses associated with **Research Question 2** is primarily reliant on a procedure to detect spatial and energy anomalies in arriving aircraft ADS-B data and associated derived metrics within

the terminal airspace. Upon detection of spatial and energy anomalies, it is straightforward to test the set of hypotheses associated with **Research Question 2** through statistical analysis. Therefore, the objective of **Experiment 2** is stated as:

Experiment 2

Detect spatial and energy anomalies in arriving aircraft ADS-B data and associated derived metrics and perform subsequent statistical analysis.

An experimental approach to detect spatial and energy anomalies in arriving aircraft ADS-B data and associated derived metrics and perform subsequent statistical tests and visual comparison requires four steps: (i) data processing, (ii) spatial anomaly detection, (iii) energy anomaly detection, (iv) statistical analysis set-up. Each of the required steps is detailed. Additionally, acceptance of the set of hypotheses (**Hypothesis 2.1**, **Hypothesis 2.2**, and **Hypothesis 2.3**) is discussed.

4.2.1 Data Processing

Similar to the experimental approach described in section 4.1, a cleaned and augmented set of ADS-B data for arriving aircraft during a specified time period is required. In the context of the **Anomaly Detection** step of the proposed offline data-driven methodology to be applied to ADS-B data for arriving aircraft within the terminal airspace, two sets of derived additional metrics are relevant: (i) the (x, y) coordinates introduced in section 4.1 for the spatial anomaly detection step and (ii) the relevant energy metrics. The two sets of derived additional metrics are discussed as well as the data processing tasks of re-sampling the data set and identifying go-arounds.

Additional Metrics

Recall, section 3.2 introduces that spatial anomaly detection is typically performed leveraging air traffic flow identification techniques. Therefore, the spatial anomaly detection step

requires that each longitude and latitude coordinate pair has been projected onto Cartesian coordinate system for trajectory clustering. On the other hand, energy anomaly detection requires appropriate energy metrics. This research proposes to place the focus on three primary energy metrics in which to detect anomalies: SPE, SKE, and STER. These energy metrics are contained in the set of energy metrics relevant to aviation anomaly detection presented by Puranik [25] and Puranik et al. [98]. Further, these metrics are contained in the set of metrics presented by Deshmukh [50] as candidates for performing energy anomaly detection.

The SPE may be taken as simply the height above ground level of the aircraft. The specific kinetic energy may be computed as:

$$SKE = \frac{V^2}{2g},$$

where V is the velocity of the aircraft and g is the gravity constant. Computation of the STER requires the computation of the STE, which may be simply the sum of the SPE and SKE:

$$STE = SPE + SKE$$

The STER may then be computed as:

$$STER = \frac{STE_{i+1} - STE_i}{\Delta t}$$

Δt is the time difference between two consecutive records, $i + 1$ and i.

Re-Sampling

As noted in section 2.2, the "scope" of the ADS-B data used for spatial and energy anomaly detection typically is different. Spatial anomaly detection is generally performed considering all trajectories within some radius of the airport. On the other hand, energy anomaly

detection is generally performed considering trajectories whose records have been cut off at some distance- or time-based threshold, i.e. 20 nautical miles cumulative ground track distance from the touchdown point or the final 15 minutes of approach. In this research, a distance-based cutoff is selected for the energy anomaly detection data set to mitigate any effect of the potentially differing approach velocities for the differing aircraft types.

Moreover, similar to the requirement for spatial anomaly detection, a re-sampling of the data is a necessary processing step before performing energy anomaly detection. The selection of a distance-based cutoff is justified for the spatial anomaly detect step, and, similarly, that justification may be applied to the selection of a distance-based cutoff for the energy anomaly detection step. Therefore, two re-sampled data set are produced to be used to anomaly detection: (i) a data set containing all trajectory records within some radius of the airport location and (ii) a data set containing all trajectory records cut off at a specified ground track distance from the touchdown point. Both data sets are of the same size and contain the same aircraft trajectories, yet the data set containing records cut off at a specified ground track distance from touchdown are thought of as having a higher "resolution".

Go-Around Identification

The final data processing step involves go-around identification. Go-arounds are a well-practiced, yet relatively rare procedure, often undertaken due to approach stability or ATC considerations [108]. Go-arounds are generally easily detected in historical data using logic-based assessments of metrics such as vertical rate, altitude, and ground track distance from touchdown [108, 125]. Several studies have focused on identifying, predicting, and/or classifying go-arounds [108, 125, 126, 127].

Go-arounds are almost always either excluded from energy anomaly detection analysis and/or detected as being anomalous by an air traffic flow identification/spatial anomaly detection method. This is because go-arounds are, inherently, non-standard operations and

the association to a specific air traffic flow is often not possible. In this research, go-arounds may be cursorily identified as trajectories in which, at a ground track distance from touchdown greater than 25 nautical miles, an aircraft's height above ground level reaches below 2,500 ft and its vertical rate reaches above 5 ft/min. This criteria does not accurately detect all possible go-arounds, yet is sufficient. Kumar et al. [125] provide more robust go-around detection criteria in ADS-B data. Therefore, the go-arounds are identified within the provided cleaned and augmented ADS-B data set such that only trajectories remain for which no spatial anomaly label has been specified a priori. The removal of go-arounds enables a spatial anomaly detection method to discover previously unknown spatial anomalies. Kumar et al. [125] perform a detailed assessment on the characterization of go-arounds, specifically. After the removal of go-arounds from the data sets used for spatial and energy anomaly detection, the data processing step is complete, where two data sets are returned: (i) *radius data* (used for spatial anomaly detection) and (ii) *distance cutoff data* (used for energy anomaly detection).

4.2.2 Spatial Anomaly Detection

As introduced in section 2.2, spatial anomaly detection generally occurs as a byproduct of applying trajectory clustering algorithms to perform air traffic flow identification. Therefore, the proposed more reliable method of identifying air traffic flows is leveraged to perform spatial anomaly detection. Specifically, spatial anomaly detection is performed by applying the HDBSCAN algorithm with the WED (as presented in Corrado et al. [128]). Identifying air traffic flows is an important step in segmenting the overall data set to produce the data sets required for energy anomaly detection. As mentioned in section 2.2, performing energy anomaly detection on a data set containing all aircraft trajectories belonging to all of the different air traffic flows may result in energy anomalies remaining undetected.

As **Research Question 2** is concerned with a quantitative investigation into the rela-

tionship between spatial and energy anomalies, it is necessary to assign detected spatial anomalies to an air traffic flow such that they may remain in the analysis. It is recalled that the lack of clarity on further consideration of spatial anomalies in the application of energy anomaly detection methods lead to the identification of **Gap 2**. Therefore, if a trajectory is detected as anomalous, the "closest" air traffic flow is also computed and recorded. The closest air traffic flow is considered to be the air traffic flow whose centroid is the shortest distance (with respect to the WED) from the spatially anomlaous trajectory. Thus, spatial anomaly detection assigned two labels to each trajectory: (i) a cluster label (-1 for anomalous trajectories and any cluster number greater than or equal to 0 for nominal trajectories) and (ii) an air traffic flow label (any number greater than or equal to 0, corresponding to the nominal cluster labels).

4.2.3 Energy Anomaly Detection

As specified, for each of the air traffic flows identified, a selected energy anomaly detection method is applied separately. Thus, the entire distance cutoff data set may be separated into several smaller data sets sets based on the air traffic flow to which the trajectory most closely belongs. Hence, each data set containing some fraction of the detected spatial anomalies. An energy anomaly detection algorithm is subsequently applied to each smaller distance cutoff data set to detect some fixed portion of energy anomalies.

With respect to clustering the energy metrics, there generally exists one cluster of nominal operations that is identified for each air traffic flow data set [98]. Consequently, a clustering algorithm is required that is capable of returning only one cluster and detecting outliers, or anomalies. The density-based algorithms such as DBSCAN or HDBSCAN are appropriate. Specifically, as only one cluster is expected, there is assumed to be no benefit of applying HDBSCAN because multiple clusters of varying densities are not anticipated. Further, the DBSCAN algorithm has been widely-utilized for the energy anomaly detection task in the aviation literature. Thus, to perform energy anomaly detection, DBSCAN

is selected. While the WED is leveraged for spatial anomaly detection, the ED is more appropriate for the energy anomaly detection application. Selection of the ED as the distance function stems from the relatively equal importance of an aircraft's energy profile the entirity of the way through the touchdown. For instance, high or low energy landings are operational events of interest to be detected an anomalous. As indicated, the metrics selected for energy anomaly detection include SPE, SKE, and STER, as combinations of these metrics are used in several energy anomaly detection studies [25, 50, 64, 93, 98, 103, 129].

In the context of performing a standardized statistical analysis of the results, it is required that a fixed fraction of outliers (anomalies) is detected for each air traffic flow data set. The parameters of DBSCAN enable straightforward iteration to a pre-defined value of the fraction of anomalies detected. For instance, there typically exists a monotonic relationship between ε and the detected fraction of anomalies for a fixed value of the $MinPts$ parameter. Specifically, as ε is increased, the detected fraction of anomalies is decreased. A standard practice within the aviation literature related to the application of DBSCAN does set a constant value of $MinPts$ and varies ε to reach a specified fraction of anomalies [47, 101, 102]. After application of DBSCAN with the ED separately to each air traffic flow distance cutoff data set, energy anomaly detection assigns a single label to each trajectory, -1 (energy anomaly) or 0 (energy-nominal).

4.2.4 Statistical Analysis Set-Up

As introduced in section 3.2, the successive detection of spatial and energy anomalies results in trajectories being placed in one of four categories, including: (i) nominal (N), (ii) only-spatial-anomaly (S), (iii) only-energy-anomaly (E), and (iv) both-spatial-and-energy-anomaly (B). **Hypothesis 2.1** states that if a trajectory is detected as a spatial anomaly, then it is more likely to also be detected as an energy anomaly. In other words, it is of interest to compute the ratio between the fraction of spatial anomaly trajectories detected as energy

anomalies and the fraction of spatially nominal trajectories detected as energy anomalies. This ratio may be referred to as the *likelihood ratio*, which is computed in terms of the trajectory categories as:

$$\frac{\frac{B}{B+S}}{\frac{E}{E+N}}$$

where, in this case, N, S, E, and B indicate the number of trajectories assigned to each category. Therefore, the statistical analysis set-up step requires the computation of the likelihood ratio.

In addition, the anomaly score is computed for each trajectory to provide more context to the "relative anomalousness" of the arriving aircraft trajectories within the terminal airspace with respect to their energy metrics. **Hypothesis 2.2** states that if only energy-nominal trajectories are considered, then only-spatial-anomaly trajectories have, on average, greater anomaly scores than nominal trajectories. In other words, it is of interest to compare the average anomaly score for only-spatial-anomaly trajectories and the average anomaly score for nominal trajectories. Therefore, the statistical analysis set-up step requires the computation of the mean of the anomaly scores of the only-spatial-anomaly trajectories and nominal trajectories.

Finally, **Hypothesis 2.3** states that if the statistical properties of the distribution of anomaly scores associated with only-energy-anomaly trajectories and both-spatial-and-energy-anomaly trajectories, then a distinct difference in statistical properties of the distribution is observed. In other words, statistical properties such as the mean, median, standard deviation, 25^{th} percentile and 75^{th} values are dissimilar for the only-energy-anomaly trajectories and both-spatial-and-energy-anomaly trajectories.

An overview of the experimental approach to support the **Anomaly Detection** step is presented in Figure 4.2.

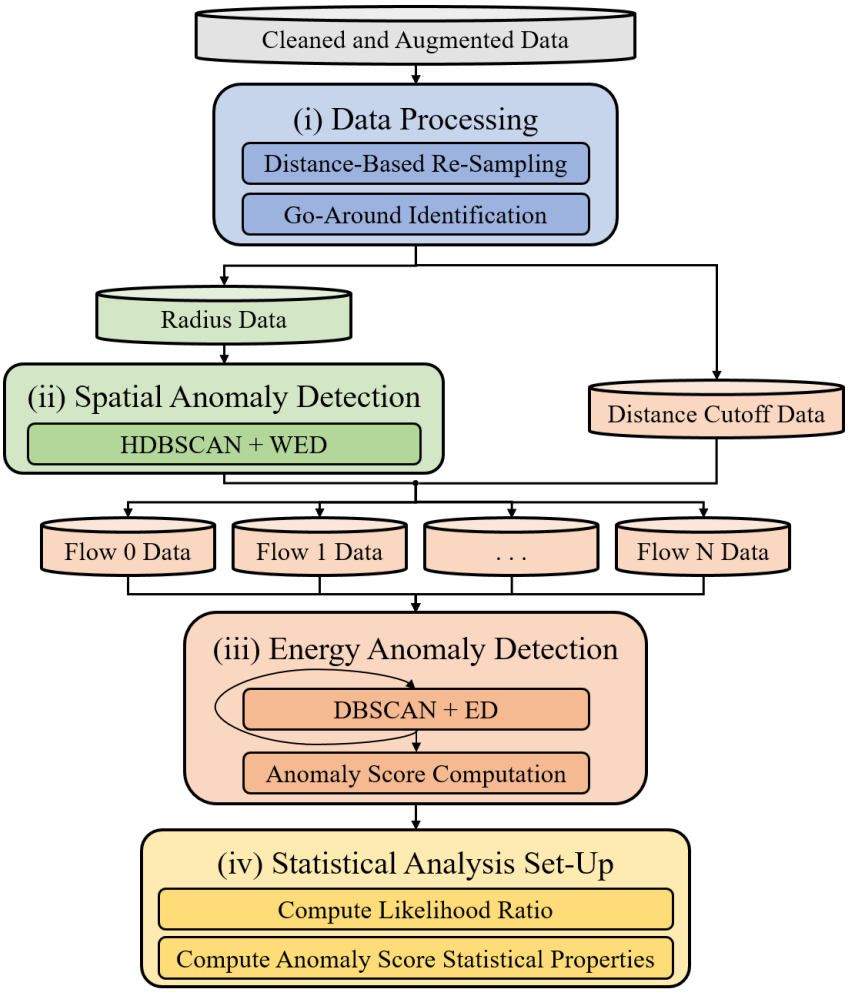

Figure 4.2: Experimental Approach: Gap 2

4.2.5 Hypothesis Validation

The requirements for validation of each of the hypotheses associated with **Research Question 2** (**Hypothesis 2.1**, **Hypothesis 2.2**, and **Hypothesis 2.3**) are discussed. Accepting each of the hypotheses associated with **Research Question 2** enables **Gap 2** to be successfully filled, where **Gap 2** corresponds to the **Anomaly Detection** step of the proposed methodology.

Hypothesis 2.1 Acceptance

Hypothesis 2.1 is accepted if the implementation of the experimental approach described corresponding to **Experiment 2** results in trajectories that are spatially anomalous being more likely to also be detected as energy-anomalous. A direct evaluation of the likelihood of spatial anomalies being more likely to be detected as energy anomalies may be made by evaluating the computed likelihood ratio. The numerator of the likelihood ratio is the fraction of spatially anomalous trajectories that also get detected as energy anomalies, while the denominator of the likelihood ratio is the fraction of spatially nominal trajectories that get detected as energy anomalies. These fractions can be thought of as probabilities of the trajectories from the spatially anomalous versus spatially nominal groups being detected as energy anomalies. Therefore, if the likelihood ratio is greater than 1, then it may be stated that spatial anomalies are more likely to be detected as energy anomalous. Thus, **Hypothesis 2.1 Acceptance** may be expressed as:

Hypothesis 2.1 Acceptance

If the likelihood ratio is greater than 1 for all air traffic flows, then **Hypothesis 2.1 is accepted**.

Hypothesis 2.2 Acceptance

Hypothesis 2.2 is accepted if the average anomaly score for only-spatial-anomaly trajectories is greater than the average anomaly score for nominal trajectories. While evaluating solely the mean anomaly scores for only-spatial-anomaly trajectories and nominal trajectories provides a reasonable starting point to accept **Hypothesis 2.2**, it is important to consider whether the comparison is statistically significant. Therefore, a one-sided Welch's t-test [130] may be conducted to determine whether the mean anomaly score for only-spatial anomaly trajectories is statistically significantly greater than the mean anomaly score for nominal trajectories. To consider the results statistically significant, a 0.05 p-

value threshold is used such that p-values below 0.05 indicate that the null hypothesis (the mean only-spatial-anomaly trajectory anomaly score is greater than the mean nominal trajectory anomaly score) is rejected. Thus, **Hypothesis 2.2 Acceptance** may be expressed as:

Hypothesis 2.2 Acceptance

If a one-sided Welch's t-test is conducted to determine whether the mean anomaly score for only-spatial-anomaly trajectories is statistically significantly greater than the mean anomaly score for nominal trajectories and the p-value returned for all air traffic flows is less than 0.05, then **Hypothesis 2.2 is accepted**.

Hypothesis 2.3 Acceptance

Hypothesis 2.3 is accepted if the statistical properties of the distributions of only-energy-anomaly trajectories' anomaly scores and both-spatial-and-energy-anomaly trajectories' anomaly scores are distinctly different. The statistical properties considered include mean, median, standard deviation, 25^{th} percentile and 75^{th} values. Each of these values must be different when comparing the two distributions. In this analysis, all anomaly scores from all air traffic flows may be considered jointly. Thus, **Hypothesis 2.3 Acceptance** may be expressed as:

Hypothesis 2.3 Acceptance

If the mean, median, standard deviation, 25^{th} percentile and 75^{th} values computed for the distributions of anomaly scores for only-energy-anomaly trajectories and both-spatial-and-energy-anomaly trajectories are different, then **Hypothesis 2.3 is accepted**.

4.3 Airspace-Level Analysis

Gap 3, which is associated with the **Airspace-Level Analysis** step in the proposed methodology, is mapped to two inter-related research questions, **Research Question 3.1** and **Research Question 3.2**, where **Hypothesis 3.1** and **Hypothesis 3.2** are formulated in response, respectively. The two research questions and associated hypotheses are reviewed such that an experiment is designed for each research question/hypothesis pair, where these experiments are connected as steps in an overall experimental approach that supports filling **Gap 3**.

Specifically, **Research Question 3.1** and the associated **Hypothesis 3.1** are related to an identification and characterization step of the overall experimental approach. **Research Question 3.2** and the associated **Hypothesis 3.2** are related to the prediction step of the overall experimental approach. It is recalled that a data processing step was discussed in section 3.3, which must occur before any experiments are performed. Therefore, the overall experimental approach that supports filling **Gap 3** requires three main steps: (i) data processing, (ii) identification and characterization, and (iii) prediction. Each of the three steps is detailed. Additionally, the acceptances of **Hypothesis 3.1** and **Hypothesis 3.2** are discussed.

4.3.1 Data Processing

To generate the airspace density matrix requires a set of cleaned and augmented data. Specifically of interest in the generation of the airspace density matrix are the (x, y) coordinates of all arriving aircraft operating that touch down during a defined time interval. Thus, the data set for an entire analysis time period is broken up into smaller data sets containing only the trajectories of arriving aircraft that touch down during specified shorter time intervals of equal length. For instance, time intervals may be selected as one hour, two hours, etc. within a broader time period of one month, six months, one year, etc.

Recall section 3.3 describes the general approach to generating an airspace trajectory matrix. For each time interval, an airspace matrix is initially generated as an $N \times N$ matrix of zeros, where N is the number of spaces within the 2D grid representing the terminal airspace's x and y coordinates. For each trajectory within the set of trajectories of arriving aircraft that touch down during the specified time interval, a second matrix, the trajectory matrix, is generated. The trajectory matrix is initially generated as an $N \times N$ matrix of zeros. Subsequently, for each set of (x, y) coordinates making up the trajectory, it is determined which grid space this point falls within. If there does not already exist a 1 within this grid space, a 1 is added to replace the 0 value. The restriction on a grid space being filled more than once per trajectory is instituted to consider situations in which the grid may be of very low resolution, i.e. N is small, and/or trajectory measurements are available at a very high resolution, as it is undesirable to skew the distribution of trajectory points towards these grid spaces if trajectories were to be counted more than once per grid space. The restriction on a grid space being filled more than once per trajectory may be a limitation in the case of holding patterns and/or go-arounds. However, holding patterns and go-arounds are not the norm, making up a relatively small percentage of the data set, so this limitation is deemed acceptable. After generation of each trajectory matrix, the trajectory matrix is added element-wise to the airspace matrix. Finally, the airspace density matrix is generated by dividing the airspace matrix by the number of trajectories represented within the airspace matrix. The final data set of airspace density matrices are the required representations of operational states required to proceed with the subsequent steps of the prescribed experimental approach.

4.3.2 Identification and Characterization

Considering **Research Question 3.1** and **Hypothesis 3.1** related to the identification and characterization step of the experimental approach, the objective of **Experiment 3.1** is stated as:

Experiment 3.1

Implement a recursive DBSCAN procedure such that the clusters identified correspond to the operational patterns and the outlying operational states are characterized as being either anomalous or transitional, where the recursion proceeds until no new transitional operational states are characterized.

However, the airspace density matrices are 2D (the set of airspace density matrices is 3D), whereas the DBSCAN clustering algorithm requires a 2D feature vector matrix as an input. Thus, the airspace density matrices should be flattened, i.e., stack the rows and/or columns of the matrix into a 1D vector. This results in a vector of length $N \times N$, where each element in the vector may be considered to be a feature to be considered in the DBSCAN clustering. Therefore, depending on the size of N, there may be a very large number of features to consider. There are a few potential issues that may arise with a large number of features. For instance, there may be an issue with the increase in computation time of the DBSCAN algorithm. Therefore, it is advantageous to *reduce the dimensionality* of the feature vectors.

Dimensionality reduction involves mathematically transforming data from a high-dimensional space into a lower-dimensional space such that the transformation preserves the defining characteristics of the high-dimensional data set. There exist a few commonly-utilized dimensionality reduction techniques within the literature, such as PCA [77], t-SNE [88], and Uniform Manifold Approximation and Projection (UMAP) [131]. PCA involves computing the principal components of a matrix and using them to perform a change of basis on the data set. Typically, only the first few principal components are considered. PCA is considered to be a linear dimensionality reduction technique, where only linear relationships are preserved in the transformation. The inability to preserve non-linear relationships in the data set is an inherent limitation of PCA. t-SNE, on the other hand, may be considered a non-linear dimensionality reduction technique. t-SNE operates by first computing the

pairwise distances between all points in the high-dimensional data set, subsequently constructing a probability distribution over pairs of high-dimensional points such that the more similar points are, the higher probability they will be assigned, and, finally, the algorithm defines a similar probability distribution over the lower-dimensional points with the objective of minimizing the Kullback-Leibler divergence (KL divergence) between the two probability distributions. A limitation associated with t-SNE is high computation time, which may make the tuning of appropriate parameters difficult. In addition, the distance between clusters of points in the reduced-dimension space does not have any physical meaning. Finally, UMAP may similarly be considered a non-linear dimensionality reduction technique, where UMAP was developed to overcome limitations that persist in using t-SNE. Similar to t-SNE, UMAP first computes the pairwise distances between points and subsequently attempts to preserve the distribution of those distances in a lower-dimensional space. However, UMAP is observed to require significantly less computation time and is said to preserve more of the global structure of a data set. Due to the ability to preserve non-linear relationships and the reduced computation time compared with t-SNE, UMAP is selected to perform the dimensionality reduction.

Once the dimensionality of the data set has been reduced, it is then possible to computationally efficiently implement the recursive DBSCAN procedure. The operational patterns are identified as the clusters identified by DBSCAN, and outlying operational states are characterized as either being anomalous operational state time intervals or transitional operational state time intervals. Transitional time intervals are characterized as such if the operational patterns of the time intervals before and after the time interval detected as outlying differ. Once a time interval's operational state has been characterized as being transitional, that time interval is removed from the data set, and the DBSCAN clustering and operational state characterization loop repeats until no new outlying time intervals are characterized as transitional.

Though, it is important to determine whether the operational patterns identified and

operational states characterized are consistent. First, considering the identification of operational patterns, it is important to determine whether the operational patterns identified are truly distinct and whether the operational patterns identified align with any official operational plans published for the terminal airspace of interest. To "summarize" the operational patterns identified, it is proposed to identify the prominent air traffic flows, and, subsequently, the centroids of the prominent air traffic flows, observed for each operational pattern. In this context, summarizing an operational pattern involves representing all the trajectories that were aggregated to represent all airspace-level operational states associated with an operational pattern. Identifying air traffic flows and their respective centroids to summarize the operational patterns *after* their identification, where each identified operational pattern is considered separately, has many benefits as opposed to identifying the air traffic flows *prior* to identifying operational patterns.

To adequately summarize operational patterns as identified air traffic flow centroids such that they are comparable, it is important to note that the fraction of outliers should be set to fall within a specified range, as described related to the **Air Traffic Flow Identification** step. DBSCAN provides greater "control" over the fraction of outliers detected as opposed to HDBSCAN, as the relationship between ε and the fraction of outliers detected is generally monotonic for a set value of $MinPts$, whereas the relationship between the optional HDBSCAN smoothing parameter and the fraction of outliers detected for a set value of the minimum cluster size parameter is often not monotonic. Varying the values of the HDBSCAN smoothing parameter does not guarantee that a specified fraction of outliers will ultimately be achieved. Considering there may be several operational patterns that are all required to have air traffic flow identified with a similar fraction of outliers detected, leveraging an algorithm for air traffic flow identification that provides greater control over the detected fraction of outliers is advantageous. Further, performing air traffic flow identification on a data set containing only arriving aircraft trajectories operating under a specified operational pattern may mitigate the impact of air traffic flows being of

varying densities. Therefore, DBSCAN is selected to perform the air traffic flow identification. Additionally, considering **Hypothesis 1**, the distance function selected to be used in the implementation of DBSCAN is the WED.

The resulting summary of the identified operational patterns as air traffic flow centroids enables a comparison of the distinctness of operational patterns with respect to each other and a comparison of the alignment of operational patterns with the distinct official arriving aircraft operational plans for a terminal airspace of interest. Specifically, to assess the distinctness of operational patterns with respect to each other, the alignment of the air traffic flow centroids identified for each operational pattern should be considered. The air traffic flow centroids may also be leveraged to assess the alignment of the operational patterns with the distinct official arriving aircraft operational plans for the terminal airspace of interest.

On the other hand, it is similarly important to consider the effectiveness of the characterization of operational states. Specifically, it is important to determine whether the operational states characterized as anomalous appear truly anomalous in the context of other operational states and, similarly, if the operational states characterized as transitional appear truly transition in nature. Related to appearing truly transitional, evidence of the presence of two or more distinct air traffic flows that are associated with one operational pattern and not the other and vice versa is required. However, determining truly anomalous operational states is more subjective, requiring assessment of the collective behavior of the individual trajectories. However, it can be assumed that the arriving aircraft trajectories operating within the terminal airspace during time intervals whose operational states have been characterized as anomalous should have little correspondence to or conformance with air traffic flows that have been identified for any of the operational patterns or contain many trajectories that appear to be spatially anomalous. Overall, completion of the identification of operational patterns and characterization of operational states enables the prediction step in the experimental approach to proceed.

4.3.3 Prediction

Considering **Research Question 3.2** and **Hypothesis 3.2** related to the prediction step of the experimental approach, the objective of **Experiment 3.2** is stated as:

> **Experiment 3.2**
>
> Train a set of multi-class classification algorithms to predict operational patterns with input features derived from recorded weather measurements and compute a classification performance metric for each multi-class classification algorithm trained.

To perform the prediction step of the experimental approach, the set of multi-class classification algorithms must be defined. While there are several options available within the literature (k-nearest neighbors, naive Bayes, decision trees, artificial neural networks, gradient boosted decision trees (XGBoost), SVMs, etc.), it is not necessary to test each one. Similar to [83], three classification algorithms are selected for the task of predicting operational patterns with input features derived from recorded weather metrics. In this research, the three classification algorithms selected are artificial neural networks, XGBoost, and SVMs. A brief description of each is presented:

- **Artificial Neural Networks**: Comprised of an input layer, various numbers of hidden layers with various numbers of nodes, and an output layer. Each node is typically connected to one or more other nodes and has an associated weight. In the training phase, the weights are initialized to random numbers and, subsequently, are refined according to the minimizing of some specified objective function.

- **XGBoost**: The XGBoost algorithm was developed as a scalable tree boosting system [132] and implements the gradient boosting decision tree algorithm. Decision trees operate by continuously splitting data according to a selected input feature, whereas boosting is an ensemble machine learning technique where new models are added to an *ensemble* to correct any errors made by models already in the ensemble. Gra-

dient boosting, specifically, involves creating new models that predict the residuals of the prior models and adding them together to generate the final prediction, where the gradient descent algorithm is leveraged to minimize the objective function when adding new models.

- **Support Vector Machines**: SVMs aim to identify a hyper-plane (decision boundary) in an n-dimensional space, where n is the number of input features. This hyper-plane should separate the prediction classes with the greatest space, or margin, between the points within the different prediction classes. Therefore, the decision boundary theoretically separates the classes "best".

Considering that each of the three classification algorithms will be leveraged to train different prediction models, it is important to determine which quantitative classification metric is appropriate to evaluate the performance of the algorithms. Murça & Hansman [83] compute an accuracy metric to compare three classification algorithms, where accuracy is defined as the percentage of correctly classified samples. In [83], the "best" prediction model achieved an accuracy of approximately 83% when forecasting operational patterns one hour ahead. Therefore, in this research, the threshold above which the accuracy of the "best" prediction is set at 85% to provide the "reasonable accuracy" specified in **Hypothesis 3.2**. Other classification performance metrics may be computed as well, such as precision, recall, and F1-score.

An overview of the experimental approach (containing **Experiment 3.1** and **Experiment 3.2**) to support filling **Gap 3** is presented in Figure 4.3.

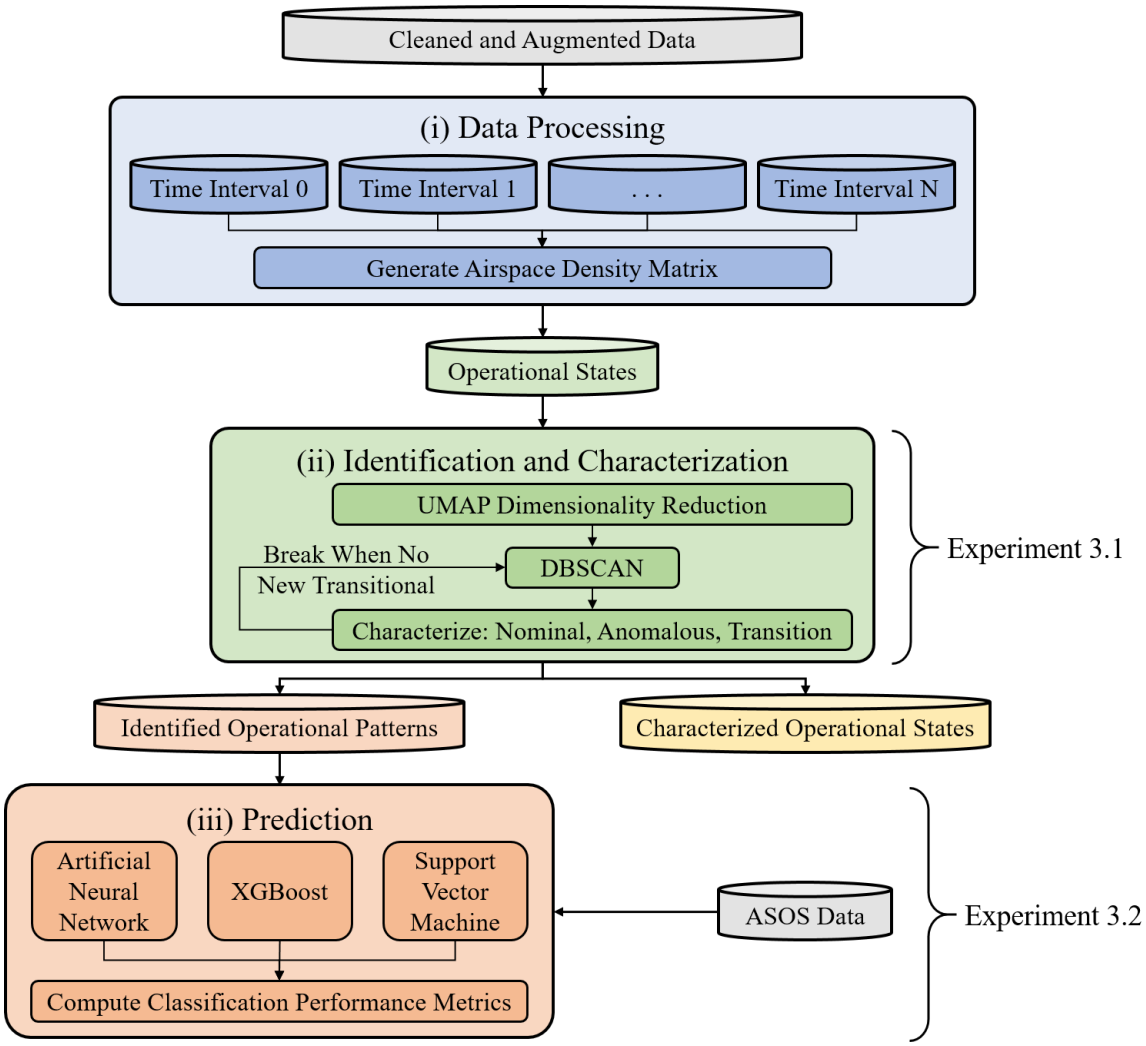

Figure 4.3: Experimental Approach: Gap 3

4.3.4 Hypothesis Acceptance

The requirements for acceptance of **Hypothesis 3.1** associated with **Research Question 3.1** and **Hypothesis 3.2** associated with **Research Question 3.2** are discussed, respectively. Acceptance of both of the hypotheses requires distinct experiments to be performed, where each of these experiments are connected to formulate an overall experimental approach aimed at successfully filling **Gap 3**, which corresponds to the **Airspace-Level Analysis** step of the proposed methodology.

Hypothesis 3.1 Acceptance

Hypothesis 3.1 is accepted if the implementation of the prescribed recursive DBSCAN procedure on the operational state representations in a reduced-dimensional space successfully identifies operational patterns and characterizes operational states as nominal, anomalous, or transitional in nature. As mentioned, the operational patterns may be summarized by identifying the prominent air traffic flows corresponding to the operational pattern and taking the air traffic flow centroids. These air traffic flow centroids for each operational pattern are able to be compared among the different operational patterns identified as well as to official operational plans published for the terminal airspace of interest. Further, an assessment of the characterization of operational states as either transitional or anomalous relies on the plotting of all trajectories occurring within the operational states characterized as such, respectively. For instance, a plot of the arriving aircraft trajectories operating during time interval whose operational states have been characterized as transitional should reveal evidence of a mix of air traffic flows distinct to each of the operational patterns on either side of the transition period. In contrast, a plot of the arriving aircraft trajectories operating during time interval whose operational states have been characterized as anomalous should reveal little evidence of trajectories corresponding to or conforming with air traffic flows that have been identified for any of the operational patterns. Thus, **Hypothesis 3.1 Acceptance** may be expressed as:

Hypothesis 3.1 Acceptance

If the dimension of the representations of the operational states as airspace density matrices is reduced and a recursive DBSCAN procedure is implemented such that operational patterns are identified that are distinct and align with distinct arriving aircraft official operational plans for the terminal airspace of interest and operational states are identified that are both transitional and anomalous in nature are characterized, then **Hypothesis 3.1 is accepted**.

Hypothesis 3.2 Acceptance

Hypothesis 3.2 is accepted if the highest-accuracy prediction model is selected from a set of three multi-class classification algorithms trained on input features derived from recorded weather measurements and the accuracy of the selected prediction model is "reasonable", i.e. above 85% considering the accuracy attained in other similar works, such as the work presented by Murça & Hansman [83]. In addition, the set of multi-class classification algorithms selected includes an artificial neural network, a gradient-boosted decision tree (XGBoost), and a SVM. Thus, **Hypothesis 3.2 Acceptance** may be expressed as:

Hypothesis 3.2 Acceptance

If an artificial neural network, gradient-boosted decision tree (XGBoost), and support vector machine are trained to predict operational patterns provided input features derived from recorded weather measurements and the algorithm providing the highest accuracy reaches or exceeds an 85% accuracy threshold, then **Hypothesis 3.2 is accepted**.

4.4 Summary

This chapter detailed the design of three experimental approaches, each corresponding to one of the gaps related to the steps in the proposed methodology. An overview of the mapping of gaps, to research questions, to hypotheses is displayed in Figure 4.4.

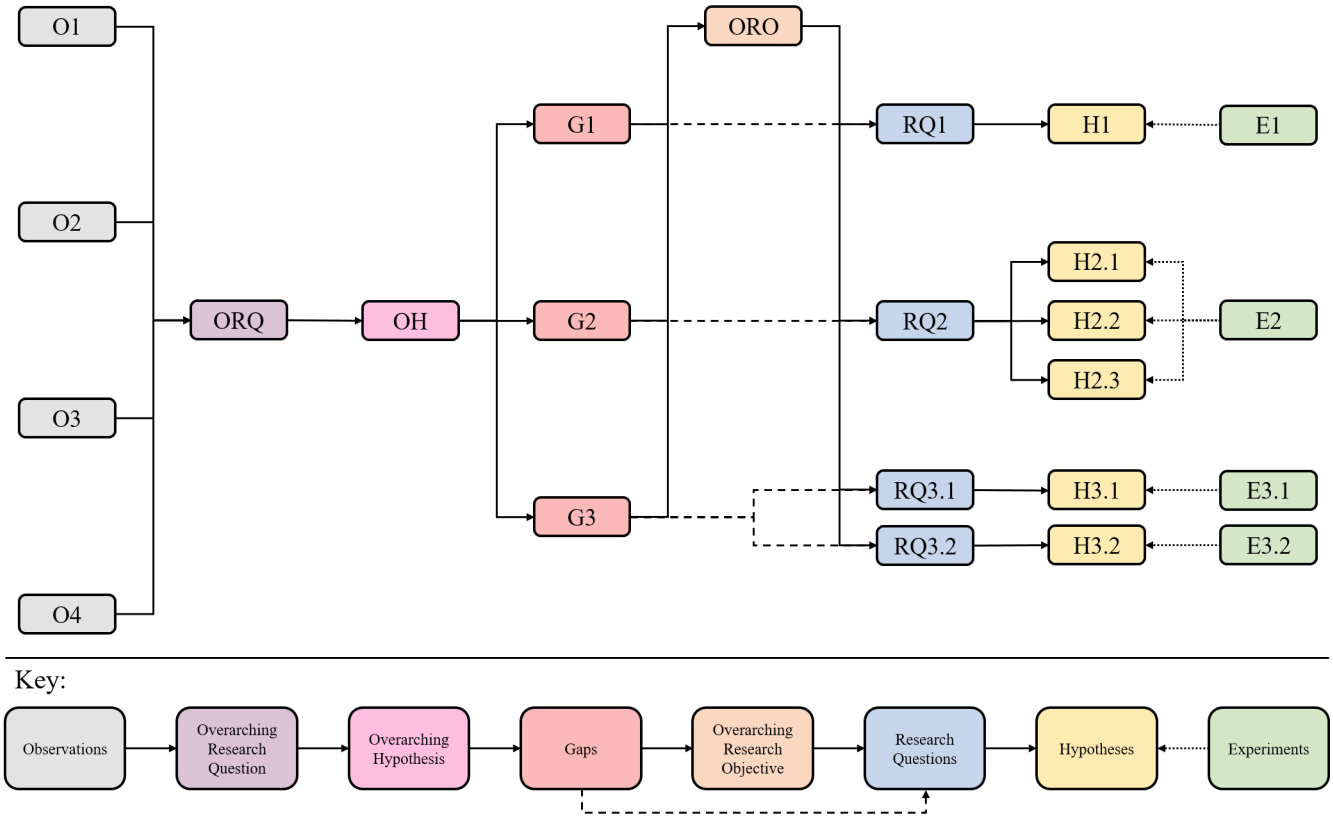

Figure 4.4: Thesis Structure: Experimental Approach

Implementation of the prescribed steps for each experimental approach requires the availability of an extracted and cleaned ADS-B data set, augmented with appropriate derived metrics, for all arriving aircraft operating within a terminal airspace of interest during a specified time period of interest.

CHAPTER 5
IMPLEMENTATION AND RESULTS

This chapter begins with a discussion on the data extraction, cleaning and augmentation process for a selected terminal airspace and time period of interest. Subsequently, the implementations of the experimental approaches corresponding to filling the gaps associated with each step in the proposed methodology are detailed, and the results that are obtained are presented and discussed.

5.1 Data Extraction, Cleaning, and Augmentation

The KSFO terminal airspace is selected as the terminal airspace of interest in this research, where the time period of interest is selected as being the full year of 2019. The time period of 2019 is selected as it is the most recent complete year that did not experience a major disruption in operations, such as due to the COVID-19 pandemic. KSFO is selected as it is a major U.S. airport, though not one of the top-five busiest. KSFO tends to experience a relatively stable Mediterranean climate throughout the year, compared to other airports in the U.S. Further, KSFO's intersecting parallel runway configuration is of interest, as KSFO has been the focus of other research efforts focusing on air traffic flow identification, such as [63]. The intersecting parallel runway configuration is displayed in Figure 5.1.

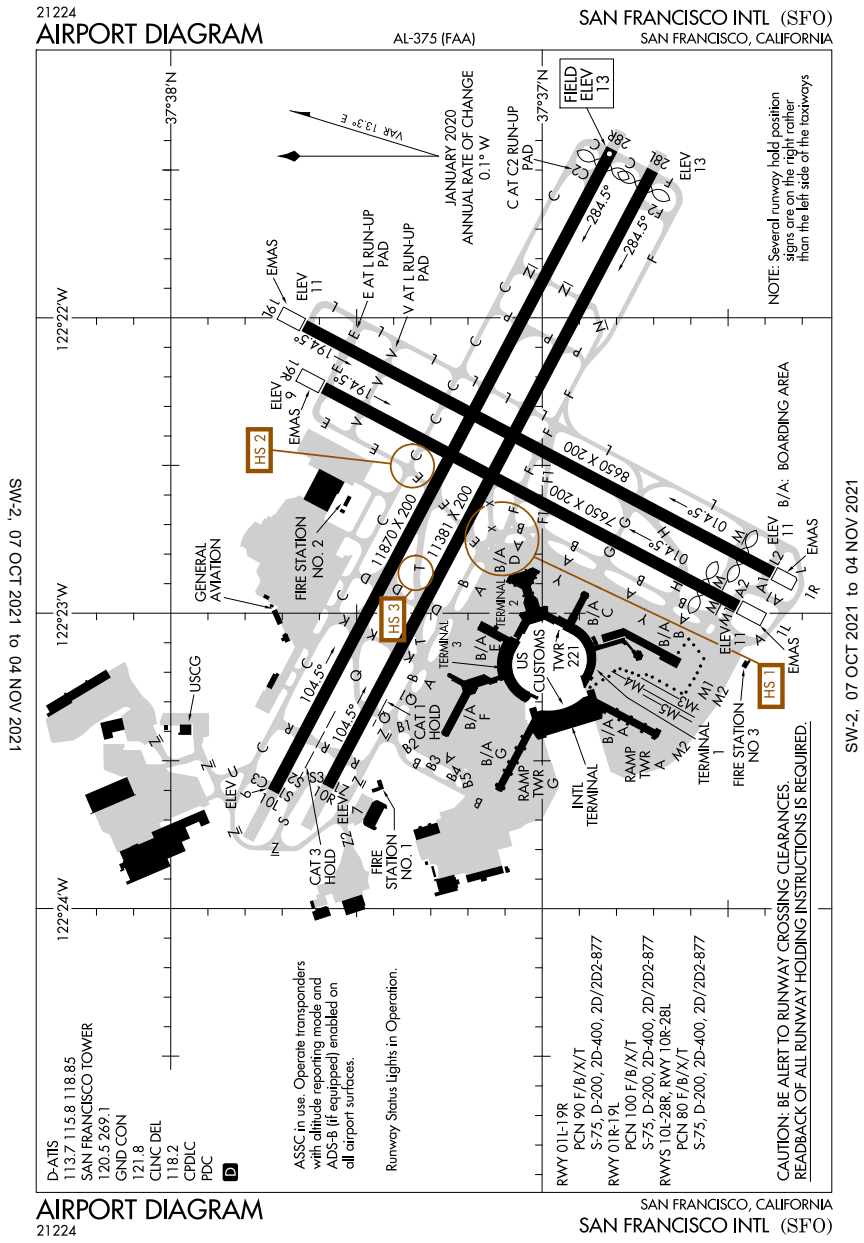

Figure 5.1: San Francisco International Airport Runway Configuration [133]

KSFO is a part of the SFO metroplex, which additionally contains Oakland International Airport (KOAK) and San Jose International Airport (KSJC). KSFO, KOAK, and KSJC collectively operate under three published official operational plans [134]: (i) west plan, (ii) south east plan, and (iii) nighttime operations. The west plan, displayed in Fig-

ure 5.2, is the dominant official operational plan due to the tendency of the winds to blow from the west 95-98% of the time. The south east plan, displayed in Figure 5.3, is observed less than 5% of the time. Finally, nighttime operations, displayed in Figure 5.4, is in place at night such that aircraft avoid populated residential areas. However, there is no difference in the flight patterns included in the official operational plans for KSFO between the west plan and the nighttime operations, as evidenced by Figure 5.2 and Figure 5.4.

Figure 5.2: SFO Metroplex West Plan [134]

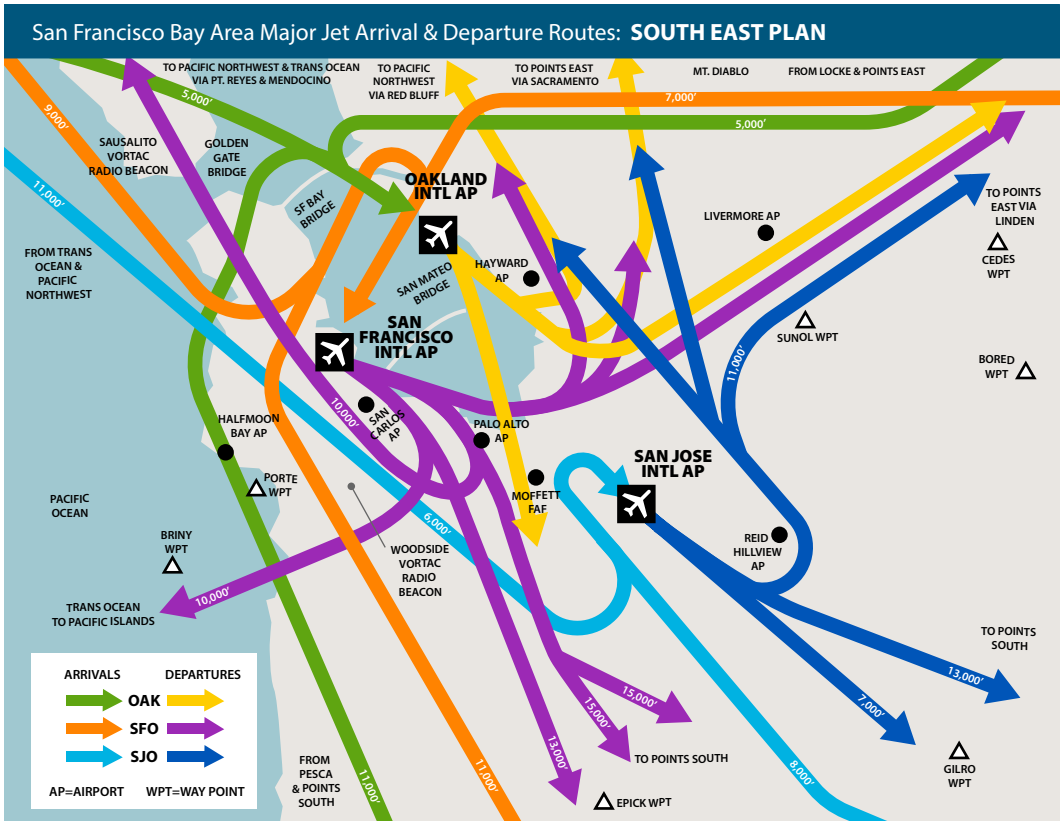

Figure 5.3: SFO Metroplex South East Plan [134]

Figure 5.4: SFO Metroplex Nighttime Operations [134]

To implement the designed experimental approaches, it was required that ADS-B data be extracted, cleaned and augmented for arriving aircraft operating within the KSFO terminal airspace in 2019. Specifically, arriving commercial aircraft operations were considered as the operations of general aviation aircraft are often distinct from commercial operations. The source of the ADS-B data considered in this work was the OpenSky Network's historical database [135]. From a global network of sensors, the non-profit OpenSky Network processes and archives ADS-B data [135]. Starting with the deployment of 11 sensor nodes in Central Europe [135], the OpenSky Network now contains over 3,000 crowd-sourced sensors located all over the world [136]. Figure 5.5 presents an overview of the OpenSky Network crowd-sourcing process, including a map of all active receivers on July 1^{st}, 2020 [136], where it is evident most receivers are concentrated in Europe and the U.S. The OpenSky Network provides full access to the historical database of aircraft *state vectors*,

broadcast by ADS-B, to researchers, where over 100 academic groups in the past five years have leveraged the OpenSky Network for a diverse range of research endeavors [136].

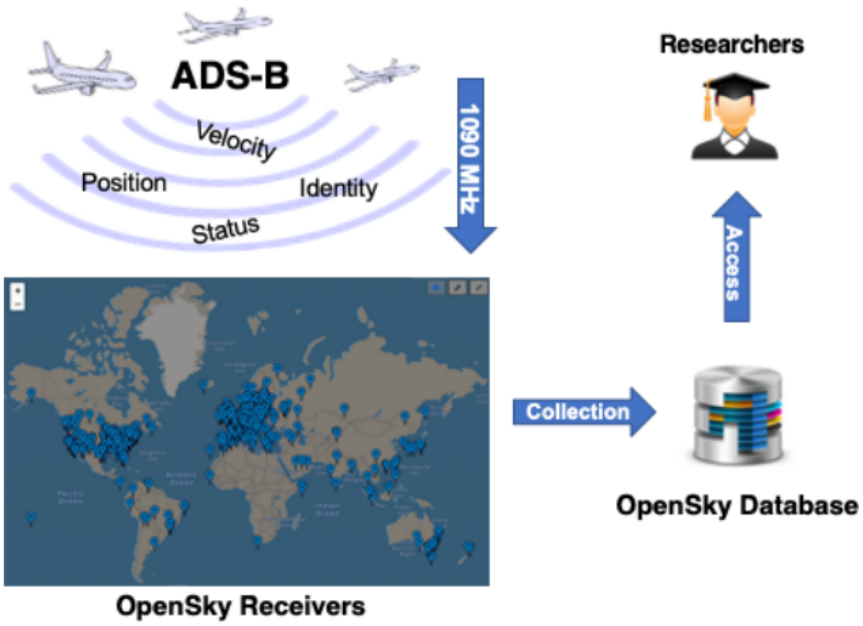

Figure 5.5: OpenSky Network Crowd-Sourcing Process (including map of active receivers on July 1^{st}, 2020) [136]

The aircraft state vectors generally include the following metrics [135]:

- $icao24$: the ICAO code (the transponders unique 24-bit address assigned by ICAO)
- $callsign$: a space-filled eight-character string typically assigned by the operating airline
- $time$: the epoch time stamp (added on the receiver side)
- lon: longitude
- lat: latitude
- alt: barometric altitude with respect to the standard atmosphere
- $track$: track angle as the clockwise angle from geographic north

- vel: east-west velocity (ground speed)
- $vrate$: north-south velocity (vertical rate)

In the OpenSky Network's historical database, the $track$ metric is referred to as "heading" [137]. The online documentation of this metric does indicate that the term "heading" is not entirely correct [137], yet it is common to (incorrectly) interchange these terms [138]. Heading may be said to be the direction in which the aircraft is pointed (simply, the direction in which the aircraft nose is pointed) [138], whereas track is the actual direction the aircraft is tracking across the ground (simply, the direction of the aircraft's path at any point) [138]. The heading of an aircraft may be different than its track is due to wind and/or magnetic variation/deviation.

A subset of state vectors may make up an individual flight or flight segment's trajectory, where each individual flight or flight segment represents a contiguous set of position reports from a given $icao24$ aircraft address. Table 5.1 displays the structure and format of the information stored in the state vector:

Table 5.1: Example Structure and Format of the State Vectors Obtained from the OpenSky Network Historical Database

icao24	**callsign**	**time**	**lon**	**lat**	**alt**	**track**	**vel**	**vrate**
a1a9f4	SKW5385	1546301514	-122.690	37.957	15250	167.28	413.15	-1664
a1a9f4	SKW5385	1546301515	-122.689	37.955	15200	167.28	413.15	-1728
a1a9f4	SKW5385	1546301516	-122.689	37.953	15175	167.24	412.17	-1920
a1a9f4	SKW5385	1546301517	-122.688	37.951	15150	167.21	411.20	-2112

While an incredibly rich source of data, due to the crowd-sourced nature of the state vectors provided by the OpenSky Network, it is possible data errors exist. Thus, an extensive data cleaning procedure is required. The extraction of the state vectors for aircraft arriving at KSFO in 2019 is detailed. Additionally, the subsequent cleaning and augmentation of the data is discussed, where the data is augmented with those derived metrics

discussed in chapter 4, such as Universal Transverse Mercator (UTM) coordinates and energy metrics. Finally, the properties and statistics related to the extracted, cleaned, and augmented ADS-B data are presented.

5.1.1 Data Extraction

The state vectors available within the OpenSky Network's historical database were extracted leveraging the *traffic* [139] Python library. The *traffic* Python library enables the direct querying of the OpenSky Network's historical database using explicit SQL queries by specifying the range of parameters such as time, longitude, latitude, altitude, etc. Accordingly, the ADS-B data for aircraft operating within a 20 nautical mile bounding box (longitude and latitude bounds) around the KSFO terminal airspace in 2019 was requested via the *traffic* Python library with a maximum altitude of 25,000 ft. Extracted state vectors were not in the form of "flight segments" or "trajectories", but are rather a large data set of mixed state vectors, generally ordered by time. Further, due to this, arrival/departure flights were not directly query-able; therefore, extracting all state vectors meeting the listed conditions resulted in state vectors corresponding to departing aircraft operating within the 20 nautical mile bounding box of the KSFO airport as well. Thus, the data cleaning procedure provided the ability to agglomerate state vectors as trajectories as well as filter out departures.

Additionally, the OpenSky Network state vectors provide measurements of barometric altitude with respect to the standard atmosphere, where this measurement may vary noticeably throughout the day based on the current atmospheric temperature and pressure. Thus, to compute a more accurate height above ground level of an aircraft at all trajectory points, the atmospheric pressure data for 2019 was also extracted. This data was extracted from the Iowa Environmental Mesonet [140], which collects historical Automated Surface Observing System (ASOS) observations at most airports across the U.S. The collection of ASOS observations arises as a result of a collaborative program between the National

Weather Service, the FAA, and the Department of Defense. Widely used by meteorologists, hydrologists, climatologists, and aviation weather experts, ASOS units are automated sensor suites designed to serve meteorological and aviation observing need and service as the U.S.'s primary surface weather observing network. Definitions of each of the weather metrics contained in the ASOS data set are presented [140]:

- $time$: epoch time
- $alti$: pressure altimeter in inches
- $drct$: wind direction in degrees from "true" north
- $dwpf$: dew point temperature in Fahrenheit, typically at two meters
- $gust$: wind gust in knots
- $mslp$: sea level pressure in millibar
- $relh$: relative humidity as a percentage
- $sknt$: wind speed in knots
- $tmpf$: air temperature in Fahrenheit, typically at two meters
- $vsby$: visibility in miles

A sample of the data extracted from the Iowa Environmental Mesonet is provided in Table 5.2.

Table 5.2: Example Structure and Format of the ASOS Data Extracted from the Iowa Environmental Mesonet

time	alti	drct	dwpf	gust	mslp	relh	sknt	tmpf	vsby
1546300800	30.02	350	30.9	30	1016.5	41.11	17	54.0	10
1546304400	30.02	340	30.9	30	1016.5	41.11	17	54.0	10
1546308000	30.03	350	32.0	30	1016.7	42.98	20	54.0	10
1546311600	30.05	350	30.0	30	1017.4	40.96	23	53.1	10

5.1.2 Data Cleaning and Augmentation

To clean the extracted ASOS data required a fairly straightforward procedure to both fill and interpolate missing values. On the other hand, the OpenSky Network data required a more comprehensive cleaning procedure. The overall data cleaning and augmentation steps are summarized in Figure 5.6. The data cleaning and augmentation procedure contains five primary steps: (i) initial cleaning, (ii) segment identification, (iii) intermediate cleaning, (iv) augmentation, and (v) final cleaning. Each step is discussed in further detail.

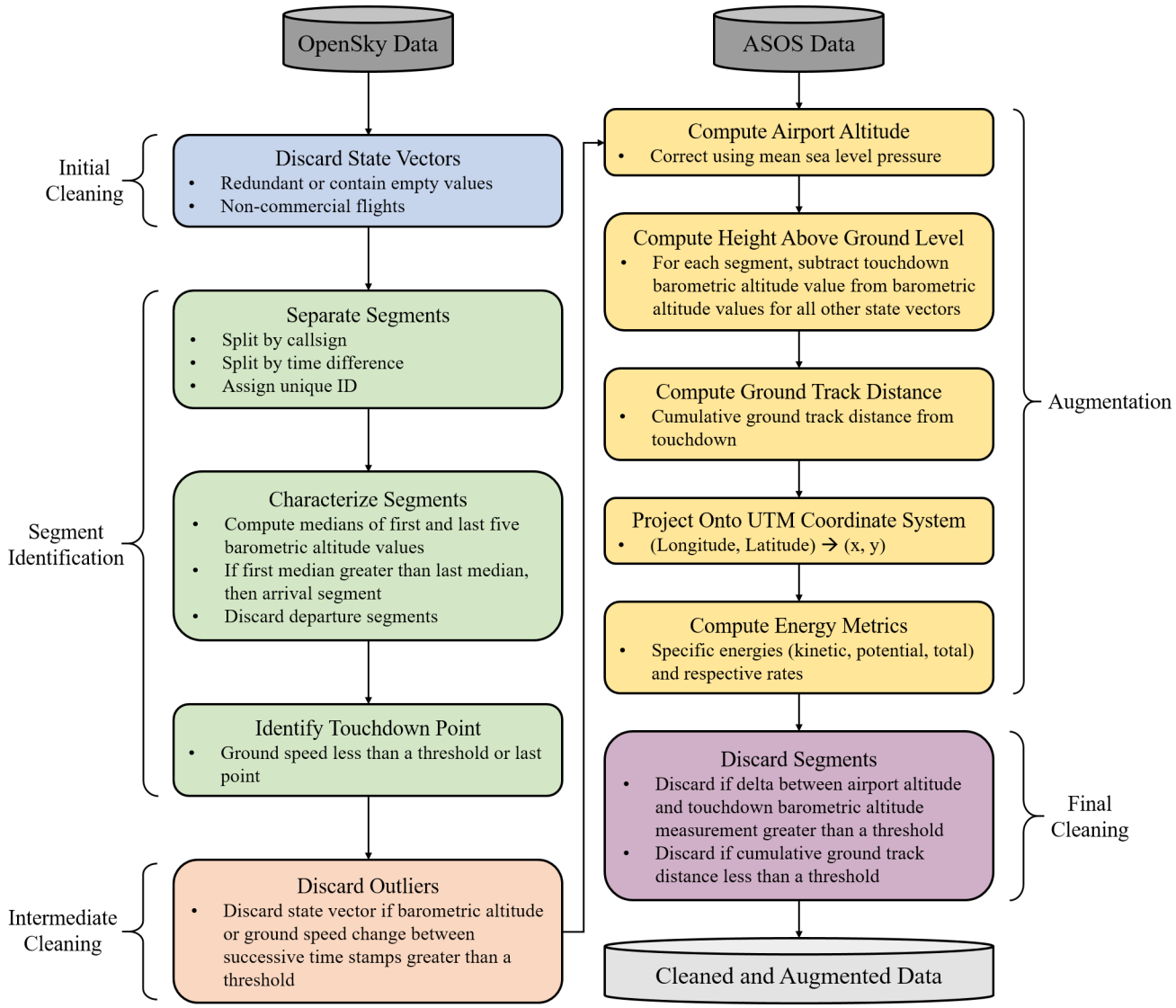

Figure 5.6: Data Cleaning and Augmentation Procedure

Initial Cleaning

The objective of the initial cleaning step is primarily to discard state vectors that are not complete, are redundant, or are not associated with commercial operations. State vectors that contained empty values or were redundant for lon, lat, alt, or vel were discarded. The $callsign$ metric enabled determination of whether the state vector was associated with a commerical flight such that state vectors not associated with a commercial flight were discarded.

Segment Identification

The objective of the segment identification step is to separate the subsets of state vectors that may be associated with each other as "flight segments" or "trajectories", characterize the flight segments are departure or arrival segments, and, finally, identify a touchdown point for each arriving aircraft trajectory. State vectors were split by value of the $callsign$ metric. Since it is possible that several flights operate under the same callsign throughout one day, if the time difference between two time-ordered state vectors was greater than five minutes, the state vectors were split into separate flight segments. To distinguish between the different identified flight segments (trajectories) within the data set, a unique identifier was assigned to each.

Subsequently, based on the medians of the first and last five altitude measurements, the flight segments were characterized as either being arrival or departure segments. The median of the altitude measurements were taken due to the possibility of data errors (large outliers that are undoubtedly erroneous) that may exist within the crowd-sourced data set. If the median of the first five altitude measurements was greater than the median of the last five altitude measurements, then the slight segment was characterized as an arrival. Otherwise, the trajectory segment was characterized as a departure and disregarded for the purpose of this work.

Finally, the state vector associated with the touchdown point was identified. The first state vector in which the ground speed reached below 100 knots was taken as the touchdown point state vector. However, if no point was recorded to reach below 100 knots, then the last point (time-wise) was taken as the touchdown point state vector. This is an approximate method to identify the touchdown point state vector, which may introduce some error. Though, this method was observed to be the most effective considering this data set.

Intermediate Cleaning

The objective of the intermediate cleaning is to discard state vectors containing outliers or erroneous state vector measurements that may exist due to the crowd-sourced and "unclean" nature of the OpenSky Network ADS-B data. Specifically, outliers in the more uncertain metrics, such as barometric altitude and ground speed, were considered. A coarse strategy to discarding state vectors due to outlying measurements in barometric altitude and/or ground speed was employed. The difference between successive state vector measurements was computed. If the difference between successive state vector measurements was greater than 1,000 ft with respect to barometric altitude and/or greater than 150 knots with respect to ground speed, then the state vector that caused the threshold to be exceeded was removed. These thresholds were selected due to the ADS-B data having a frequency of roughly 1 Hz, where it is unrealistic to assume these large changes occurring with that time frame or even a slightly larger time frame.

Augmentation

The objective of the augmentation step is to augment the ADS-B data set with relevant derived metrics. Additionally, in this step, the ASOS weather data metrics are considered. First, a corrected airport barometric altitude, or reference altitude (ref_alt), was computed for each trajectory leveraging the mean sea level pressure. Specifically, the recorded mean sea level pressure ($mslp$) from the ASOS data set associated with a trajectory's touchdown point's time stamp was leveraged to perform this correction, which proceeded as:

$$ref\_alt = (145366.45 * (1 - (\frac{mslp}{1013.25})^{0.190284})) * airport\_alt$$

where $airport\_alt$ is the altitude of the airport considering standard atmospheric conditions.

Next, for each trajectory, the touchdown point state vector's barometric altitude mea-

surement was subtracted from all other state vectors' barometric altitude measurements such that a height above ground level ($hagl$) is computed. Additionally, for each trajectory, the ground track distance between each state vector was computed such that the ground track distance from the touchdown point was able to be computed for each trajectory point. Additionally, the longitude and latitude coordinates were projected onto an (x, y) Cartesian coordinate system.

There are several options available regarding projecting the measured longitude and latitude coordinates onto an (xy) Cartesian coordinate system. A "perfect" projection would preserve area, shape, direction, bearing, and distance. However, preserving all of these properties is not feasible. Conformal map projections preserve angles locally, which enables the local scale in every direction around any one point to be constant, meaning that the ratio of two lengths in the local domain are preserved [141]. In the context of clustering trajectories within a terminal airspace radius, a conformal map projection is most appropriate as local distances are preserved. Two of the most widely-used conformal map projections are the UTM projection and the Lambert conformal conic projection. The UTM projection divides the earth into 60 segments, or "zones", where each of these zones are then flattened out with a transverse cylinder. The UTM projection is not appropriate in cases in which the area of interest spans multiple UTM zones. On the other hand, the Lambert conformal conic projection seats a cone over the sphere of the Earth and projects the surface conformally onto the cone, where the cone is then unrolled such that the parallel that was touching the sphere is assigned unit scale [141]. While no projection is necessarily more "accurate" than another, the preservation of different properties and differing levels of distortion are of consideration. Considering that KSFO is located nearly in the center (east-west-wise) of UTM Zone 10, and distortion is least closest to the center, a UTM projection was selected to project longitude and latitude coordinates onto an (x, y) Cartesian coordinate system.

Finally, section 4.2 introduced the need for augmentation of the data set with derived energy metrics. Related to **Experiment 2**, the SPE (SPE), SKE (SKE), and STER ($STER$)

were computed, where the SPE was simply considered to be the height above ground level. The velocity of the aircraft leveraged to compute the SKE was the ground speed of the aircraft available in ADS-B data. The choice of speed (true airspeed, ground speed, etc.) to use to compute the energy metrics of interest is not straightforward. Ground speed was selected due to relevance in certain ATM situations and data limitations. For instance, runway excursions, which are considered anomalies, occur in instances of excess ground speed on touchdown. Further, estimates related to time an aircraft reaches a specific point, whether it be touchdown or a specific point related to separation of two aircraft, are calculated using an aircraft's ground speed. The selection of other speeds, such as true airspeed, may be more relevant and necessary when aerodynamic effects, such as stall, are considered, which was not the focus of this work. Finally, other speed measurements were not readily available within the OpenSky Network historical database. The extracted ASOS data did include wind speed and direction, which could be leveraged to approximate an airspeed value. However, the ASOS data is measured at a ground station, which would require an assumption of homogeneous atmosphere within the 20 nautical mile radius of KSFO, which is not realistic. Accordingly, the ground speed was selected to compute the selected energy metrics.

Final Cleaning

The objective of the final cleaning step is to discard any trajectories that may be deemed incomplete. For instance, a trajectory was discarded if the difference between the touchdown point barometric altitude and the corrected airport altitude (ref_alt) was greater than a certain threshold, i.e. the trajectory did not reach within a certain $hagl$ threshold. These segments were discarded as not enough data exists to consider the approach trajectory as being complete. The setting of the $hagl$ threshold may be varied depending on the required application. For instance, in the context of the flight-level analysis (the **Air Traffic Flow Identification** and **Anomaly Detection** steps in the proposed methodology), it is advanta-

geous to have higher-fidelity, more complete ADS-B data and associated derived metrics due to the consideration of the individual flight trajectories. Therefore, to generate the data sets used to perform **Experiment 1** and **Experiment 2** a threshold of 100 ft was set. On the other hand, in the context of the **Airspace-Level Analysis** step of the proposed methodology, it is not as important that each individual trajectory reach exactly the airport location. Thus, to generate the data set used to perform **Experiment 3.1** and **Experiment 3.2**, a threshold of 1,000 ft was set. The discarding of trajectories based on $hagl$ accounts for trajectory records that do not reach the airport location.

However, it was also considered that trajectory data may not be available at the borders of the airspace, i.e. the trajectory state vectors exist only beginning somewhere in the center of the 20 nautical mile radius. Therefore, trajectories were discarded if the maximum ground track distance from touchdown was less than 20 nautical miles. The threshold of 20 nautical miles was selected as if an aircraft had entered the terminal airspace and followed a direct, straight path to the airport location, the minimum ground track distance from touchdown that could have been observed was 20 nautical miles.

5.1.3 Data Set Properties and Statistics

At the conclusion of these processing steps, two cleaned and augmented ADS-B data sets are generated: one to be used in the experimental approaches associated with the **Air Traffic Flow Identification** and **Anomaly Detection** steps and one to be used in the experimental approach associated with the **Airspace-Level Analysis** step. Both data sets contain arriving aircraft ADS-B data for both domestic and international flights. While the airspace-level analysis data set contains slightly more trajectories than the flight-level analysis data set due to a less stringent threshold value set for discarding trajectories based on $hagl$, both data sets contain arriving aircraft trajectories for around 180,000 both domestic and international flights. The flights were operated by 190 unique airlines, where United Airlines operated approximately one-third of the flights. Further, the data set contained 110

unique aircraft types, where the A320 aircraft was the most commonly operated aircraft. A bar char of the top five most common airlines and aircraft types for the approximately 180,000 flights in the two data sets is displayed in Figure 5.7.

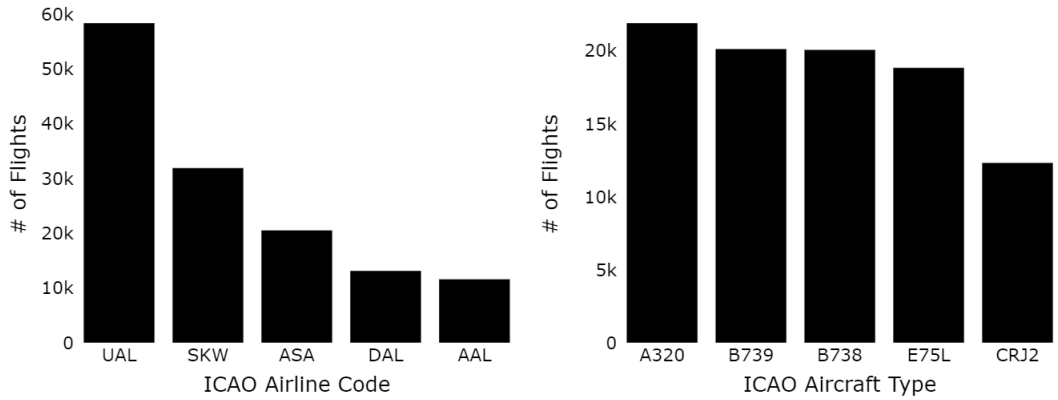

Figure 5.7: Top Five Most Common Airlines and Aircraft Types for Flight-Level and Airspace-Level Data Sets

5.2 Air Traffic Flow Identification

Executing the experimental approach designed in section 4.1 that is associated with filling **Gap 1** provides results that may be interpreted to enable the assessment of **Hypothesis 1 Acceptance**. Thus, the implementation of the experimental approach is detailed and the corresponding results are presented and discussed.

5.2.1 Implementation

The **Air Traffic Flow Identification** experimental approach involves **Experiment 1** and includes two main steps: (i) data processing and (ii) trajectory clustering. Thus, the implementations of both the data processing and trajectory clustering steps of the experimental approach to accomplish the objective of **Experiment 1** are described.

Data Processing

The flight-level analysis data set of all arriving aircraft operating within the KSFO terminal airspace in 2019 was leveraged to implement the experimental approach. This cleaned and augmented data set contained 178,890 trajectories.

A distance-based re-sampling method created the n-dimensional feature vectors for clustering, where the ground track distance from touchdown is the distance-based metric that the re-sampling was performed on. Each trajectory was re-sampled on 50 uniform intervals between 0 ground track distance from touchdown and the maximum ground track distance from touchdown associated with that trajectory. The average maximum ground track distance from touchdown value over all trajectories within the data set was approximately 32.7 nautical miles. The average distance between trajectory points over all trajectories within the data set was approximately 0.67 nautical miles.

While the trajectory clustering step may theoretically be implemented on the entire data set of 178,890 flights, this would only provide one "data point" from which to draw conclusions. Therefore, full-year data set was split into smaller data sets corresponding to smaller time intervals such that an adequate assessment of performance of the implementation of HDBSCAN with the WED may be performed. Specifically, the full-year data set was split into both daily and weekly data sets. Each daily data set contained around 500 trajectories, on average, whereas each weekly data set contained around 3,500 trajectories, on average.

Trajectory Clustering

For each data set in the data groups (daily and weekly),HDBSCAN was implemented with both the WED and ED, separately. The *hdbscan* [124] Python library was leveraged to perform the trajectory clustering with the two distance functions. The ED is the default distance function implemented by the *hdbscan* Python library. Therefore, a method of using the WED was required. The *hdbscan* Python library is capable of implementing any distance function that exists within the *sklearn* [142] Python library as well as a custom

distance function. Within the *sklearn* Python library there exists the implementation of a Standardized Euclidean Distance (SED). The SED is intended to scale the different features by their respective variances, represented by a variance vector $V = [v_1, v_2, ...v_k]$. Computing the SED between trajectories T^i and T^j proceeds as:

$$D_{SED}^{i,j} = \sqrt{\sum_{k=1}^{n} [\frac{(x_k^i - x_k^j)^2 + (y_k^i - y_k^j)^2}{v_k}]}$$

To produce the desired WED effect, the values of $\frac{1}{w_k}$, where w_k are the values of the weight vector, W, were assigned to V such that an effective weighting results. Selecting the "optimal" values of the weight vector is not straightforward and may require iteration. Therefore, to assess the performance of the WED compared with the ED, several different weight vectors, or weighting schemes, were evaluated. These different weight vectors were created by varying the values of w_k according to functions inspired by the shapes of the normal and beta distribution probability density and cumulative distribution functions. Four different weighting schemes were selected, where these weighting schemes are displayed in Figure 5.8. The point 0 is the touchdown point, whereas point 49 is the first point in which the aircraft enters the terminal airspace radius.

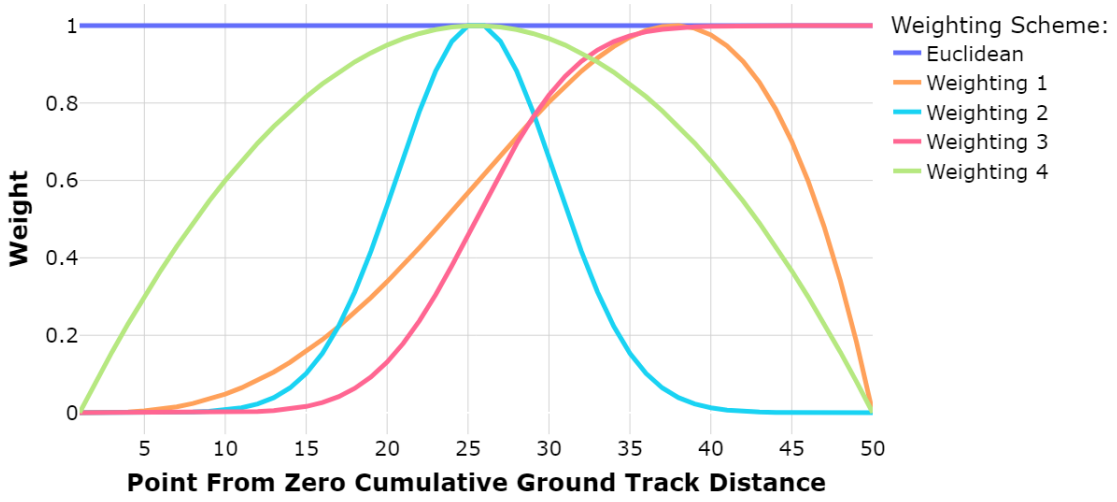

Figure 5.8: Weighting Schemes Evaluated

TO explore the method's robustness and underscore the importance of selecting an appropriate weighting scheme, Weighting 2 was included. Considering Weighting 2, only the trajectory points in the "center" of the trajectory are given any significant weight, whereas the majority of the points are weighted near zero. It was anticipated that a "less reliable" clustering results from usage of Weighting 2 when compared with the performance of the other weighting schemes due to the emphasis on a limited sub-sequence of trajectory points.

The parameter minimum cluster size parameter was set to 2% of the total number of trajectories within each data set. On average, for the daily data sets, this was an average of 10, whereas, for the weekly data sets, this was an average of 70. These values were relatively consistent with other values set for the minimum cluster size parameter (analogous to the DBSCAN $MinPts$ parameter) for other studies' applications of DBSCAN or HDBSCAN [73, 87]. The optional HDBSCAN smoothing parameter is varied to provide some "control" over the fraction of outliers detected. For context on the scale of this fraction, Gariel et al. [63] find typical fractions of outlying trajectories per day to be between 0.02 to 0.16. As this range is too broad to enable a fair comparison between the two implementations of HDBSCAN, a tighter range of fraction of outliers detected between 0.10 to

0.15 was selected to remove the outlier bias. In this research, for each data set, an iterative loop was implemented to adjust the smoothing parameter such that the fraction of outlying trajectories was restricted to between 0.10 and 0.15.

Finally, as introduced in section 4.1, to compare the WED weighting schemes' performance with that of the ED, the mean ED from the assigned cluster centroid was computed for each data set within the daily and weekly data groups for the multiple implementations of HDBSCAN with the WED considering the different weighting schemes and the implementation of HDBSCAN with the ED. Computation of the mean ED from the assigned cluster centroid for each HDBSCAN implementation enabled a direct comparison of performance of the different WED weighting schemes with respect to the ED as well as the performance of each of the weighting schemes with respect to each other.

5.2.2 Results and Discussion

Leveraging the *hdbscan* Python library to cluster both the daily and weekly data set using the various WED weighting schemes and the ED, the computation time difference between using the WED versus the ED was negligible. The results are presented for both the daily and weekly data groups separately. Only 139 daily data sets and 30 weekly data sets were ultimately available for comparison due to an inability to reach a fraction of outliers between 0.10 and 0.15 for all WED weighting schemes and/or the ED. For each data group, for each data set, value of the mean ED from the assigned cluster centroid for was compared for each WED weighting scheme to the value of the mean ED from the assigned cluster centroid for the ED. For each data group, the percentage of data sets for which a weighting scheme performed "better" than the ED with respect to the mean ED from the assigned cluster centroid is presented in a bar chart. Results for the daily and weekly data groups are presented.

Daily Data Group

Figure 5.9 displays the percentage of time the mean ED from the assigned cluster centroid is greater for the implementation of HDBSCAN with the ED as compared to the implementation of HDBSCAN with the WED, for each weighting scheme. With respect to Figure 5.9, three of the weighting schemes, Weighting 1, Weighting 3, and Weighting 4, appear to identify air traffic flows more reliably (perform better) than the ED on the majority of data sets in the daily data group. Specifically, Weighting 1 notably performs the "best", followed by Weighting 3, where both of these weighting schemes have a similar skewed shape, with the primary difference being in the weighting values of the ten trajectory points closest to the terminal airspace border. As expected, Weighting 2 does not perform well compared to the other weighting schemes. Weighting 4 does not perform as well as Weighting 1 and Weighting 3, yet still performs better than the ED.

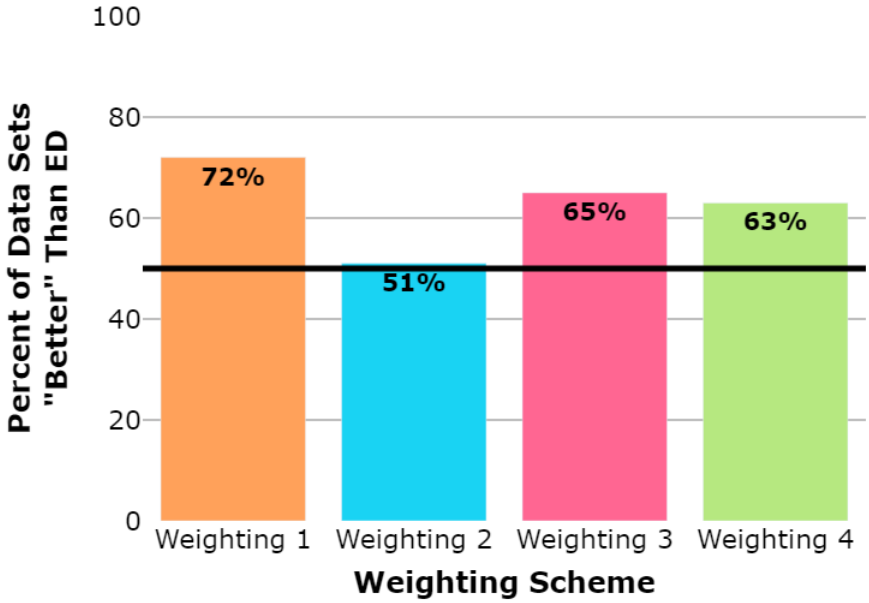

Figure 5.9: Daily Data Group, Percentage of Time the Mean Euclidean Distance from Assigned Cluster Centroid is Greater for the Implementation of HDBSCAN with the Euclidean Distance as Compared to the Implementation of HDBSCAN with the Weighted Euclidean Distance

Weighting 1 was selected as the weighting scheme best-suited to identify air traffic flows for arriving aircraft at KSFO. To further assess results for individual daily data sets, clusters identified for a single daily data set are graphically presented in Figure 5.10 and Figure 5.11 for the HDBSCAN implementation with the WED Weighting 1 and the HDBSCAN implementation with the ED, respectively. Both the implementation of HDBSCAN with the WED Weighting 1 and implementation of HDBSCAN with the ED identified five distinct clusters.

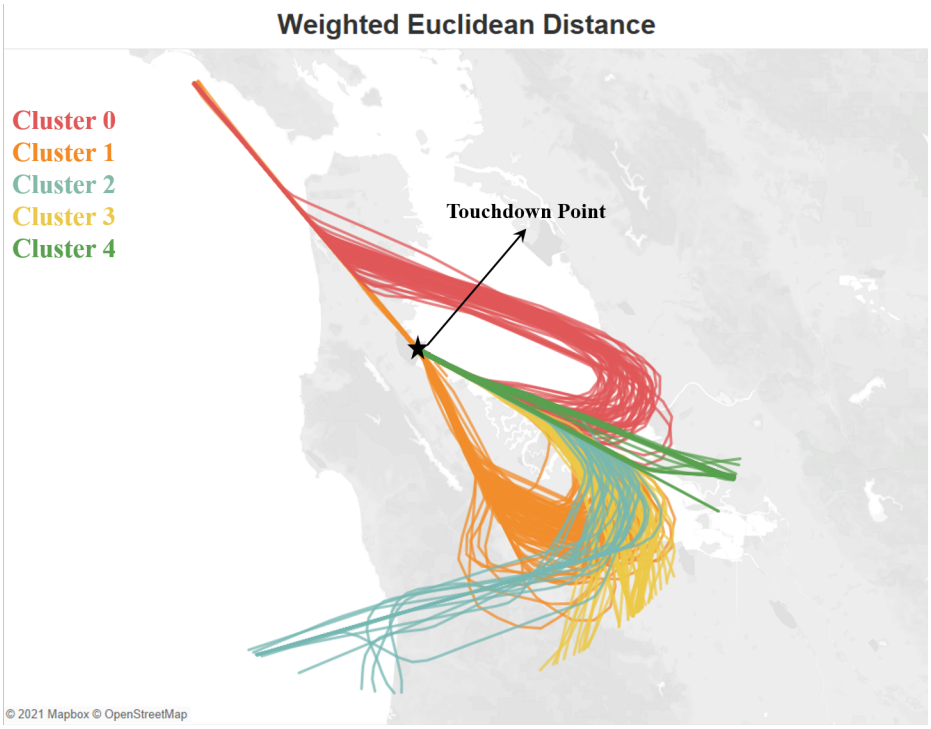

Figure 5.10: Clusters Identified Implementing HDBSCAN with the Weighted Euclidean Distance Weighting 1 for a Single Day, Jun 10^{th}, 2019

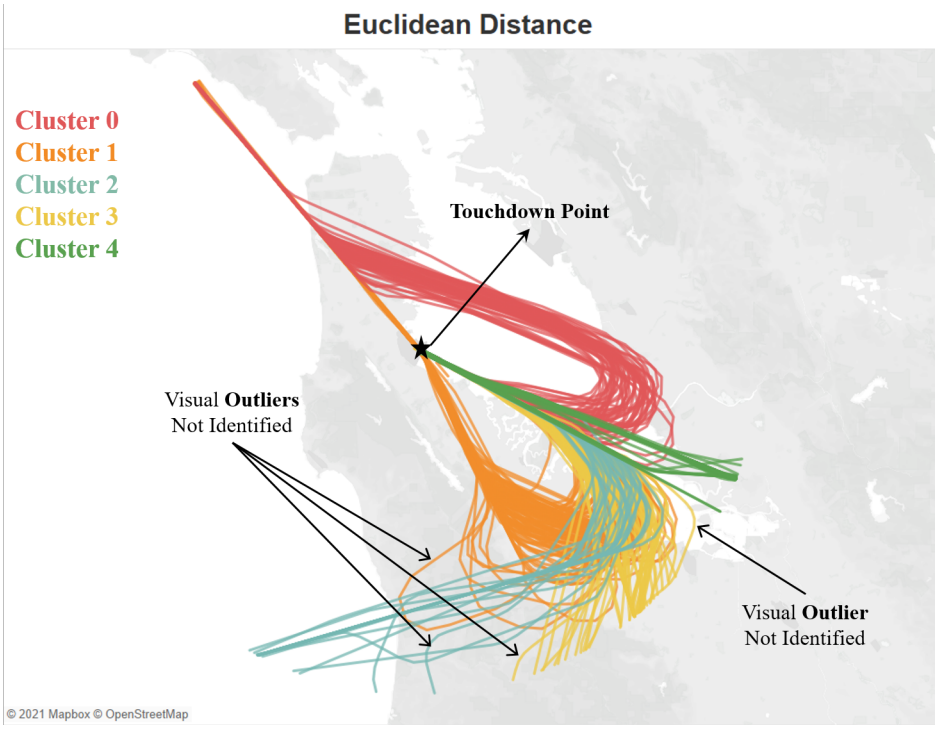

Figure 5.11: Clusters Identified Implementing HDBSCAN with the Euclidean Distance for a Single Day, June 10^{th}, 2019

Considering the implementation of HDBSCAN with the ED, some trajectories were assigned to clusters that, visually, appear as likely outliers. On the other hand, considering the implementation of HDBSCAN with the WED Weighting 1, these trajectories were identified as outliers. MAny of the clusters (representative of the basis of air traffic flows) identified implementing HDBSCAN with WED Weighting 1 appeared to be “tighter”, or more reliable, which was one of the primary objectives in implementing HDBSCAN with the WED. Figure 5.12 and Figure 5.13 display Cluster 0 and Cluster 4, respectively, in more detail such that the differences may be further evaluated.

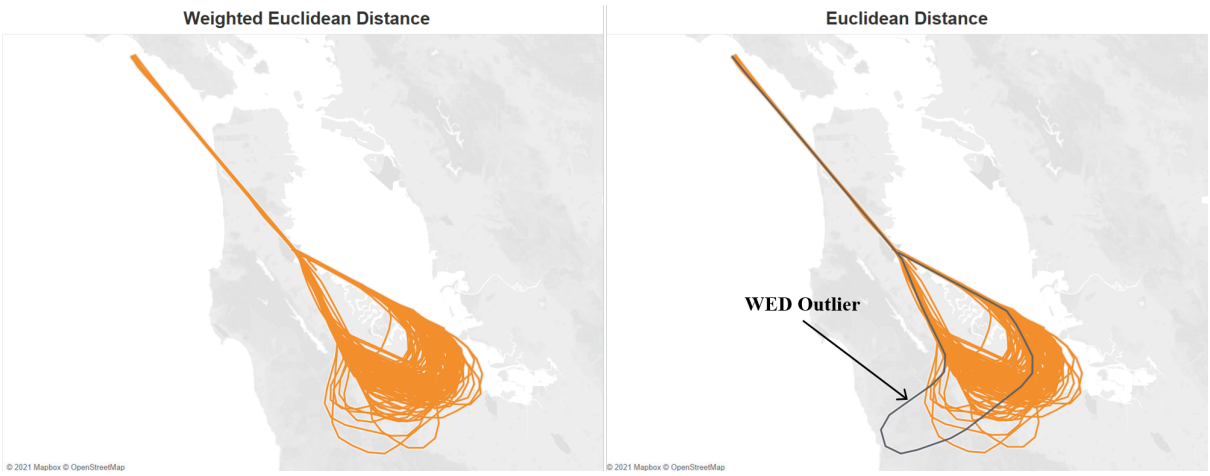

Figure 5.12: Cluster 0 Identified Implementing HDBSCAN with the Weighted Euclidean Distance Weighting 1 versus Implementing HDBSCAN with the Euclidean Distance for the Single Day, June 10^{th}, 2019

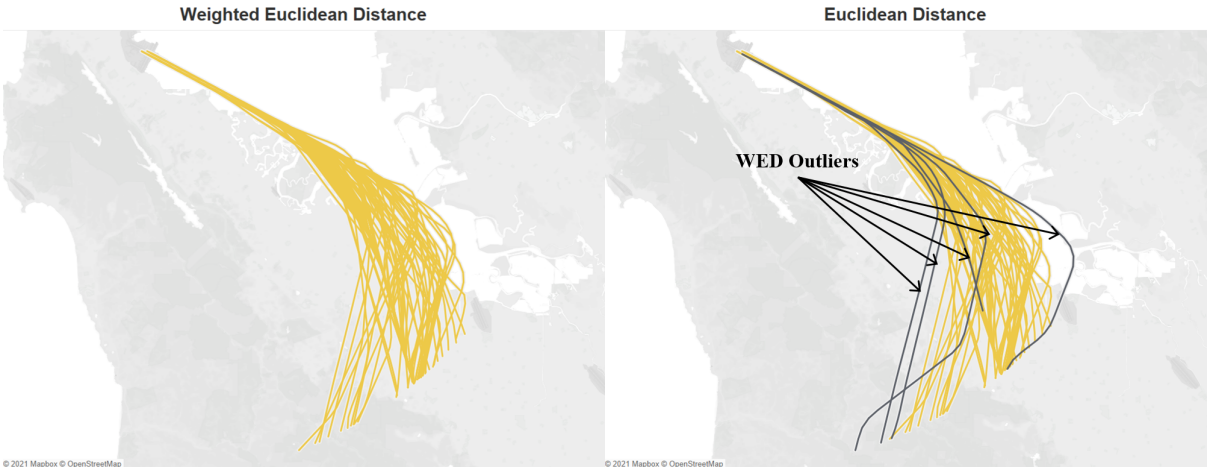

Figure 5.13: Cluster 4 Identified Implementing HDBSCAN with the Weighted Euclidean Distance Weighting 1 versus Implementing HDBSCAN with the Euclidean Distance for the Single Day, June 10^{th}, 2019

The dark gray lines in Figure 5.12 and Figure 5.13 represent trajectories that were not identified as outliers implementing HDBSCAN with the ED, yet were identified as outliers implementing HDBSCAN with the WED Weighting 1. In Figure 5.12 and Figure 5.13, the trajectories represented by the dark gray lines were not as nominally conforming with the dominant air traffic flow paths or other trajectories within the cluster, and, therefore, should have been flagged as outliers (anomalies). However, implementation of HDBSCAN with

the ED was unable to detect these trajectories appropriately as outliers, which potentially results in the skewing of the overall air traffic flows identified. Overall, the quantitative results obtained through assessment of the mean ED from the assigned cluster centroid (displayed in Figure 5.9) are further supported by a visual inspection of Figure 5.10, Figure 5.11, Figure 5.12, and Figure 5.13.

Weekly Data Group

Figure 5.14 displays the overall results for all 30 data sets included in the weekly data group. Two of the weighting schemes, Weighting 1 and Weighting 3, appear to identify air traffic flows much more reliably than the ED. Notably, Weighting 2 performs worse than the ED, where its undesirable performance, again, is expected. For the weekly data group data sets, Weighting 4 does not perform better for a majority of the weeks considered, which is likely due to the inferior performance of symmetric weighting schemes being more pronounced when clustering data sets including many more trajectories and, potentially, more air traffic flows.

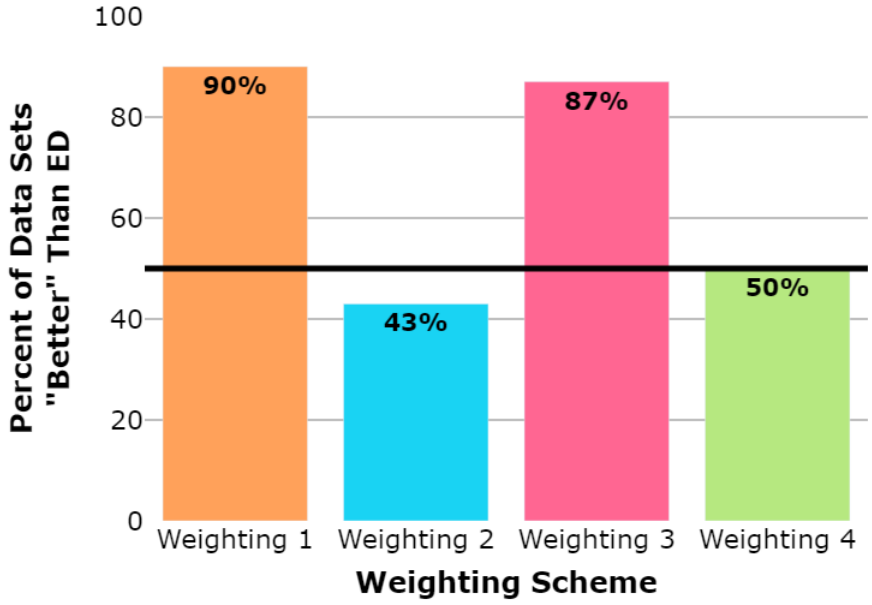

Figure 5.14: Weekly Data Group, Percentage of Time the Mean Euclidean Distance from Assigned Cluster Centroid is Greater for the Implementation of HDBSCAN with the Euclidean Distance as Compared to the Implementation of HDBSCAN with the Weighted Euclidean Distance

Similar to the analysis of the daily data set results, an individual weekly data set is visually assessed to further support the quantitative results. Figure 5.15 and Figure 5.16 display the clusters identified through implementation of HDBSCAN with the WED Weighting 1 and the implementation of HDBSCAN with the ED, respectively. Both the implementation of HDBSCAN with the WED Weighting 1 and the implementation of HDBSCAN with the ED identify seven distinct clusters. Again, it is observed that considering the implementation of HDBSCAN with the ED that some trajectories were assigned to clusters that, visually, appear as likely outliers. On the other hand, considering the implementation of HDBSCAN with the WED Weighting 1, these trajectories were identified as outliers. Additionally, visually, some clusters identified implementing HDBSCAN with WED Weighting 1 appear to be “tighter”, or more reliable. Therefore, results for the weekly data group were similar to those for the daily data group.

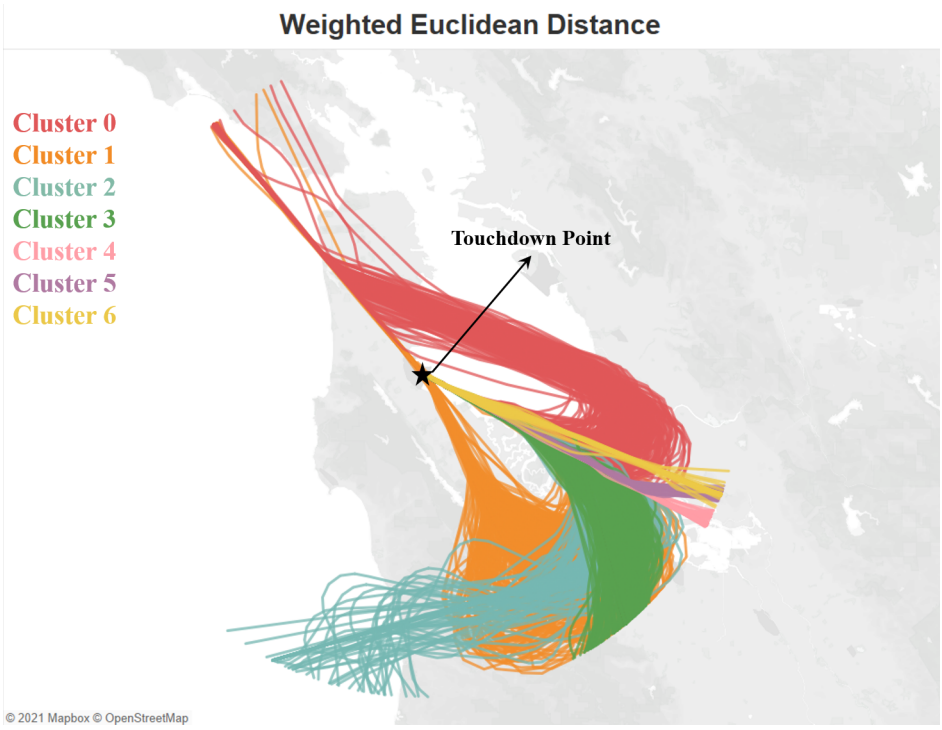

Figure 5.15: Clusters Identified Implementing HDBSCAN with the Weighted Euclidean Distance Weighting 1 for a Single Week, Week of June 11^{th}, 2019

Figure 5.16: Clusters Identified Implementing HDBSCAN with the Euclidean Distance for a Single Week, Week of June 11^{th}, 2019

The results of the execution of the experimental approach involving **Experiment 1** outlined in section 4.1 indicate that, in the majority of cases, there does exist a weighting scheme for implementation of HDBSCAN with the WED that out-performs the implementation of HDBSCAN with the ED. This indicates a more reliable identification of air traffic flows. Quantitatively, the mean ED from the assigned cluster centroid was lower in the majority of cases in which HDBSCAN was implemented with the WED than in which HDBSCAN was implemented with the ED, where each implementation results in the fraction of outliers detected between 0.10 and 0.15. Therefore, **Hypothesis 1 is accepted**.

It is observed that the weighting schemes that placed more weight on the trajectory points between the middle of the trajectory point sequence and those approaching the bor-

der of the airspace, yet not necessarily at the border, performed better when considering aircraft arriving at KSFO. However, the most appropriate weighting scheme, or type of weighting scheme, may be different for different airports depending on the runway configuration and other terminal airspace-specific factors. Further, the more reliable identification of air traffic flows inherently involves a more robust identification of outlying, or spatially anomalous, trajectories, which is relevant in the context of the **Anomaly Detection** step in the proposed methodology.

The acceptance of **Hypothesis 1** fills **Gap 1**. The primary contribution of filling **Gap 1** is a more reliable method of identifying air traffic flows considering the converging nature of arriving aircraft trajectories.

5.3 Anomaly Detection

Executing the experimental approach designed in section 4.2 that is associated with filling **Gap 2** provides results that may be evaluated to enable the assessment of **Hypothesis 2.1 Acceptance**, **Hypothesis 2.2 Acceptance**, and **Hypothesis 2.3 Acceptance**. Thus, the implementation of the experimental approach is detailed and the corresponding results are presented and discussed.

5.3.1 Implementation

The **Gap 2** experimental approach involves **Experiment 2** and includes four main steps: (i) data processing, (ii) spatial anomaly detection, (iii) energy anomaly detection, and (iv) statistical analysis set-up. Thus, the implementations of the data processing, spatial anomaly detection, energy anomaly detection, and statistical analysis set-up steps of the experimental approach to accomplish the objective of **Experiment 2** are described.

Data Processing

A subset of the flight-level analysis data set of all arriving aircraft operating within the KSFO terminal airspace in 2019 was selected to implement the experimental approach due to computational limitations, i.e. a smaller number of trajectories were considered. The time period considered was four months of 2019 in that trajectories of aircraft arriving at KSFO between June 1^{st}, 2019 and September 30^{th}, 2019 were considered. This data set contained 63,630 total flights.

As introduced in section 4.2, two data sets were required: (i) radius data for spatial anomaly detection and (ii) distance cutoff data for energy anomaly detection. Re-sampling the trajectories within the 20 nautical mile radius of KSFO to 50 uniformly-spaced points generated the radius data set, whereas re-sampling the trajectories cutoff at the last 20 nautical miles of flight to 50 uniformly-spaced points generated the distance cutoff data set. The radius data set contains trajectories an average of around 0.65 nautical miles ground track distance between trajectory points. The distance cutoff data set contains trajectories with approximately 0.4 nautical miles ground track distance between trajectory points.

Go-arounds were identified such that within the data set, 337 of the 63,360 flights were classified as go-arounds, making up only about 0.53% of the data set. Therefore, a higher go-around rate than the 0.29% presented in [143] was observed. The go-around trajectories were removed from the data set leaving each data set (radius data and distance cutoff data) containing 63,023 trajectories.

Spatial Anomaly Detection

As prescribed in section 4.2 and considering the acceptance of **Hypothesis 1**, the trajectory clustering supporting the spatial anomaly detection step was performed by implementing HDBSCAN with the WED. Specifically, the Weighting 1 weighting scheme (displayed in Figure 5.8) was selected due to its superior performance observed in **Experiment 1**. Recall, a beta probability distribution inspired the Weighting 1 weighting scheme.

Similar to **Experiment 1**, the HDBSCAN Python library was leveraged to implement HDBSCAN with the WED such that air traffic flows were identified and spatial anomalies were detected. The selection of the required minimum cluster size parameter is not straightforward and is dependent somewhat on operational intuition as well as the evaluation of the fraction of anomalies HDBSCAN detects at differing parameter settings. In the context of operational intuition, it is relevant to consider the fraction of trajectories that could comprise a single air traffic flow. During the analysis period, approximately 525 trajectory records exist per day, where, within a single day, it is possible to observe multiple air traffic flows. It is introduced that SFO's most prominent official operational plan, the west plan (Figure 5.2, is observed 95-98% of the time. The west plan contains five flight arrival patterns, which may generally represent possible air traffic flows. If all five possible "air traffic flows" are active, it is unlikely the average of 525 trajectories would be distributed evenly among them and some portion of the trajectories will likely be spatial anomalies. Despite these considerations, the average number of trajectories per "air traffic flow" was considered to be 105. Therefore, it was asserted to the be considered an air traffic flow over the entire 120-day time period, the minimum value for the minimum cluster size parameter should be no lower than the average number of trajectories per "air traffic flow" for a 6-day period (5% of 120 days). Hence, a minimum value of the minimum cluster size parameter was set at 630, which is equivalent to approximately 1% of the total number of trajectories within the data set. An HDBSCAN hyper-parameter sensitivity analysis was performed for minimum cluster size parameter values of approximately 1% of the total number of trajectories to approximately 3% of the total number of trajectories, in 0.5% increments. The results of this sensitivity analysis are presented in Figure 5.17. Further, it was relevant to consider that anomalies are, by definition, rare event and should not make up a large fraction of the data set. Because the number of clusters identified remains constant for the minimum cluster size parameter values evaluated, the selected value of the minimum cluster size parameter was 630 (1% of the total number of trajectories) as

this was the value producing the lowest fraction of outliers (0.128).

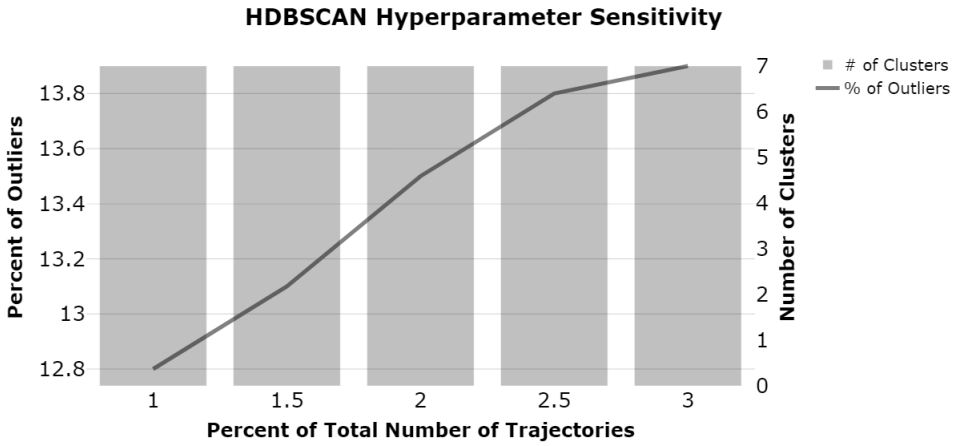

Figure 5.17: HDBSCAN Sensitivity to Percent of Total Number of Trajectories Taken as the Minimum Cluster Size Parameter

After implementing HDBSCAN with the WED Weighting 1 and minimum cluster size parameter equal to 630, 8,082 trajectories were detected as spatial anomalies. Additionally, seven distinct clusters were identified, where these cluster are displayed in Figure 5.18. Figure 5.19 presents the percentages of trajectories assigned to one of the seven clusters, assigned a spatial anomaly label by HDBSCAN, or assigned a go-around label a priori. The cluster sizes are not uniform. For instance, Cluster 4 contained almost 30% of the total number of trajectories, while Cluster 5 contained only about 4% of the total number of trajectories.

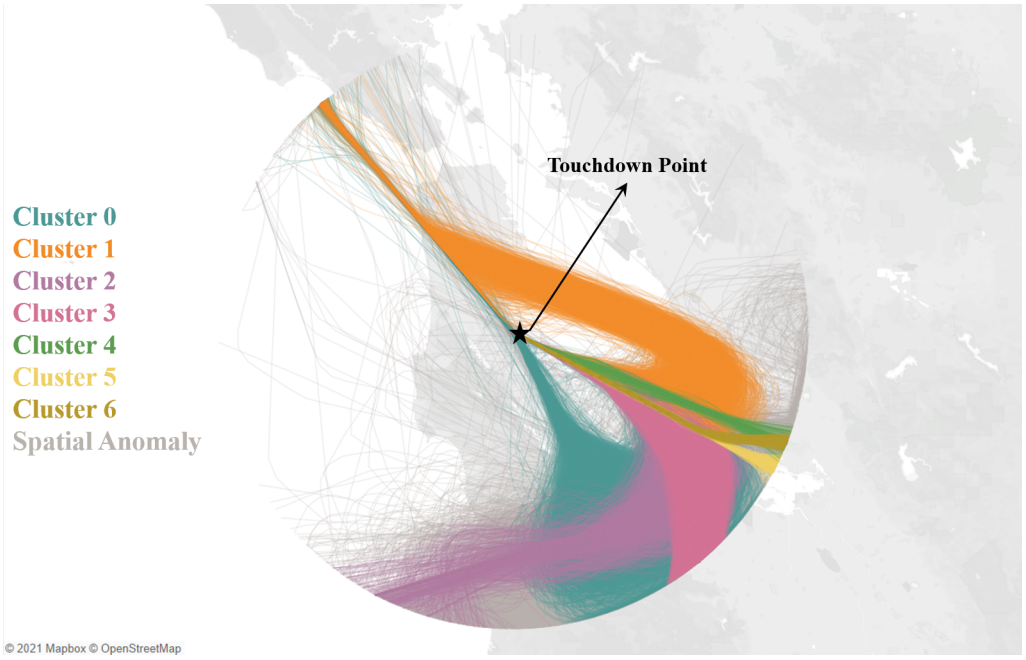

Figure 5.18: Clusters Identified by Implementing HDBSCAN with the Weighted Euclidean Distance Weighting 1

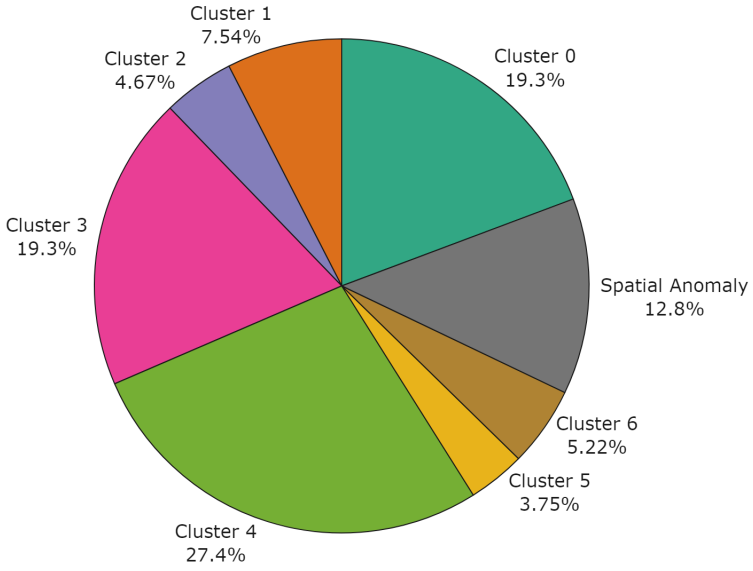

Figure 5.19: Distribution of Trajectories

The clusters identified by the implementation of HDBSCAN with the WED Weighting 1 are representative of seven distinct air traffic flows. Spatially anomalous trajectories were subsequently associated with the clusters to which they most closely belong to form air traffic flows and to enable robust energy anomaly detection. Thus, *air traffic flow*, in this context, refers to all spatially nominal trajectories belonging to a specific cluster as well as all spatially anomalous trajectories closest (according to the WED Weighting 1) to the specific cluster's centroid compared with all other air traffic flow centroids. Figure 5.20 displays the air traffic flows identified by implementing HDBSCAN with the WED Weighting 1.

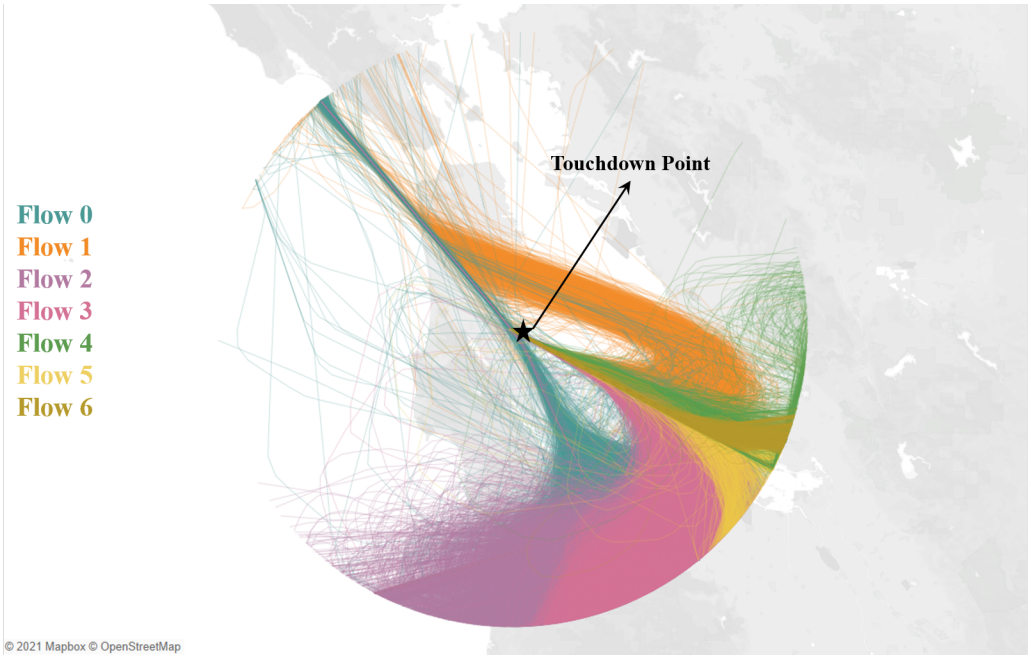

Figure 5.20: Air Traffic Flows Identified by Implementing HDBSCAN with the Weighted Euclidean Distance Weighting 1

Energy Anomaly Detection

The entire distance cutoff data set was separated into smaller air traffic flow data sets. The number of trajectories assigned to each energy anomaly detection data set and the percentage of trajectories detected as spatial anomalies associated with each air traffic flow are presented in Table 5.3.

Table 5.3: Composition of Distance Cutoff Air Traffic Flow Data Sets

Data Set	# Total Trajectories	% Spatial Anomalies
Flow 0	12,764	4.8
Flow 1	4,953	4.1
Flow 2	5,078	42.0
Flow 3	15,217	20.2
Flow 4	18,059	4.2
Flow 5	3,075	23.2
Flow 6	3,877	15.1

To prevent feature bias in the implementation of DBSCAN with the ED due to varying magnitudes of the features (SPE, SKE, $STER$), a normalization was performed on each air traffic flow data set prior to the clustering. Specifically, a z-normalization was conducted for each feature x^i within each data set, which is computed as:

$$x^i_{norm} = \frac{x^i - \overline{x^i}}{s_{x^i}}$$

where s_{x^i} is the standard deviation associated with feature x^i and $\overline{x^i}$ is the mean associated with feature x^i. The *sklearn* Python library provided the module necessary for the z-normalization of the energy metrics.

DBSCAN is implemented leveraging the *sklearn* Python library's DBSCAN module. The $MinPts$ parameter was set at approximately 1% of the total number of trajectories within each air traffic flow data set, which was consistent with the minimum number of samples setting in [66]. The exact value for each air traffic flow data set was dependent on the number trajectories associated with each identified air traffic flow in the spatial anomaly detection step. The ε parameter was varied until approximately a fraction of 0.10 (between 0.09 and 0.11) of all trajectories in each data set were detected as energy anomalies. This

fraction was consistent with the fraction of anomalies detected in previous studies [47, 100]. As anticipated, for all air traffic flows, DBSCAN did identify only one cluster of nominal operations. Finally, the anomaly score for each trajectory was computed, as specified in section 3.2.

Statistical Analysis Set-Up

The statistical analysis set-up step in the experimental approach is relatively straightforward. First, within each data set, the trajectories were assigned to one of the four categories: (i) nominal (N), (ii) only-spatial-anomaly (S), (iii) only-energy-anomaly (E), or (iv) both-spatial-and-energy-anomaly (B). The distribution of these categories for each air traffic flow is displayed in Figure 5.21. These distributions do differ between air traffic flows; although, the sum of the proportions of the both-energy-and-spatial-anomaly (B) category and only-energy-anomaly (E) category were constant and equal to approximately 0.10 (the specified fraction of energy anomalies detected).

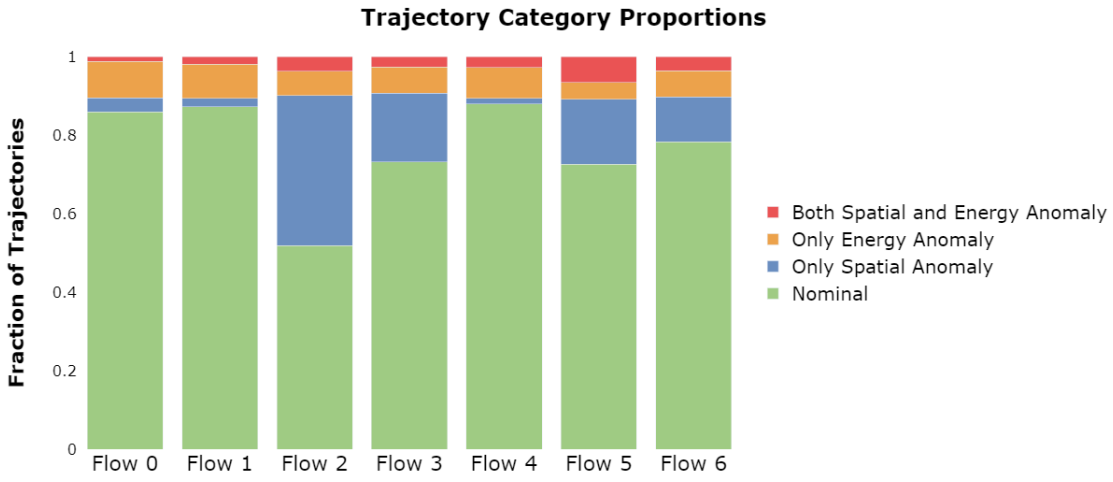

Figure 5.21: Trajectory Category Distributions within Each Air Traffic Flow Data Set

For each air traffic flow, the fraction of energy anomalies detected considering only spatial anomalous trajectories and the fraction of energy anomalies considering only spatially

nominal trajectories are plotted in Figure 5.22. The likelihood ratio was then computed as the ratio of these two values for each air traffic flow.

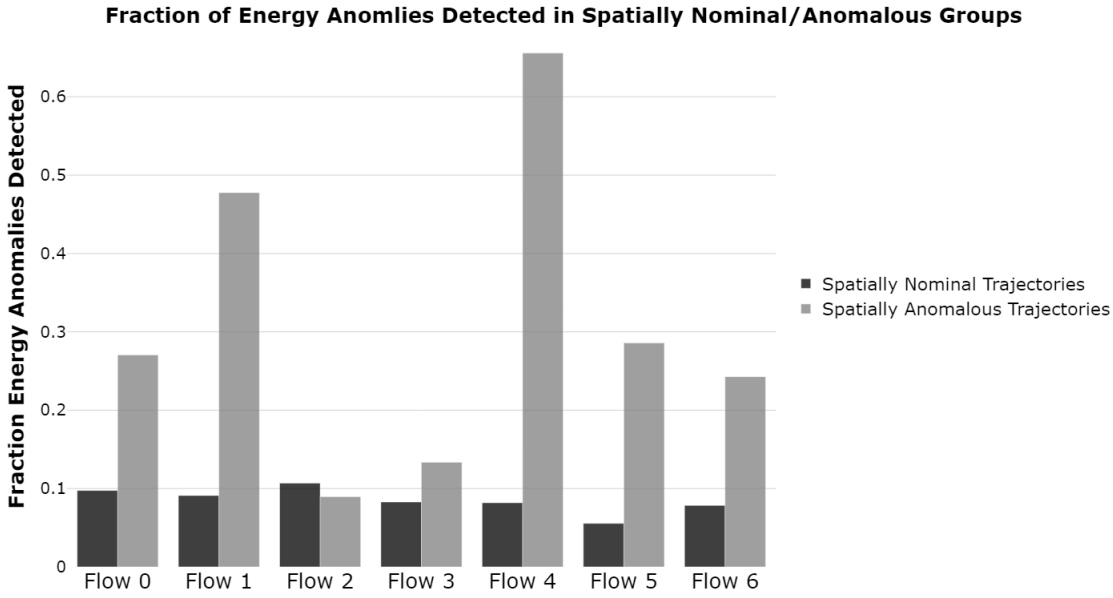

Figure 5.22: Fraction of Energy Anomalies Detected Among Spatially Anomalous Trajectories and the Fraction of Energy Anomalies Detected Among Spatially Nominal Trajectories

In addition, for each air traffic flow, the mean value of the anomaly scores with respect to trajectory category were computed. Further considering all trajectories across all air traffic flows, the mean, median, standard deviation, 25^{th} percentile, and 75^{th} percentile values of anomaly scores with respect to trajectory category were computed, where hese values are displayed in Table 5.4. Finally, Figure 5.23 displays the distributions of the base-10 logarithm of aforementioned grouping of anomaly scores. The base-10 logarithm of the anomaly scores is presented such that distinctions between the anomaly score distributions are more readily observable.

Table 5.4: Anomaly Score Statistical Properties by Trajectory Category

Category	Mean	Median	Std	25^{th}	75^{th}
N	13.67	13.43	1.56	12.53	14.56
S	13.96	13.73	1.64	12.72	14.93
E	21.43	18.95	9.41	17.01	22.05
B	34.28	27.21	21.64	19.39	42.71

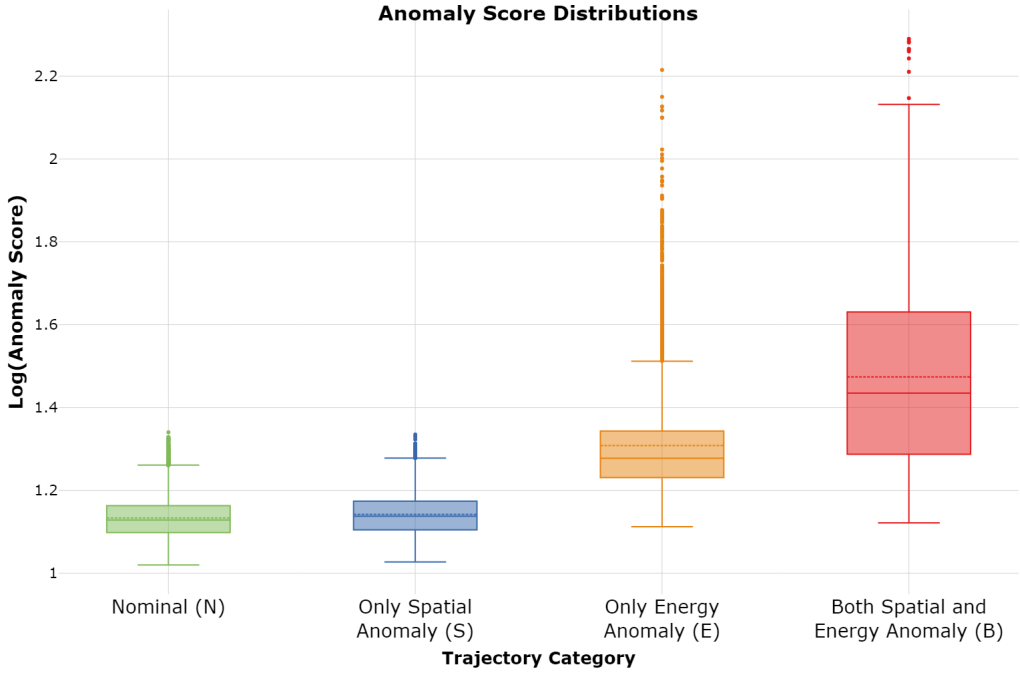

Figure 5.23: Trajectory Category Anomaly Score Distributions with the Base-10 Logarithm of the Anomaly Score Presented on the Y-Axis

5.3.2 Results and Discussion

Figure 5.22 reveals that, generally, the fraction of spatial anomalies detected as energy anomalies is greater than the fraction of spatially nominal trajectories detected as energy anomalies. For example, within the Flow 4 data set, 65.6% of spatial anomalies were detected as energy anomalies, whereas only 8.2% of spatially nominal trajectories were detected as energy anomalies. In general, for all air traffic flows flows, with the exception

of Flow 2, the fraction of energy anomalies detected is noticeably higher in the group of spatial anomalies than it is in the group of spatially nominal trajectories. Due to the unique results associated with Flow 2, a further investigation proceeded.

Both Table 5.3 and Figure 5.21 indicate the Flow 2 consists of an usually large percentage of spatial anomalies (42.0%). Therefore, the trajectories associated with Flow 2 were plotted. It was evident that the majority of the spatial anomalies detected that were associated with Flow 2 *should* have been initially assigned by HDBSCAN to strictly belong to Flow 2 (rather than be detected as anomalies). All other air traffic flows were plotted to assess if similar issues were present, yet no issues were observed. Because the large majority of *detected* spatial anomalies do not appear to be *actual* spatial anomalies associated with Flow 2, the results were skewed and conclusions were difficult to draw. It is likely that the fraction of spatial anomalies detected as energy anomalies was driven down for Flow 2, whereas the fraction of spatially nominal trajectories detected as energy anomalies was driven up. Thus, this explained the lack of alignment of the Flow 2 results with the other air traffic flow results and the results for Flow 2 are generally disregarded.

Figure 5.22 indicates that spatial anomalies were more likely to be detected as energy anomalies than spatially nominal trajectories. While the discovery of this relationship is significant, it is equally relevant to note that not all spatial anomalies are subsequently detected as energy anomalies. Table 5.5 displays the likelihood ratio, where this metric provides a quantitative measure of how much more or less likely it was for a trajectory to be detected as an energy anomaly provided that the trajectory was detected as a spatial anomaly. Considering Flow 4, the likelihood ratio was 8.01, which indicates that spatial anomalies associated with Flow 4 are just over eight times more likely to be subsequently detected as energy anomalies than the spatially nominal trajectories.

Table 5.5: Spatial Anomaly Detection and Energy Anomaly Detection Relationship Exploration

Data Set	Likelihood Ratio	(S Mean)/(N Mean)	Welch's t-test p-value
Flow 0	2.77	1.02	0.000
Flow 1	5.24	1.02	0.024
Flow 2	0.84	0.99	1.000
Flow 3	1.61	1.02	0.000
Flow 4	8.01	1.03	0.001
Flow 5	5.15	1.02	0.000
Flow 6	3.10	1.04	0.000

The results of the execution of the experimental approach involving **Experiment 2**, indicate that, on average, across all flows, excluding the Flow 2 data set due to skewed results,spatial anomalies trajectories are between 1.6 to 8 times more likely to be identified as energy anomalies than spatially nominal trajectories. Therefore, **Hypothesis 2.1 is accepted**.

Considering trajectories belonging to Flow 3, the horizontal speed (ground speed) profile and the vertical speed (vertical rate) profile for spatially nominal trajectories and spatially anomalous trajectories are displayed in Figure 5.24. These profiles represent the interquartile range of the ground speed and vertical rate, respectively. It is expected that both the spatial profiles and speed profiles of the spatially nominal trajectories and spatially anomalous trajectories would converge for the final five nautical miles of flight in the case that a successful touchdown occurs. Therefore, if the final five nautical miles of flight are neglected, it is evident that the interquartile range of both the ground speed and vertical rate were not aligned. For instance, spatially nominal trajectories typically experienced ground speeds that were higher than those of spatially anomalous trajectories, where the converse was true considering vertical rate. Necessarily, several factors may impact this observation; however, it is possible that aircraft experience spatially anomalous conditions were

provided different instructions regarding ground speed and vertical rate than those aircraft that experience spatially nominal conditions. Thus, ATC and other ATM system operators should be aware of the potential for non-alignment of the speed profiles and, hence, off-nominal energy management, to manifest in situations in which aircraft are experiencing off-nominal spatial operations, regardless of whether the off-nominal spatial operations are a result of ATC instructions or pilot actions. Moreover, the prediction of aircraft arrival times could be less accurate when the more rare spatially anomalous trajectories are observed., which may have efficiency implications.

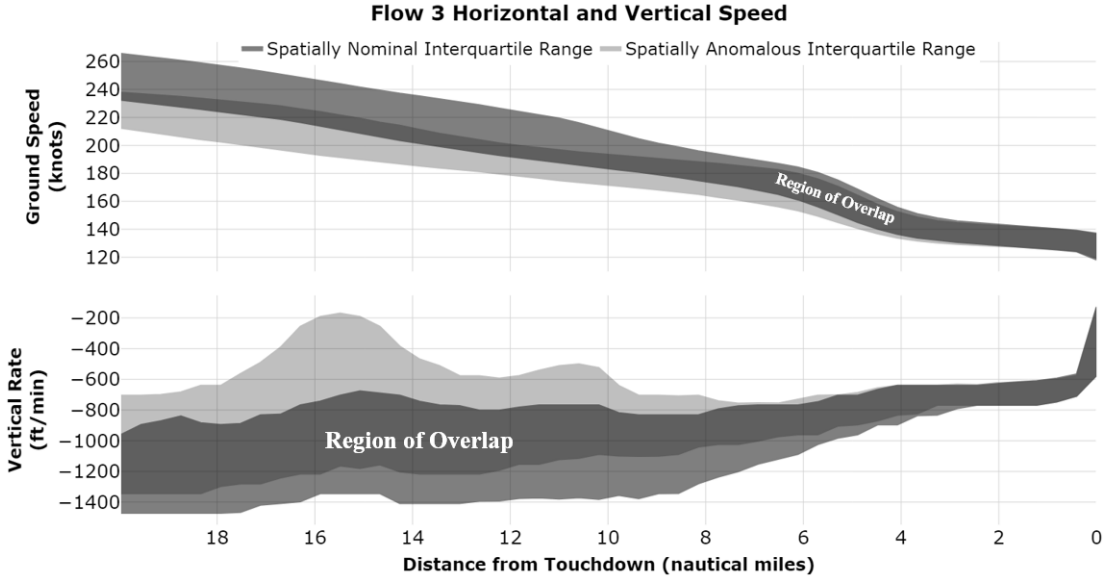

Figure 5.24: Flow 3, Speed Profiles of Spatially Nominal Trajectories versus Spatially Anomalous Trajectories

Further, the ratios between the mean anomaly score of each three categories containing trajectories (S, E, and B) identified as anomalous and the mean anomaly score of nominal (N) trajectories, respectively, were computed. Figure 5.25 displays these ratios of mean anomaly scores for each air traffic flow. Neglecting Flow 2, the ratios between the mean anomaly scores of only-spatial-anomaly (S) trajectories and the mean anomaly scores of nominal (N) trajectories were greater than 1. A value greater than 1 indicates that the anomaly scores of only-spatial-anomaly trajectories were, on average, greater than the

anomaly scores of nominal trajectories. Hence, this implies that, despite not being detected as energy-anomalous, only-spatial-anomaly trajectories were "relatively more anomalous" than nominal trajectories. Therefore, a higher degree of caution should still be exercised with respect to aircraft experiencing off-nominal spatial operations, even if their energy profiles appear to conform to nominal operations.

However, due to the closeness of the ratios between the mean anomaly scores of only-spatial-anomaly (S) trajectories and the mean anomaly scores of nominal (N) trajectories, it is advantageous to assess the statistical significance of the observation. Accordingly, the one-sided Welch's t-test was conducted to determine whether the mean anomaly scores of only-spatial-anomaly (S) trajectories were statistically significantly greater than the mean anomaly scores of nominal (N) trajectories. The *scipy* [144] Python library was leveraged to perform the one-sided Welch's t-test, which tests for the null hypothesis that the two samples (S and N) have identical mean values [130].

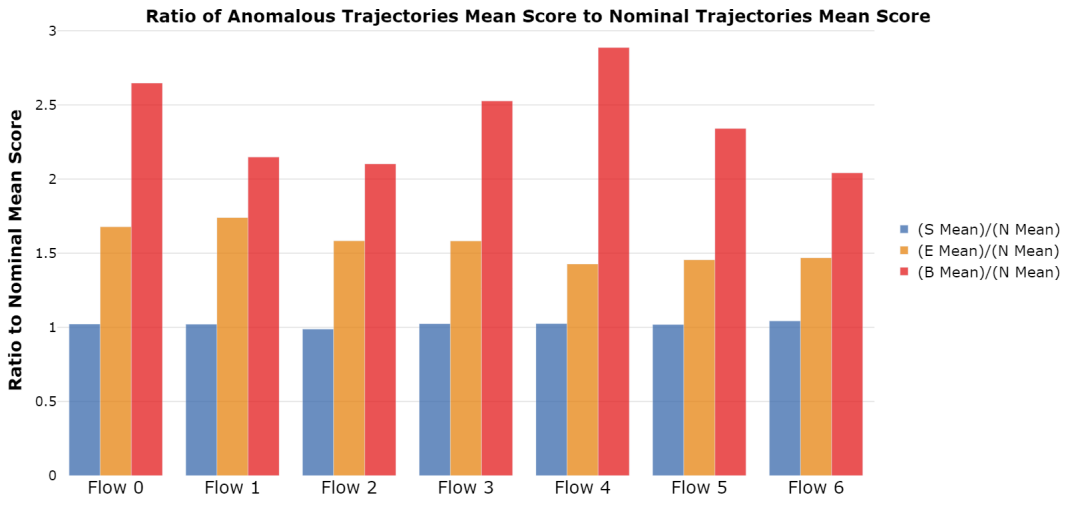

Figure 5.25: Anomaly Scores Mean Ratios by Air Traffic Flow

The computed p-values for the one-sided Welch's t-test are displayed in Table 5.5. Again, neglected Flow 2, the p-values computed were less than the commonly-applied 0.05 statistical significance threshold, which indicated that the null hypothesis (the mean

anomaly scores of the only-spatial-anomaly trajectories were greater, in fact, greater than the mean anomaly scores of nominal trajectories) was rejected. Therefore, **Hypothesis 2.2 is accepted**.

Figure 5.25 also implies that both-spatial-and-energy-anomaly (B) trajectories experience more "severe" energy anomalies compared with only-energy-anomaly (E) trajectories. Comparing the severity of the energy anomalies detected in the both-spatial-and-energy-anomaly category and only-energy-anomaly category provides insight into whether a distinction should be made between the two categories, which would imply that these two categories of trajectories should be considered independently in the context of decision-making.

Considering Table 5.4 and Figure 5.23 it is evident the statistical properties of the distributions of the anomaly scores of the both-spatial-and-energy-anomaly (B) trajectories and only-energy-anomaly (E) trajectories were distinct. Specifically, the distribution of the anomaly scores for the both-spatial-and-energy-anomaly trajectories was skewed toward much higher anomaly scores. Breaking the distributions down by individual air traffic flows provided similar results. Therefore, **Hypothesis 2.3 is accepted**. Therefore, the underlying mechanisms may be dissimilar for both-spatial-and-energy-anomaly (B) and only-energy-anomaly (E) trajectories. Moreover, it appeared that more severe energy anomalies occurred if an aircraft was observed to experience off-nominal spatial states, which is of interest for ATM system operators to consider to facilitate the safe and efficient management of air traffic in all situations.

However, it is important to note that the proportion of trajectories labeled as spatial anomalies was significantly less than the proportion of trajectories labeled as spatially nominal. Therefore, the range of nominal energy metrics was likely skewed more towards those are experienced by spatially nominal trajectories, which may have resulted in the assessment of energy anomalies being more severe (with respect to anomaly score) for both-spatial-and-energy-anomaly trajectories when compared with only-energy-anomaly trajec-

tories. An investigation into the rate of unable approaches for the both-spatial-and-energy-anomaly and only-energy-anomaly categories could shed more light on the relationship. However, this research is limited by a lack of availability of aircraft weight information in ADS-B data, which supports the full evaluation of stabilized approach criteria.

Considering all the statistical evaluation metrics, Flow 5 produced metrics that were often in the "middle" of the metric ranges. As such, Flow 5 is selected to present a more in-depth analysis. Considering the three anomaly categories (S, E, and B), an individual trajectory was selected belonging to each category. Figure 5.26 displays the spatial profiles of the selected trajectories overlaid on the spatial profiles of all spatially nominal trajectories. Analogously, Figure 5.27 displays the energy profiles (with respect to SPE, SKE, STER) of the selected trajectories overlaid on the bounds of energy-nominal trajectories. The selected only-spatial-anomaly trajectory (in blue) clearly deviated from the nominal trajectories' spatial profiles (in green). Although, despite the readily observable spatial deviation, the energy profiles of this trajectory remained nominal throughout the entire approach. This supports the concept that an aircraft experiencing off-nominal spatial operations does not necessarily experience poor energy management. The complement was true considering the selected only-energy-anomaly trajectory in that the trajectory (in orange) was in near-perfect alignment with the other spatially nominal trajectories, yet experienced large deviations in its energy profiles. Finally, considering the selected both-spatial-and-energy-anomaly trajectory, both its spatial and energy profiles deviated significantly from observed nominal operations. Additionally, evaluation of the trajectory's spatial and energy profiles appeared to indicate the that the deviations seemed more severe than the deviations observed in the only-spatial-anomaly trajectory and only-energy-anomaly trajectory with respect to spatial and energy metrics, respectively.

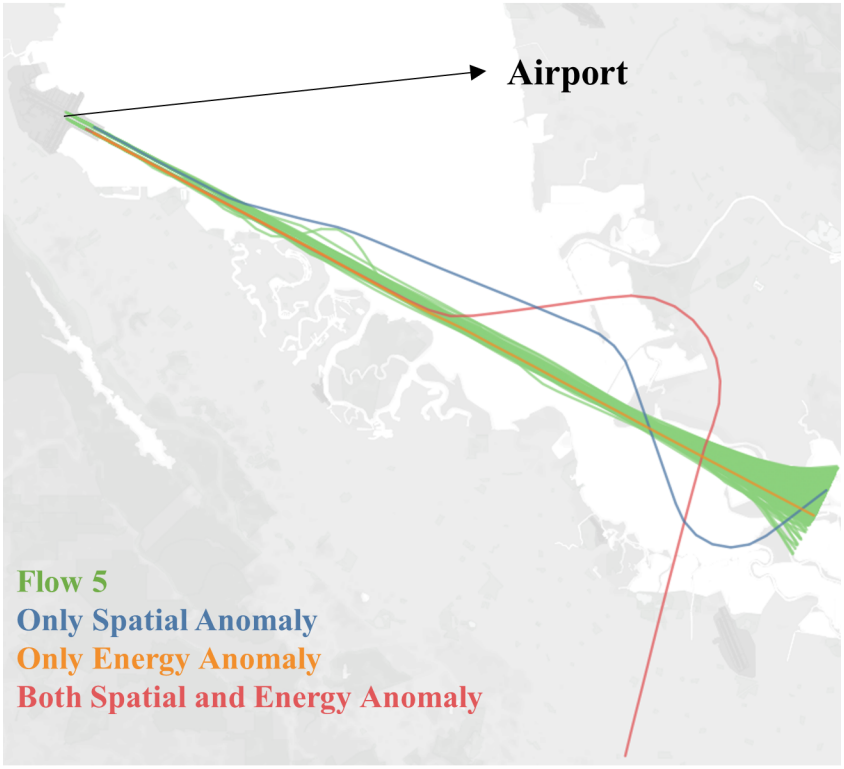

Figure 5.26: Flow 5, Spatial Profiles and Selected Anomalous Trajectories

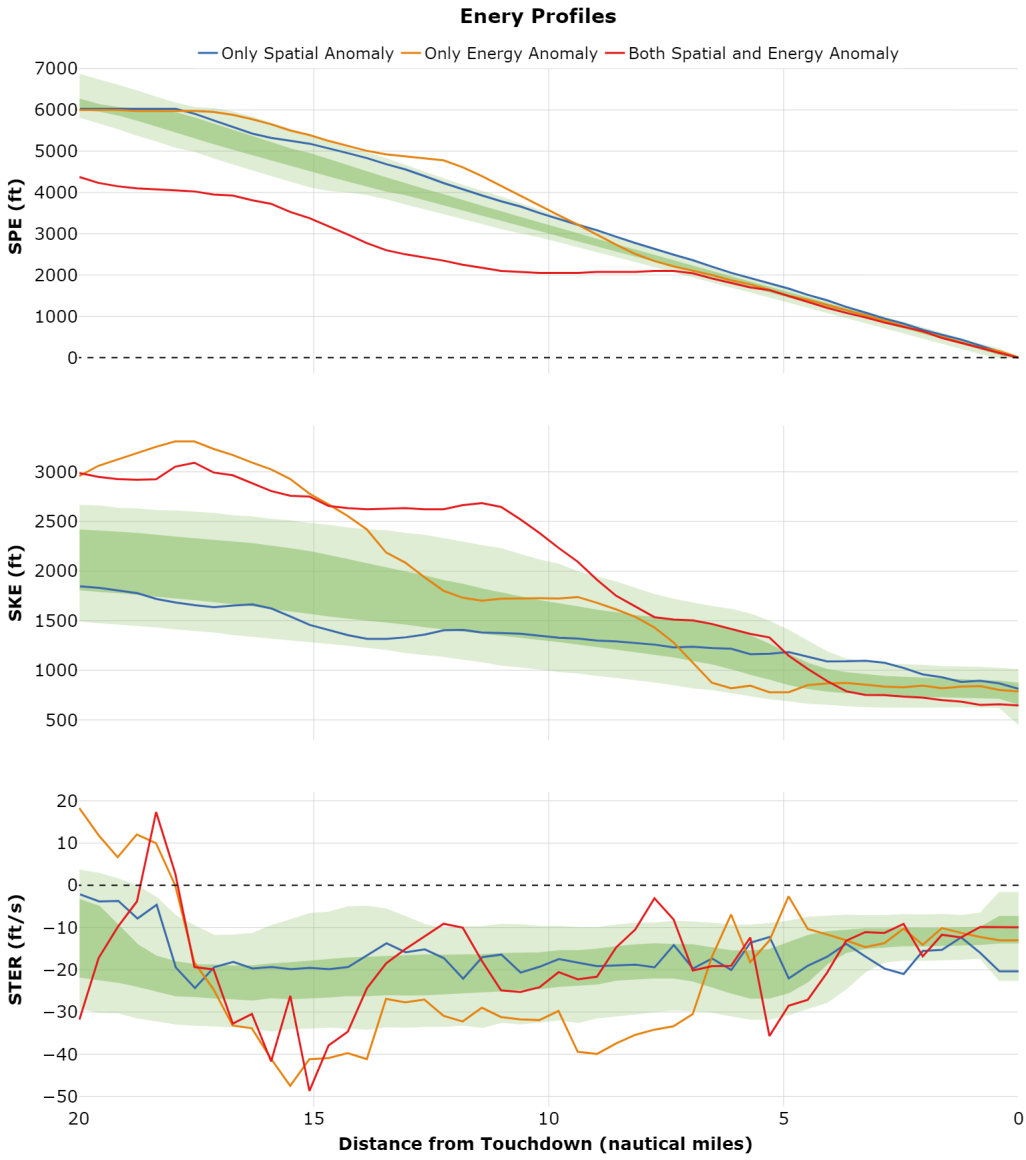

Figure 5.27: Flow 5, Energy Profiles and Selected Anomalous Trajectories

The acceptance of **Hypothesis 2.1**, **Hypothesis 2.2**, and **Hypothesis 2.3** fills **Gap 2**. The primary contribution of filling **Gap 2** is knowledge of the relationship between spatial and energy anomalies for arriving aircraft.

5.4 Airspace-Level Analysis

Executing the experimental approach designed in section 4.3 that is associated with filling **Gap 3** provides results that may be interpreted to enable the assessment of **Hypothesis 3.1 Acceptance** and **Hypothesis 3.2 Acceptance**. Thus, the implementation of the experimental approach is detailed and the corresponding results are presented and discussed.

5.4.1 Implementation

The **Gap 3** experimental approach involves **Experiment 3.1** and **Experiment 3.2** and includes three main steps: (i) data processing, (ii) identification and characterization (**Experiment 3.1**), and (iii) prediction (**Experiment 3.2**). Thus, the implementations of each step of the experimental approach to accomplish the objectives of **Experiment 3.1** and **Experiment 3.2** are described.

Data Processing

The entire airspace-level analysis data set of all arriving aircraft operating within the KSFO terminal airspace in 2019 was leveraged to implement the experimental approach. Therefore, the time period is the full year of 2019. One hour was selected as length of the time intervals to consider within the time period. Thus, an operational state is an aggregation of the flight-level operations, or time series trajectory data, for all arriving aircraft arriving within the KSFO terminal airspace during one hour. One hour was selected due to the tendency of some airspace-level metrics to be computed on an hourly basis. For instance, the FAA's ASPM database contains flight metadata information, such as number of scheduled arrivals, percentage of on-time arrivals, average arrival delay time, etc., on an hourly basis [107]. Therefore, the data set was split into hourly data sets containing all trajectories observed during the hour, resulting in 8,760 time intervals. However, most airports do not operate 24 hours per day. Hence, some hourly data sets do not contain any trajectories

(approximately 5% of the hourly data sets are empty). Therefore, these hourly data sets were discarded, leaving 8,283 time intervals in which the corresponding data set contains at least one trajectory.

Though, it is asserted that a single trajectory record does not contain sufficient information to perform an airspace-level analysis as no aggregation of multiple flight-level operations occurs. For the full year of 2019, all available metrics provided on an hourly basis within the FAA's ASPM database were extracted, where the number of scheduled arrivals was specifically of interest. Figure 5.28 displays a histogram of the number of scheduled arrivals per hour at KSFO for all hours in 2019, excluding hours containing zero scheduled arrivals. The distribution of number of scheduled arrivals per hour is bi-modal. Specifically, there is a peak in the distribution of number of scheduled arrivals per hour at two to four scheduled arrivals per hour as well as around 29 to 31 scheduled arrivals per hour. Excluding number of scheduled arrivals per hour below around 15, the distribution of number of scheduled arrivals per hour appeared relatively normal.

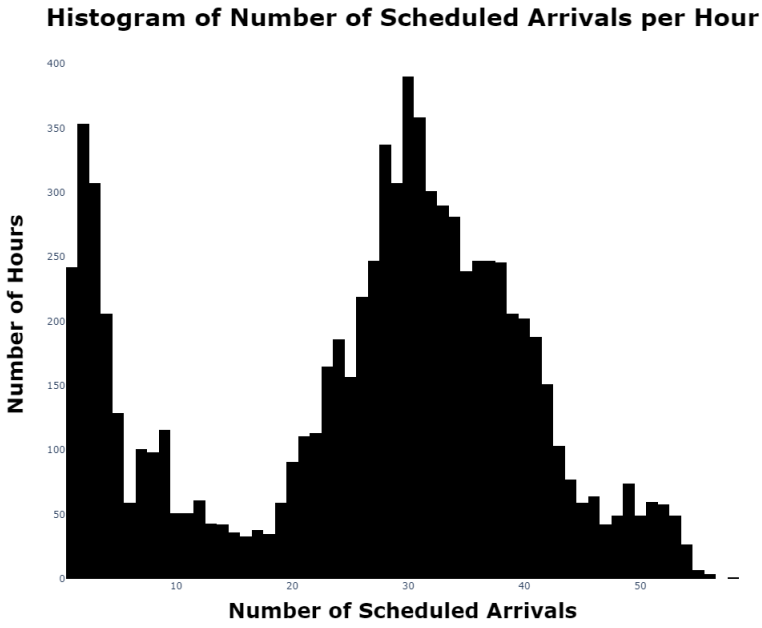

Figure 5.28: Histogram of Number of Scheduled Arrivals per Hour at San Francisco International Airport for All Hours in 2019, Excluding Hours Containing Zero Scheduled Arrivals

It is important to consider the potential outcomes of generating an airspace density matrix for an hour in which a relatively small number of aircraft arrive at an airport. For instance, considering a relatively small number of arriving aircraft trajectories, it is difficult to discern the *distribution* of the trajectories as the sample size is so small. Further, the impact of spatially anomalous trajectories on the representation of the terminal airspace operational states may be much more pronounced if only a few trajectories are considered. For instance, if only three trajectories are contained within an hourly data set, and one is spatially anomalous, that spatially anomalous trajectory's path within the terminal airspace will have a density of one-third, which is a density that should more realistically be expected of a true air traffic flow. Consequently, the generation of several airspace density matrices containing a relatively small number of flights may skew results of the operational pattern identification and/or operational state characterization. It is undesirable to have a single trajectory significantly impact the representation of the terminal operational states. Recall the limitations of current approaches discussed in section 2.3, which include specifying an air traffic flow as being active if only one trajectory is observed to be associated with that air traffic flow. Thus, it is asserted that a reasonable sample size of trajectories must be aggregated to adequately represent the distribution of arriving aircraft trajectories to enable proper identification of operational patterns and characterization of operational states. Moreover, in the context of the ATM system modernization goal of increasing capacity, the operational patterns of time intervals in which a small number of aircraft arrive are more relevant than those with fewer arrivals.

Therefore, a threshold was set below which time intervals were discarded if the time interval data set did not contain at least this many trajectories. While this threshold may be variable depending upon the configuration of and operations within a specific terminal airspace, for the KSFO terminal airspace this threshold was set at 15. A threshold of 15 was selected considering Figure 5.28, where 15 was approximately the mid-point between the two modes of the distribution of the number of scheduled arrivals per hour. After discarding

all hourly data sets that did not contain at least 15 trajectories 5,861 data sets remained for analysis.

To generate an $N \times N$ airspace density matrix requires the specification of N, which is dependent upon the *resolution* of the airspace density matrix desired. The resolution should enable corresponding trajectory points along a similar, yet not exactly the same, spatial path to be grouped in the same grid spaces, while placing those that are divergent in different spaces. Too high a resolution may result in all corresponding trajectory point being placed in a separate grid space, while too low a resolution may result in all corresponding trajectory points being placed in a single grid space. A sensitivity study was performed to determine an appropriate value for N. The sensitivity study involved generating an airspace density matrix considering the trajectories within all 5,861 hourly data sets for different values of N. The airspace density matrix was plotted as a heatmap and the value of N for which the "spread" of the trajectories appeared to be best represented was selected. A value of N equal to 50 was selected, where the heatmap of the airspace density matrix values for N equal to 50 is displayed in Figure 5.29.

Figure 5.29: Airspace Density Matrix Generated Considering Trajectories within All 5,861 Hourly Data Sets

It is evident the dominant air traffic flows identified in the **Anomaly Detection** step and plotted in Figure 5.20 (corresponding with the official SFO west plan) are visible within Figure 5.29 with the addition of at least one air traffic flow that may be associated with the official SFO south east plan.

Identification and Characterization

The generated airspace density matrices provided the basis for the input into the recursive DBSCAN procedure. The airspace density matrices were flattened, resulting in vectors of length 2,500 ($50 * 50 = 2500$). The vectors generated as a result of the flattening of the airspace density matrix were subsequently stacked to generate a feature vector matrix of size 5861×2500. Additionally, following the steps of the experimental approach outlined in section 4.3, the dimensionality of the feature vector matrix was reduced by applying the

UMAP dimensionality reduction technique. UMAP was implemented leveraging the *umap* [145] Python library such that two components were returned, i.e. the dimension of the feature vectors was reduced from 2,500 to 2. A plot of each of the operational states in the reduced-dimensional space is presented in Figure 5.30. Figure 5.30 appears to indicate three very distinct clusters, or operational patterns.

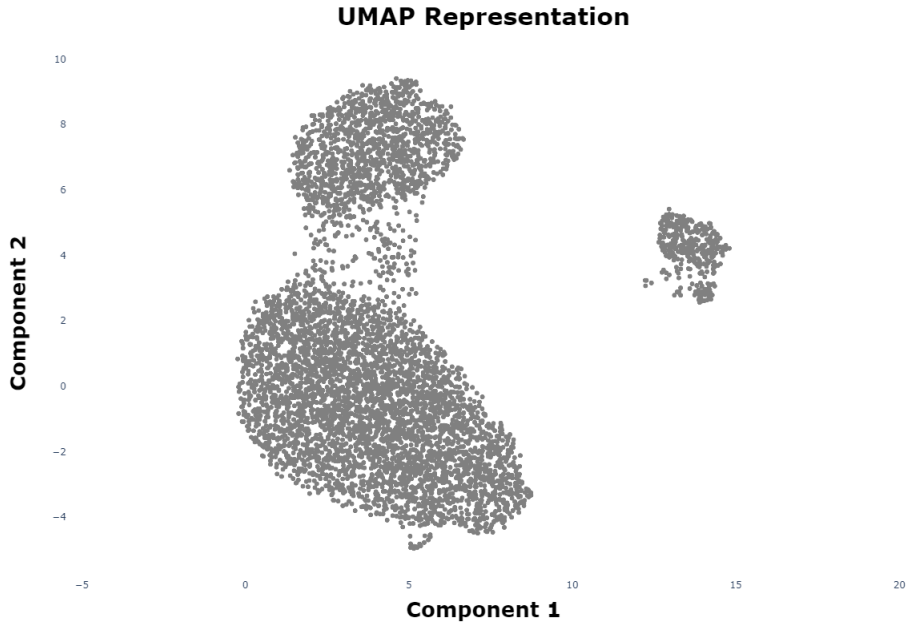

Figure 5.30: Operational States After UMAP Dimension Reduction

Before implementing DBSCAN, the UMAP components were normalized leveraging the *sklearn* Python library's z-normalization module. Then, the recursive DBSCAN procedure was implemented leveraging the *sklearn* Python library's DBSCAN module. The $MinPts$ parameter was set at 2% of the total number of time intervals. The value of 2% of the total number of time intervals was selected considering that the official SFO west plan is observed 98-95% of the time, which would leave, potentially, approximately 2% of the time that the official SFO south east plan may occur. Therefore, setting a value of $MinPts$ greater than 2% of the total number of time intervals could result in operational states corresponding to operational patterns associated with the official SFO south east plan being identified as being outliers by DBSCAN.

As introduced, there exists monotonic relationship between fraction of outliers identified and the ε parameter for a fixed value of $MinPts$. Therefore, the value of ε was varied until the fraction of outliers identified was between 0.01 and 0.02 for each recursion step in the recursive DBSCAN procedure. The fraction of outliers identified was specified as between 0.01 and 0.02 due to the property of the outlying operational states being characterized as anomalous in the final recursion step, after no more transitional operational states are characterized. Anomalies are, inherently, rare events. Thus, it is desirable to detect only a small portion of a data set as anomalies, especially when there exists no ground truth or prior work related to detection of airspace-level anomalous operational states. An objective in detecting anomalous operational states is to narrow down the set of time intervals for SMEs, operators, planners, and decision-makers to place the most emphasis on understanding due to their abnormal characteristics. Accordingly, it is desirable to begin with a smaller set of anomalous operational states.

An outlying operational state may only be characterized as transitional if the time intervals directly before and after the outlying time interval are available in the data set. For instance, it was possible, due to the threshold requiring least 15 trajectories to exist within an hourly data set, that the time intervals directly before or after the outlying time interval were not available. In the reduced-dimensional space, the results of the implementation of the recursive DBSCAN procedure in terms of the identified operational patterns and characterized operational states are displayed in Figure 5.31.

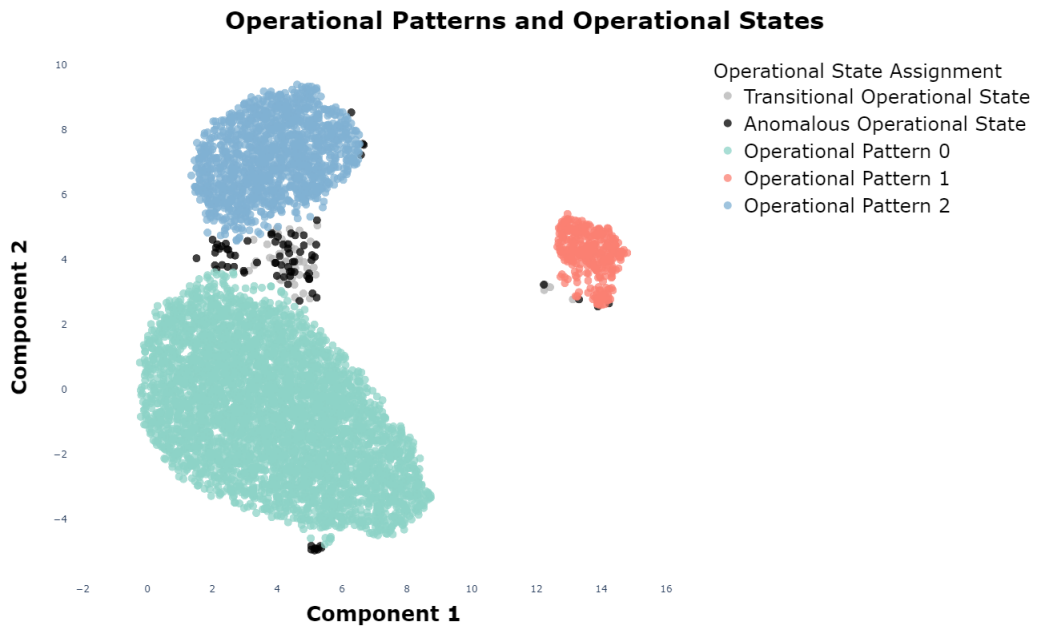

Figure 5.31: Identified Operational Patterns and Characterized Operational States in UMAP Reduced-Dimension Space

Table 5.6 presents the number of time intervals (hours) assigned to each class of operational state: transitional, anomalous, or specific operational pattern if the operational state was classified as nominal. Additionally, the percentage of the total number of time intervals (5,861 hours) assigned to each operational state is displayed in Table 5.6.

Table 5.6: Operational State Characterization Including Specific Operational Pattern Identified If Operational State Is Nominal

Operational State Assignment	**Number of Hours**	**Percentage of Hours**
Transitional Operational State	45	0.77%
Anomalous Operational State	76	1.30
Operational Pattern 0	4,086	69.72%
Operational Pattern 1	353	6.02%
Operational Pattern 2	1,301	22.20%

Prediction

Before the training of any of the selected multi-class classification algorithms, the input features were specified and derived. The features were able to be derived from the extracted ASOS data. The ASOS data was aggregated by hour such that the means of each metric over the hour were computed if the ASOS data for the hour contained multiple measurements. As a prediction capability is not especially valuable if the prediction is not able to occur in advance of the prediction time interval of interest, ASOS data measurements associated with the prediction time interval of interest were not leveraged.

The features used to train the prediction models included the ASOS data aggregated by hour for the hour prior to the prediction time interval of interest. Additionally, derived features from the recorded weather measurements included the mean of the ASOS data aggregated by hour for the two hour prior to the prediction time interval of interest. The ASOS metrics considered included: $alti$, $drct$, $dwpf$, $gust$, $mslp$, $relh$, $sknt$, $tmpf$, $vsby$. Finally, categorical features indicating the month ($month$) and hour of day ($hour$) of the time interval of interest were included in the input feature vector. Therefore, 20 total features were included in the input feature vector matrix used to train the prediction models.

The set of multi-class classification algorithms trained included an artificial neural network, a gradient-boosted decision tree (XGBoost), and an SVM, as specified in section 4.3. The artificial neural network and SVM were trained leveraging the *sklearn* Python library modules, whereas the gradient-boosted decision tree was trained leveraging the *xgboost* Python library [132]. The primary classification performance metric of interest was accuracy, i.e. the percentage of correctly classified samples. However, other metrics such as precision, recall, and F1-score were also computed as assessing at all of these measures together may enable a more complete evaluation of the performance of the prediction models, especially considering that the classes were not completely balanced.

The data set of the labels (assigned operational pattern, if operational state is nominal) for the time intervals and their corresponding input features were split into training and

testing data sets to assess the various algorithms' classification performance metrics. The training data set contained 80% of the time intervals assigned to an operational pattern (4,592 time intervals), whereas the testing data set contained the remaining 20% of the time intervals assigned to an operational pattern (1,148 time intervals). A breakdown of the number of time intervals associated with each operational pattern for the training and testing data sets is displayed in Table 5.7.

Table 5.7: Operational Pattern Breakdown in Training and Testing Data Sets

Operational Pattern	**Training Samples**	**Testing Samples**
Operational Pattern 0	3,264	822
Operational Pattern 1	283	70
Operational Pattern 2	1,045	256

5.4.2 Results and Discussion

The results are presented and discussed separately for each step in the experimental approach associated with filling **Gap 3** related to the **Airspace-Level Analysis** step in the proposed methodology.

Data Processing

For each identified operational pattern, time intervals observed to have differing trajectory densities were plotted. Both a plot of the arriving aircraft trajectories within the terminal airspace and a heatmap of the airspace density matrix values are displayed side-by-side such that it was evident how the trajectories were aggregated to generate an airspace density matrix. Airspace density matrices corresponding to operational states associated with Operational Pattern 0, Operational Pattern 1, and Operational Pattern 2 are presented and discussed.

Figure 5.32 and Figure 5.33 both display a plot of the arriving aircraft trajectories operating during a specified time interval and a corresponding heatmap of airspace density matrix values for time intervals associated with the identified Operational Pattern 0. The same few air traffic flows would have been considered "active" if existing methods for airspace-level analysis were applied, as there exists at least one trajectory corresponding to the paths that would likely be identified as air traffic flows. However, the airspace density matrix was able to make a distinction between the density of the different air traffic flows. For instance, the highest density of trajectories displayed within Figure 5.32 appeared to be those trajectories approaching from the south, whereas the highest density of trajectories displayed within Figure 5.33 appeared to be those approaching from the south east. These density differences were accounted for in the generation of the airspace density matrix by the values of the grid spaces corresponding to those trajectories approaching from the south in Figure 5.32 being highest and by the values of the grid spaces corresponding to those trajectories approaching from the south east in Figure 5.33 being lower.

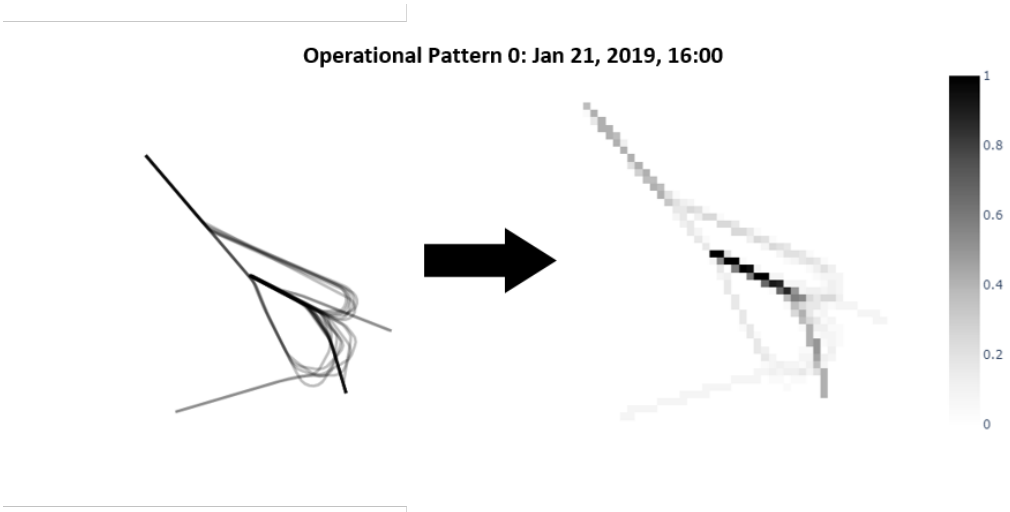

Figure 5.32: Operational Pattern 0, Jan 21st, 2019, 16:00

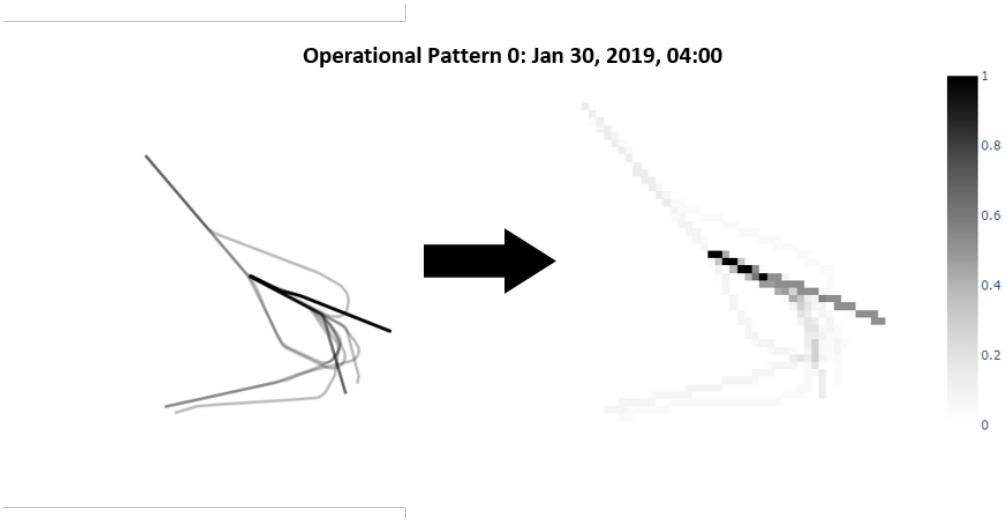

Figure 5.33: Operational Pattern 0, Jan 30^{th}, 2019, 04:00

Figure 5.34 and Figure 5.35 both display a plot of the arriving aircraft trajectories operating during a specified time interval and a corresponding heatmap of airspace density matrix values for time intervals associated with the identified Operational Pattern 1. Again, it is noted that the same couple of air traffic flow would have been considered "active" if existing methods for airspace-level analysis were employed. The primary difference between the two time intervals displayed in Figure 5.34 and Figure 5.35 is that Figure 5.34 displays a plot in which the density of the trajectories approaching from the south/west was higher than the density of trajectories approaching from the south/west in Figure 5.35. The trajectories approaching from the south/west displayed in Figure 5.35 were much more "spread out" than those displayed in Figure 5.34. The differing "spreads", or densities, of the trajectories approaching from the south/west was captured in the generation of the airspace density matrix as in Figure 5.35 many more grid spaces appeared to have non-zero values, while in Figure 5.34 the values in the south/west grid spaces were higher when they were non-zero, yet many more values in these grid spaces were equal to zero.

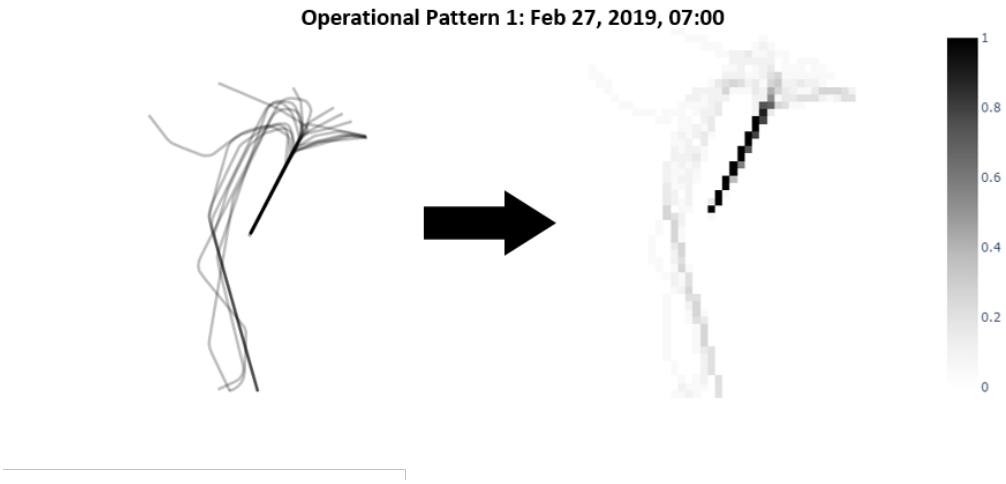

Figure 5.34: Operational Pattern 1, Feb 27^{th}, 2019, 07:00

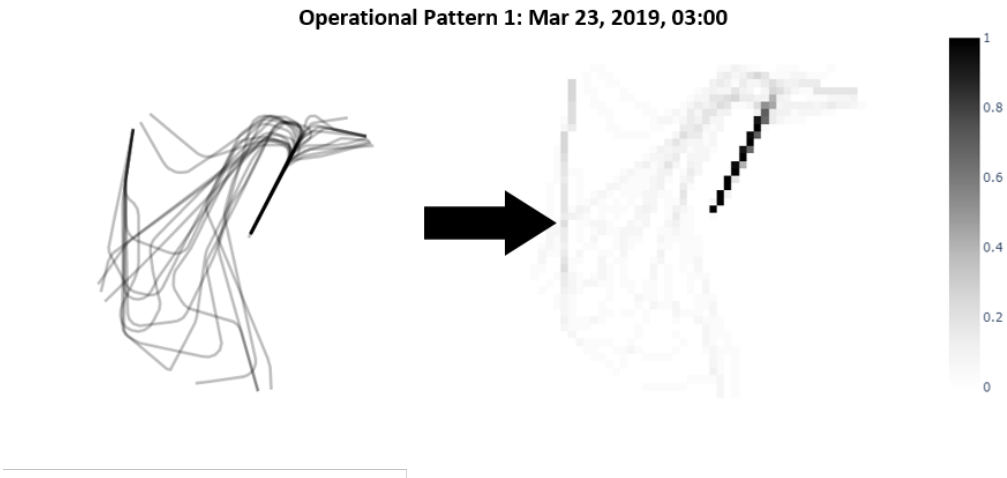

Figure 5.35: Operational Pattern 1, Mar 23^{rd}, 2019, 03:00

Finally, Figure 5.36 and Figure 5.37 both display a plot of the arriving aircraft trajectories operating during a specified time interval and a corresponding heatmap of airspace density matrix values for time intervals associated with the identified Operational Pattern 2. Considering Figure 5.36 and Figure 5.37, there were distinct differences in the arriving aircraft trajectories observed. Specifically, a go-around appeared to occur within the time interval corresponding to Figure 5.37. Further, the trajectories displayed in Figure 5.37 were much more scattered than the relatively "tight" nature of the trajectories displayed

in Figure 5.36. This difference in densities of the trajectories was evident in the airspace density matrix values, as the heatmap displayed in Figure 5.36 indicates that there are several more spaces within the terminal airspace that were equal to zero than in the heatmap displayed in Figure 5.37.

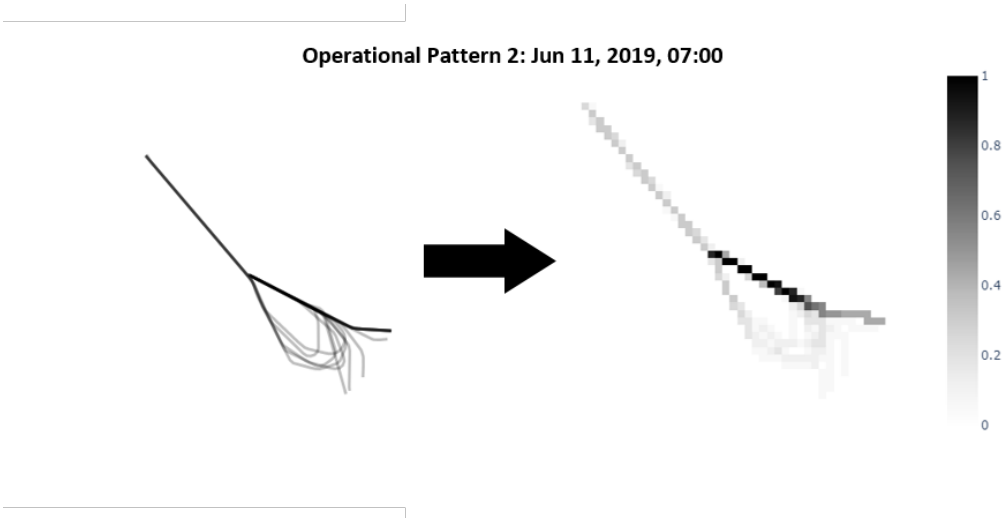

Figure 5.36: Operational Pattern 2, Jun 11^{th}, 2019, 07:00

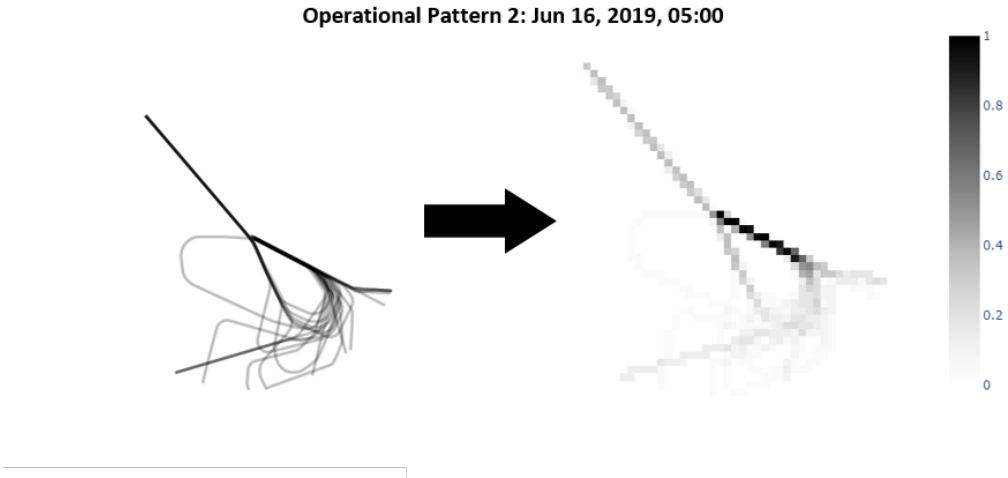

Figure 5.37: Operational Pattern 2, Jun 16^{th}, 2019, 05:00

Identification and Characterization

The three identified operational patterns were summarized by leveraging DBSCAN with the WED to identify air traffic flow centroids considering all individual trajectories associated with time intervals that had been identified to belong to the operational pattern of interest. Figure 5.38, Figure 5.39, and Figure 5.40 display the summaries of the identified Operational Pattern 0, Operational Pattern 1, and Operational Pattern 2, respectively.

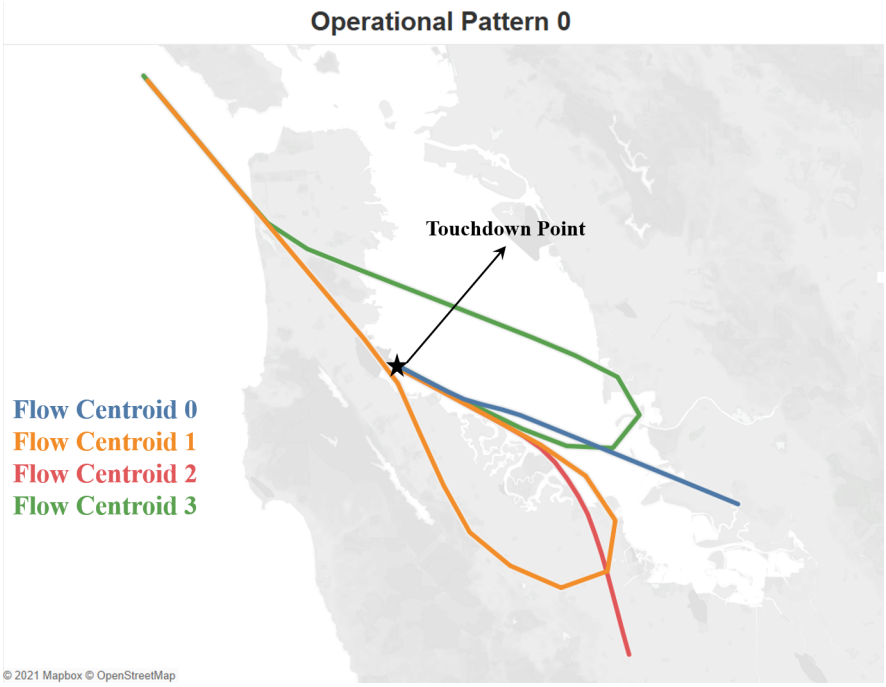

Figure 5.38: Operational Pattern 0, Identified Air Traffic Flow Centroids

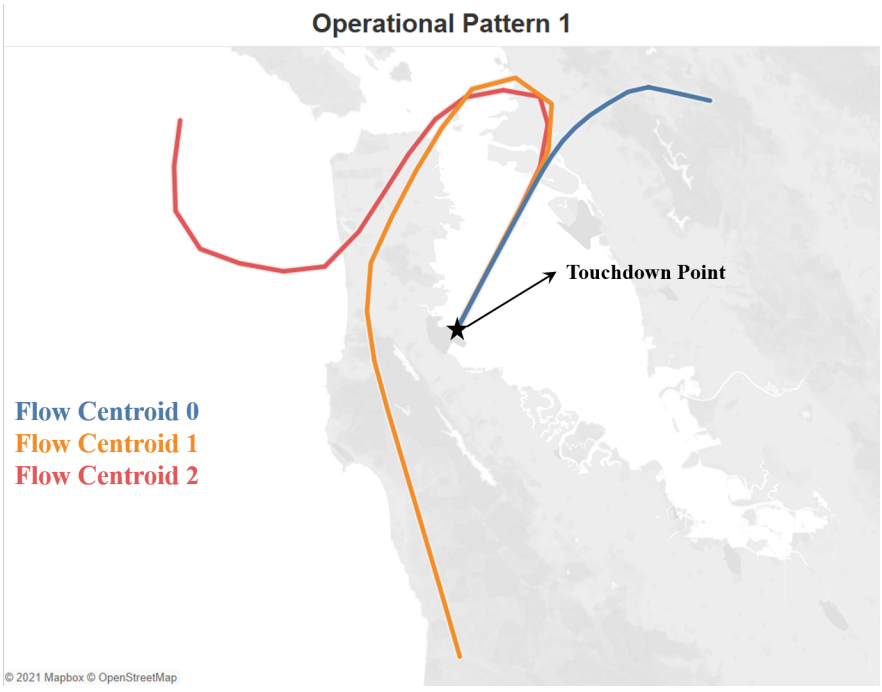

Figure 5.39: Operational Pattern 1, Identified Air Traffic Flow Centroids

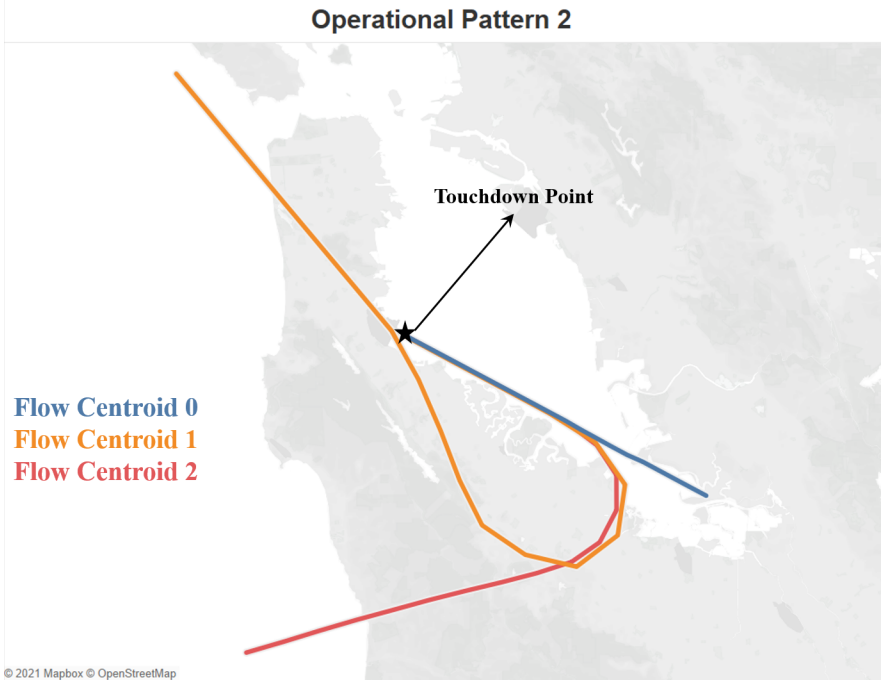

Figure 5.40: Operational Pattern 2, Identified Air Traffic Flow Centroids

Operational Pattern 0 and Operational Pattern 2 appeared to conform with the official SFO west plan, while Operational Pattern 1 conformed with the official SFO south east

plan. While Operational Pattern 0 and Operational Pattern 2 both conformed with the official SFO west plan, there existed important distinctions between the two operational patterns. For instance, while both Operational Pattern 0 and Operational Pattern 2 contained a prominent air traffic flow approaching from the south east, the air traffic flow centroids are distinct for Operational Pattern 0 and Operational Pattern 2. Figure 5.38 displays an air traffic flow approaching from the south east (Flow Centroid 0 in Figure 5.38) that was slightly "above" the air traffic flow approaching from the south east in Figure 5.40 (Flow Centroid 0 in Figure 5.40), where the Operational Pattern 0 air traffic flow approaching from the south east made a slight adjustment towards the end of the approach to align with the runway rather than being aligned throughout the approach as in Operational Pattern 2.

Further, Figure 5.38 for Operational Pattern 0 displays two air traffic flows (Flow Centroid 1 and Flow Centroid 3) that originated from the northwest, whereas Figure 5.40 displays only one air traffic flow (Flow Centroid 1) that originated from the northwest. Moreover, the loop present for Flow Centroid 1 in Figure 5.38 was "tighter" than the loop present for Flow Centroid 1 in Figure 5.40. Finally, Operational Pattern 0 and Operational Pattern 2 had two distinct remaining prominent air traffic flows observed. Flow Centroid 2 in Figure 5.38 approached from the south, whereas Flow Centroid 2 in Figure 5.40 approached from the southwest. Thus, despite the conformance of Operational Pattern 0 and Operational Pattern 2 to the official SFO west plan, there existed important differences revealed in the structure of the airspace in the two operational patterns, which the recursive DBSCAN procedure was able to uncover. Additionally, this recursive DBSCAN procedure was able to separate and identify time intervals corresponding to Operational Pattern 1, which conforms with the official SFO south east plan. Therefore, distinct operational patterns were identified that align with the arriving aircraft official SFO west plan and official SFO south east plan.

Moreover, assessing the relative positions of the clusters within Figure 5.31, Operational Pattern 0 and Operational Pattern 2 were "closer", and appeared to have more overlap

and transition periods between them. Considering both Operational Pattern 0 and Operational Pattern 2 both conformed to the official SFO west plan, it is reasonable to assert that these operational patterns were, actually, "closer", whereas Operational Pattern 1, which conformed to the official SFO south east plan and had a significantly different structure was much "further" as displayed in Figure 5.31. Additionally, there was at least one air traffic flow that was relatively similar in position between the two operational patterns. Therefore, it was anticipated that the higher-dimension airspace density matrix representation of the operational states would display more global similarities between Operational Pattern 0 and Operational Pattern 2, which the UMAP dimensionality reduction technique preserved, where the global structure of the data set being preserved in the reduced-dimensional space is a specific benefit of leveraging the UMAP dimensionality reduction technique [131].

Additionally, the recursive DBSCAN procedure sought to characterize operational states that were not associated with a cluster (operational pattern) as transitional or anomalous. With respect to operational states being characterized as transitional, Figure 5.41 and Figure 5.42 display the sequence of operational states surrounding an operational state characterized as transitional for a transition between Operational Pattern 0 to Operational Pattern 1 and Operational Pattern 2 to Operational Pattern 0, respectively.

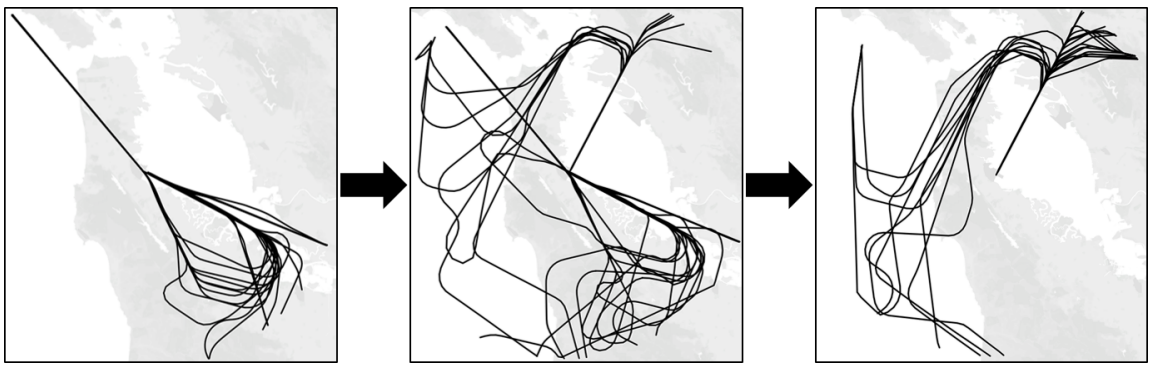

Figure 5.41: Operational Pattern 0 to Operational Pattern 1 Sequence, Mar 25^{th}, 2019, 16:00, 17:00, 18:00

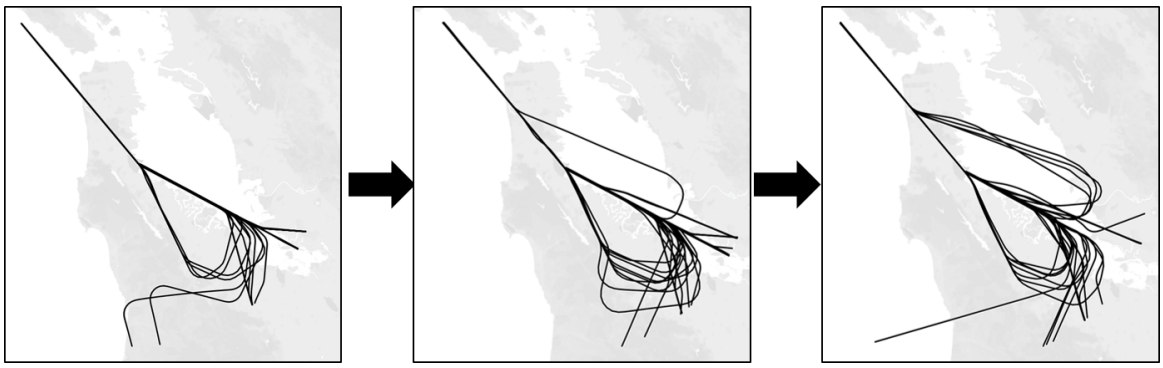

Figure 5.42: Operational Pattern 2 to Operational Pattern 0 Sequence, Jan 30^{th}, 2019, 19:00, 20:00, 21:00

Figure 5.41 displays a transition between Operational Pattern 0 and Operational Pattern 1. The operational state was observably characterized as having been transitional due to the existence of trajectories that align with both the prominent air traffic flows associated with Operational Pattern 0 and Operational Pattern 1. Further, it appeared the arriving aircraft operations were experiencing a transition from the official SFO west plan to the official SFO south east plan. Figure 5.42 displays a transition between Operational Pattern 2 and Operational Pattern 0. The time interval of the operational state characterized as having been transitional contains trajectories that align with the most prominent air traffic flows associated with Operational Pattern 0 and for Operational Pattern 2. Both Operational Pattern 0 and Operational Pattern 2 conformed to the official SFO west plan. However, if the official SFO west plan and official SFO south east plan were considered to be the only operational patterns, the transitional nature of the terminal airspace from two distinct operational structures (operational patterns) would have been neglected. Overall, the recursive DBSCAN procedure appeared to be capable of characterizing obvious transitional operational states as such.

Finally, with respect to operational states having been characterized as anomalous, Figure 5.43 and Figure 5.44 display plots of two operational states that were characterized as anomalous by the recursive DBSCAN procedure. Both of the operational states displayed in Figure 5.43 and Figure 5.44 observably contain one or more trajectories that do not align

with the identified prominent air traffic flows associated with the identified operational patterns and/or were potentially experiencing either a go-around or holding pattern. Further, the time interval displayed in Figure 5.43 contained no trajectories approaching from the northwest, which was abnormal, and likely contributing to the time interval's operational state representation having been detected as anomalous. One trajectory alone experiencing a go-around or holding pattern should not result in an operational state being characterized as anomalous. However, in Figure 5.44 it appears there may have been more than one or two trajectories that experienced go-around or holding patterns; thus, resulting in the operational state having been detected as anomalous. Overall, the recursive DBSCAN procedure appeared to be capable of characterizing obvious anomalous operational states as such in that multiple trajectories observed within the anomalous operational state time interval appeared to be abnormal.

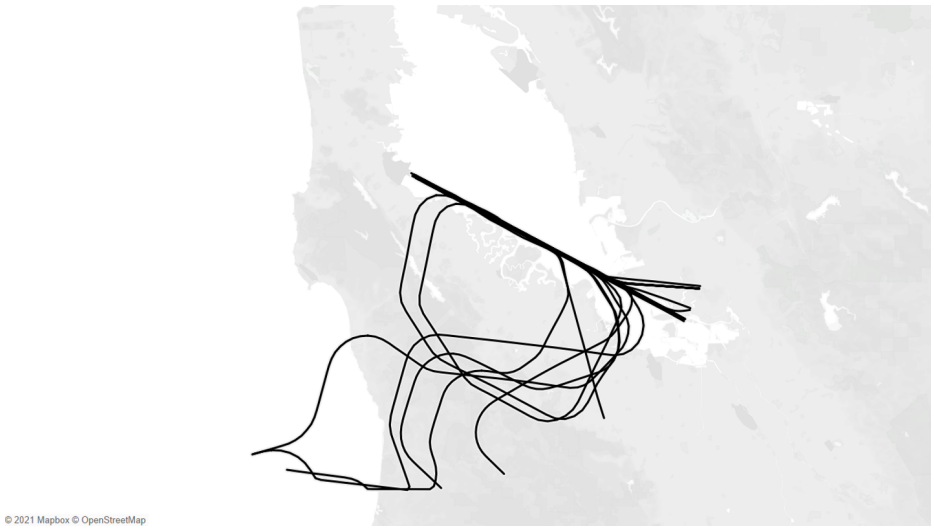

Figure 5.43: Anomalous Operational State, May 16^{th}, 2019, 06:00

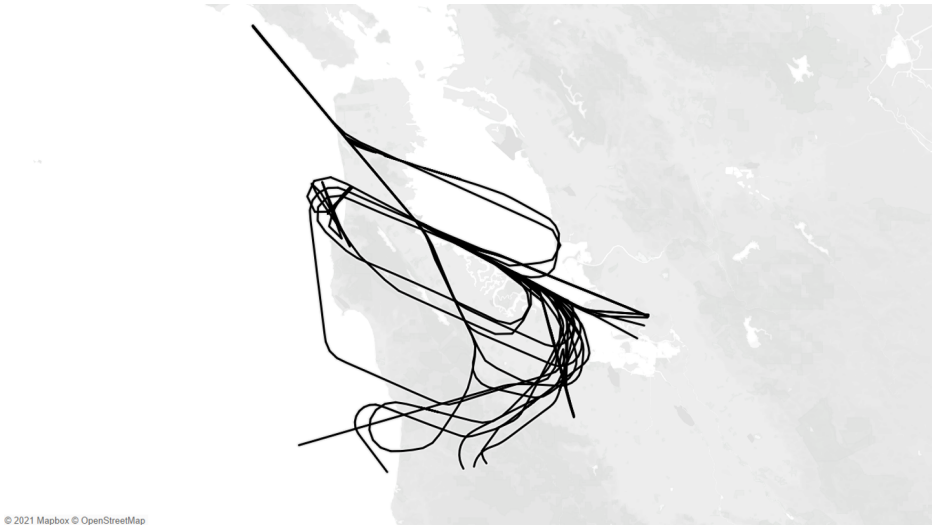

Figure 5.44: Transition Operational State, August 30^{th}, 2019, 15:00

An evaluation of the summaries of the operational patterns through their associated air traffic flow centroids indicated that the operational patterns identified are both distinct and align with the distinct official SFO operational plans. Additionally, Figure 5.41 and Figure 5.42 that corresponded to two differing operational patterns as transitional. Finally, Figure 5.43 and Figure 5.44 indicate the robust characterization of operational states containing multiple trajectory operations that corresponded to off-nominal spatial behavior as anomalous. Therefore, **Hypothesis 3.1 is accepted**.

A time series of the either the operational pattern identified or operational state characterization, if it does not belong to an identified operational pattern, is displayed in Figure 5.45. This time series displays each hour within a day on the Y-axis and each day within the year on the X-axis. It is evident that Operational Pattern 2 tended to dominate for early morning flights and those later in the evening. If flights were occurring very late into the night/early morning (00:00 to 02:00), Operational Pattern 0 was generally not identified. Thus, despite all the trajectory patterns in the official SFO nighttime operations being a replicate of the trajectory patterns in the official SFO west plan, it actually did not appear that during the night time that the full official SFO nighttime operations were ever observed.

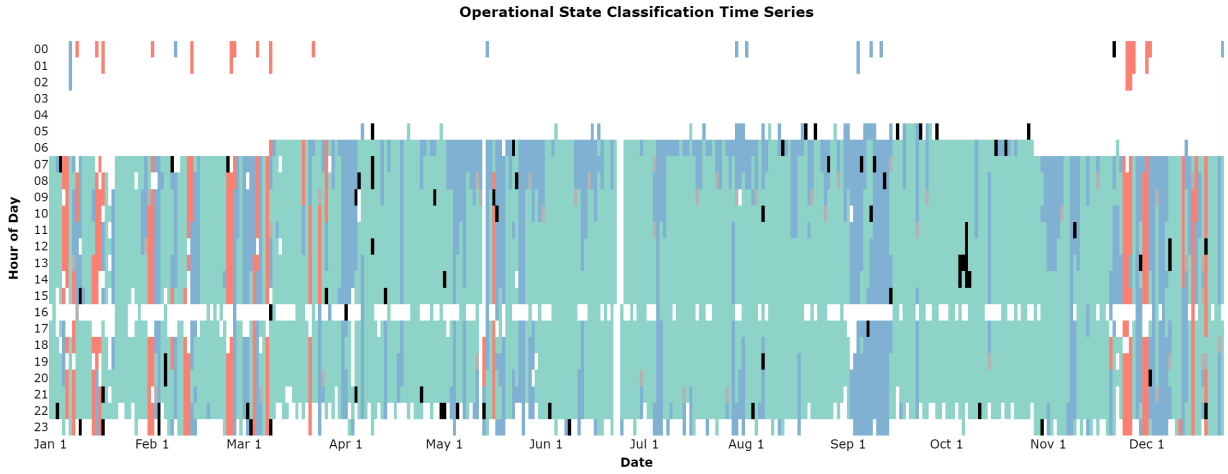

Figure 5.45: Full Year of 2019 Time Series of Operational State Classification By Hour of Day and Day of Year

The typical distribution of operational patterns that were identified by hour of the day within the KSFO terminal airspace is presented in Figure 5.46. As observed in Figure 5.45, Operational Pattern 0 was most commonly observed during the day, while Operational Pattern 1 and Operational Pattern 2 appeared to have been more prominent in the night time/early morning.

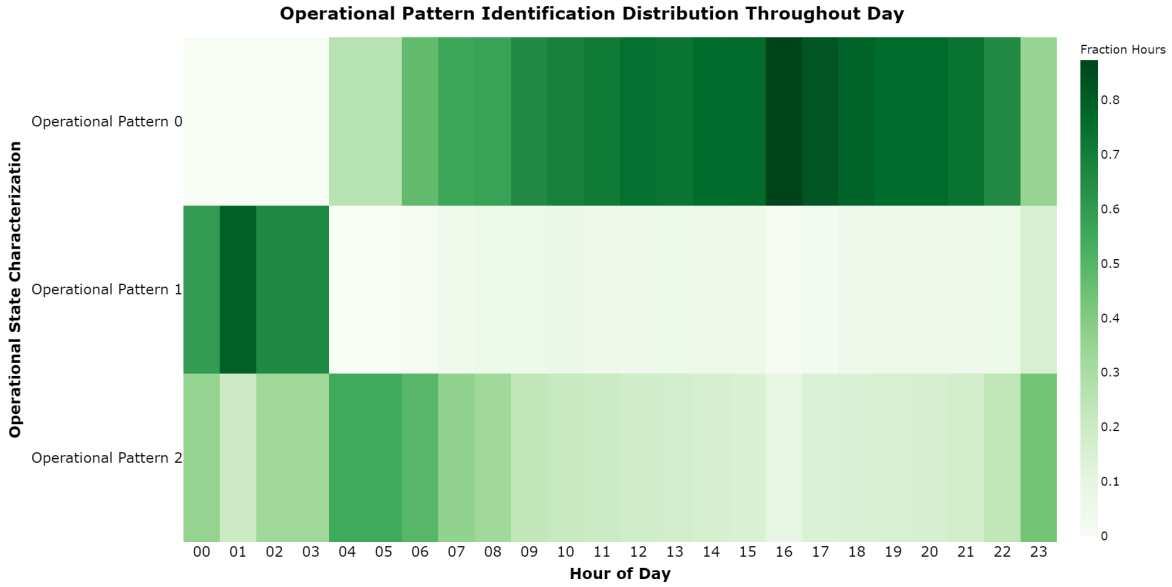

Figure 5.46: Frequency of Operational Patterns Identified by Hour of Day

Further, the distribution of operational states characterized as either anomalous or transitional by hour of the day within the KSFO terminal airspace is presented in Figure 5.47. By definition, transitional and anomalous operational states occurred much less frequently than nominal operational states that were assigned to an identified operational pattern. However, out of the few time intervals that existed corresponding to 04:00 and 05:00, a relatively larger portion of those operational states were identified as being anomalous. This may have been due to extenuating circumstances that lead aircraft to have had to arrive at the airport within those hours. Further, the highest portion of time intervals experiencing transitional operational states occurred around 08:00 and 09:00, which may indicate that a structural change in the terminal airspace somewhat-regularly occurred during one of those hours.

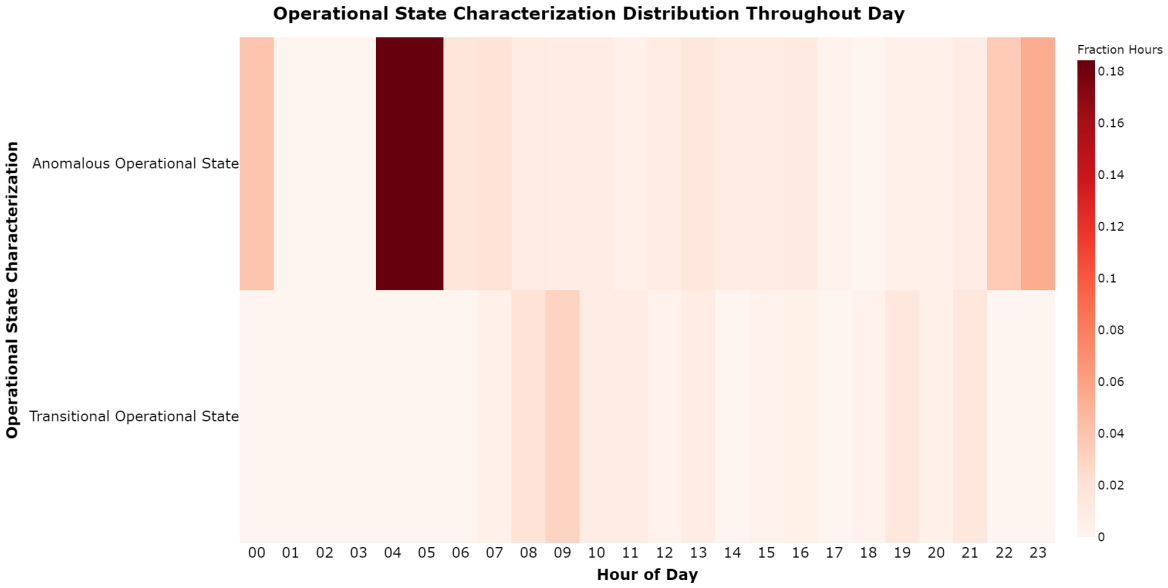

Figure 5.47: Frequency of Operational State Characterized by Hour of Day, If Not Characterized as Nominal

Moreover, Table 5.8 displays the number and percentage of time intervals transitioning between different operational patterns. There were no transitional operational states that appeared to transition from Operational Pattern 1 to either Operational Pattern 0 nor Operational Pattern 2. However, there were five time intervals in which Operational Pat-

tern 0 or Operational Pattern 2 transitioned to Operational Pattern 1. Hence, a transition to Operational Pattern 1 comprised 11.11% of the transitional operational states characterized, while the remaining 88.89% of the transitional operational states characterized were between Operational Pattern 0 and Operational Pattern 2 and vice-versa.

Table 5.8: Number and Percentage of the Time Intervals Transitioning Between Different Operational Pattern

Time Interval Before	**Time Interval After**	**# Intervals**	**% Intervals**
Operational Pattern 0	Operational Pattern 1	4	8.89%
Operational Pattern 0	Operational Pattern 2	14	31.11%
Operational Pattern 1	Operational Pattern 0	0	0.00%
Operational Pattern 1	Operational Pattern 2	0	0.00%
Operational Pattern 2	Operational Pattern 0	26	57.78%
Operational Pattern 2	Operational Pattern 1	1	2.22%

To further expand upon the acceptance of **Hypothesis 3.1**, the output of the experimental approach associated with filling **Gap 2**, related to the **Anomaly Detection** step, was considered. The execution of the **Gap 2** experimental approach resulted in a total of five labels having been assigned to each trajectory. The data processing step resulted in a label having been assigned to a trajectory indicating whether a go-around was performed. Related to the spatial anomaly detection step, the second label indicated whether the trajectory was a spatial anomaly, or not and the third label specified the air traffic flow to which the trajectory was assigned. The fourth label was related to the energy anomaly detection step and indicated whether the trajectory was an energy anomaly, or not. Finally, the fifth label was the label indicating the category to which the trajectory belonged (N, S, E, or B). Each operational state classification (transitional, anomalous, Operational Pattern 0, Operational Pattern 1, and Operational Pattern 2) was considered such that an additional label was added to each trajectory to indicate the operational state classification for the time interval

in which the trajectory occurred.

For each operational state classification, the trajectories associated with the particular operational state classification were divided into sets based upon the air traffic flow to which the trajectory was assigned. The percentage of trajectories belonging to each air traffic flow within the operational classification sets of trajectories was computed, where Table 5.9 displays these percentages.

Table 5.9: Percentage of Trajectories Belonging to Each Air Traffic Flow Considering the Operational State Class to Which the Time Interval in Which a Trajectory Operates Belongs

Air Traffic Flow	**Transitional**	**Anomalous**	**Pattern 0**	**Pattern 1**	**Pattern 2**
0	16.69%	23.89%	19.45%	0.00%	23.11%
1	3.51%	2.78%	9.54%	0.00%	3.26%
2	11.02%	11.39%	6.14%	0.00%	10.93%
3	20.70%	19.72%	25.49%	0.00%	20.38%
4	21.70%	15.56%	37.08%	0.00%	1.82%
5	11.85%	6.11%	1.02%	0.00%	18.50%
6	14.02%	18.61%	0.69%	0.00%	21.60%

During the time period of June 1^{st}, 2019 through September 30^{th}, 2019, Operational Pattern 1 did not occur, which was also evident when assessing Figure 5.45. Evaluating Figure 5.20, it was possible to associate the air traffic flows identified in the execution of the spatial anomaly detection step with the air traffic flows identified to summarize the operational patterns identified. Considering the Operational Pattern 0 air traffic flow centroids summary in Figure 5.38, the associations were as follows:

- Figure 5.38 Flow Centroid 0 was associated Figure 5.20 Flow 4.
- Figure 5.38 Flow Centroid 1 was associated Figure 5.20 Flow 0.
- Figure 5.38 Flow Centroid 2 was associated Figure 5.20 Flow 3.

- Figure 5.38 Flow Centroid 3 was associated Figure 5.20 Flow 1.

Considering the Operational Pattern 2 air traffic flows centroids summary in Figure 5.39, the associations were as follows:

- Figure 5.40 Flow Centroid 0 was associated Figure 5.20 Flow 5 and Flow 6.
- Figure 5.40 Flow Centroid 1 was associated Figure 5.20 Flow 0.
- Figure 5.40 Flow Centroid 2 was associated Figure 5.20 Flow 2.

Thus, Operational Pattern 0 was generally associated with Flow 0, Flow 1, Flow 3, and Flow 4 displayed in Figure 5.20 and Operational Pattern 2 was generally associated with Flow 0, Flow 2, Flow 5, and Flow 6. The percentages of trajectories belonging to each air traffic flow displayed in Table 5.9 generally supported these associations. Further, the percentage of trajectories belonging to each air traffic flow displayed in Table 5.9 generally supported the notion that there was an important operational distinction between Operational Pattern 0 and Operational Pattern 2 that is not captured in the specification of the official SFO operational plans. For instance, considering only trajectories in the Operational Pattern 0 data set, the percentage of trajectories belonging to Flow 5 and Flow 6 was relatively negligible at 1.02% and 0.69%, respectively, whereas, considering only trajectories in the Operational Pattern 2 data set, the percentage of trajectories belonging to Flow 5 and Flow 6 was substantial, at 18.50% and 21.60%, respectively. Further, considering only trajectories in the Operational Pattern 0 data set, the percentage of trajectories belonging to Flow 4 was very substantial at 37.08%, whereas, considering only trajectories in the Operational Pattern 2 data set, the percentage of trajectories belonging to Flow 4 was relatively negligible at 1.82%. Knowledge of the distinction in the distribution of trajectories among air traffic flows depending on the observation of certain operational patterns may be relevant in ATM system operation and planning activities.

Further, for each operational state classification, the trajectories associated with the particular operational state classification were divided into sets based upon the trajectory category label (N, S, E, or B), where trajectories that were identified as having been go-arounds were also considered as a category. The percentage of trajectories that belonged to each trajectory category within the operational classification sets of trajectories was computed, where Table 5.10 displays these percentages.

Table 5.10: Percentage of Trajectories Associated with Each Operational Classification Belonging to One of the Four Trajectory Categories or Having Been Identified as a Go-Around

Category	**Transitional**	**Anomalous**	**Pattern 0**	**Pattern 2**
N	62.94%	66.94%	81.26%	71.78%
S	22.54%	17.50%	6.98%	21.43%
E	7.18%	7.50%	8.49%	3.93%
B	6.84%	6.11%	2.69%	2.47%
Go-Around	0.50%	1.94%	0.58%	0.39%

Overall, the trajectories within the nominal operational state data sets (Operational Pattern 0 and Operational Pattern 2) tended belong to the nominal (N) trajectory category with a higher frequency. Further, the trajectories within the nominal operational state data sets tended belong to the both-spatial-and-energy-anomaly (B) trajectory category with a noticeably lower frequency. It was observed that the Operational Pattern 0 data set had the highest proportion of trajectories in the only-energy-anomaly (E) category. Therefore, it may be important for ATM system operators to pay more attention to aircraft energy state management in the case that the spatial metrics of trajectories appear normal both at the flight level and the airspace level during time intervals conforming to Operational Pattern 0. Moreover, trajectories within the transitional operational state data set had the highest proportion of trajectories in the only-spatial-anomaly (S) category. This may have been

due to the fact that transitions between distinct airspace structures results in some adjustment to individual trajectories that may have caused them to appear spatially anomalous. Thus, ATM system operators should be aware of the potential consequences to multiple trajectories' spatial metrics conformance during transitional time intervals. Though, it was observed that trajectories within the Operational Pattern 2 data set also had a high proportion of only-spatial-anomaly (S) trajectories, which may be of interest to further investigate in the context of developing new ATM concepts of operation at KSFO. Finally, trajectories associated with the anomalous operational state data set had the highest proportion of trajectories experiencing go-arounds, where the proportion was nearly four times higher than for the remaining operational state classification data sets. As the airspace density matrix was generated as an aggregation of the trajectories spatial metrics' during a specified time interval, it was expected that those time intervals experiencing a higher proportion of go-arounds would have been detected as having anomalous operational states. The results of the flight-level analysis that occurred in the experimental approach related to the **Gap 2** in the context of the airspace-level analysis performed enables additional acceptance of the analysis at the airspace level to occur as well as provides new insights that may be of interest to ATM system operators, planners, and decision-makers.

Prediction

The accuracy, overall precision, overall recall, and overall F1-score computed on the testing data set for each multi-class classification algorithm (artificial neural network, XGBoost, and SVM) are summarized in Table 5.11.

Table 5.11: Summary of Classification Performance Metrics

Algorithm	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Artificial Neural Network	85.63%	0.86	0.76	0.80
XGBoost	89.02%	0.89	0.84	0.87
SVM	85.45%	0.86	0.77	0.81

All prediction models provided a testing accuracy that exceeded the 85% accuracy threshold specified. However, the XGBoost algorithm performed the best on the testing data set across all trained multi-class classification algorithms according to all classification performance metrics computed.

As an XGBoost prediction model was trained to produce a testing accuracy of 89.02%, **Hypothesis 3.2 is accepted**. Though, it was further of interest which features derived from the ASOS weather metrics were most significant in predicting the operational pattern for a time interval one hour ahead. The importance of each feature in the trained prediction model, as determined by the XGBoost algorithm [132] is displayed in Figure 5.48. The most important feature was the average relatively humidity over the past two hours, followed by the average wind speed over the past two hours, and wind direction in the previous hour.

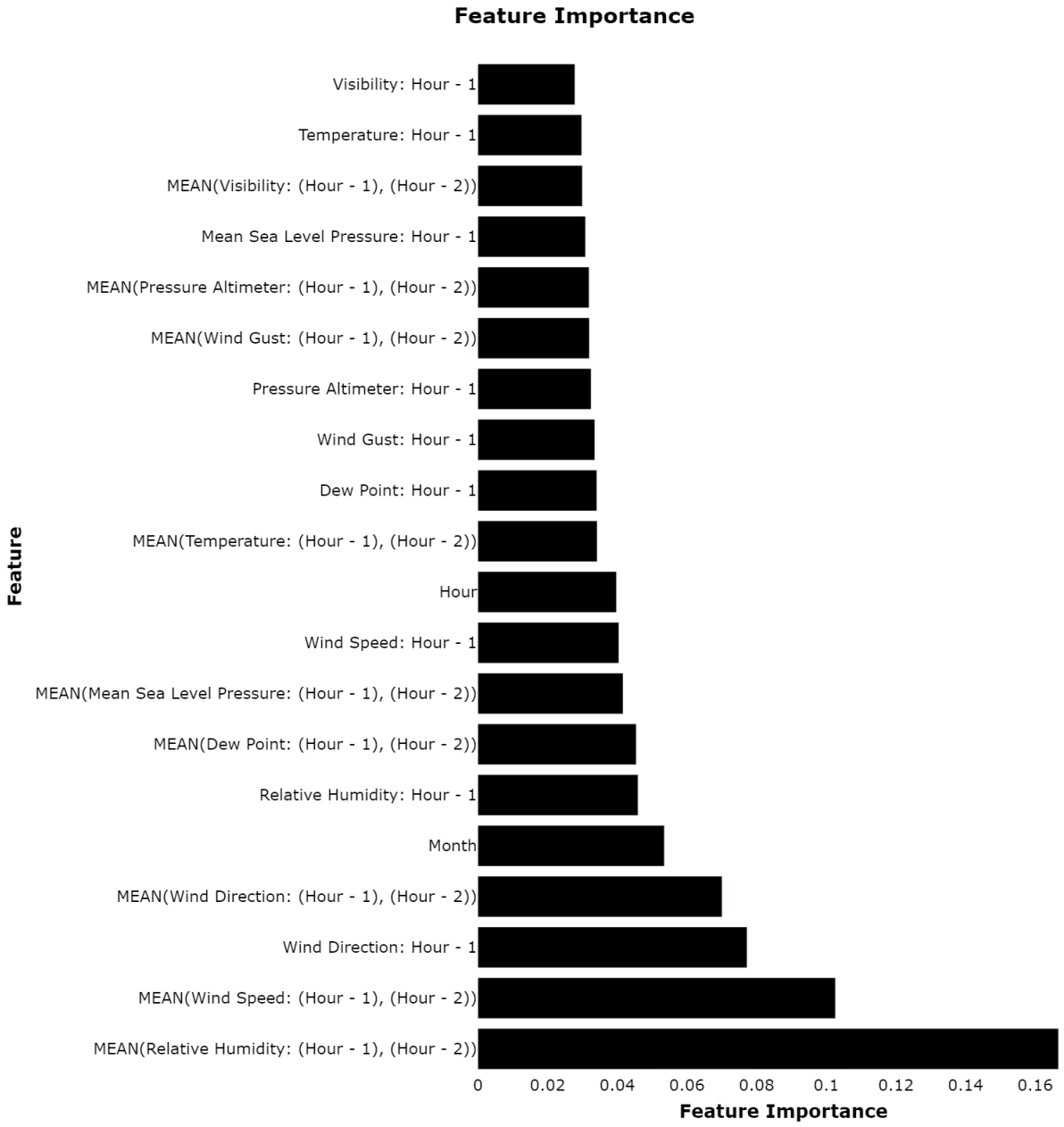

Figure 5.48: XGBoost Feature Importance

The acceptance of **Hypothesis 3.1** and **Hypothesis 3.2** fills **Gap 3**. The primary contribution of filling **Gap 3** is an airspace-level method to identify operational patterns, characterize whether the operational state is nominal, anomalous, or transitional, and predict operational patterns considering arriving aircraft operating during specified time intervals.

5.5 Summary

Three experimental approaches were implemented to test and accept the sets of hypotheses formulated in response to the three sets of research questions associated with the three gaps identified, where each gap was associated with a step in the proposed methodology. Each experimental approach was implemented considering ADS-B data that was extracted, cleaned, processed, and augmented for arriving aircraft at KSFO during the full year of 2019 in addition to the corresponding extracted and processed ASOS weather data.

The primary contribution of filling **Gap 1** is a more reliable method of identifying air traffic flows considering the converging nature of arriving aircraft trajectories. Additionally, the primary contribution of filling **Gap 2** is knowledge of the relationship between spatial and energy anomalies for arriving aircraft. Finally, the primary contribution of filling **Gap 3** is an airspace-level method to identify operational patterns, characterize whether the operational state is nominal, anomalous, or transitional, and predict operational patterns considering arriving aircraft operating during specified time intervals. The implementation of the experimental approaches outlined supports the adequacy of the steps in the proposed methodology. Therefore, the **Overarching Hypothesis is accepted**. For arriving aircraft at KSFO in 2019, an analysis of ATM system operations at the flight level and the airspace level was completed.

CHAPTER 6
CONCLUSION

This research was motivated by the overarching emphasis on increasing ATM system efficiency and capacity, while maintaining a high level of safety as the current systems undergo a transformation in concept of operations under global ATM system modernization plans. Considering the global modernization efforts' current state and future milestones, gaining a comprehensive understanding of both flight-level and airspace-level operations enables ATM system operators, planners, and decision-makers to make better-informed and more robust decisions related to the implementation of future operational concepts. Though, the complexity of ATM systems restricts the use of classical physics-based methods for the analysis of new operational concepts. However, an increased availability of operational data provides several new analysis opportunities. Specifically, ADS-B data provides the basis for offline data-driven methods to be applied to analyze ATM system operations at both the flight level and the airspace levels. Further, the analysis of ATM system arriving aircraft operations has the highest potential to impact system safety, capacity, and efficiency levels. Therefore, the **Overarching Research Question** was formulated as:

Overarching Research Question

How can the offline application of data-driven methods to ADS-B data be leveraged to analyze Air Traffic Management system arriving aircraft operations at both the flight level and the airspace level?

Through an initial literature survey, three relevant research areas were identified that may be considered steps in a proposed offline data-driven methodology to be applied to ADS-B data to analyze ATM system arriving aircraft operations at botht the flight level and the airspace level. Therefore, the **Overarching Hypothesis** was formulated as:

Overarching Hypothesis

If an offline data-driven methodology to be applied to ADS-B data for arriving aircraft is developed that performs (i) **Air Traffic Flow Identification**, (ii) **Anomaly Detection**, and (iii) an **Airspace-Level Analysis**, then Air Traffic Management system arriving aircraft operations are analyzed at both the flight level and the airspace level.

Subsequently, an in-depth literature review of the methods related to each step of the proposed methodology was performed in which three gaps were identified, each associated with a step in the proposed methodology. The gaps identified are expressed as follows:

Gap 1

A reliable method to identify air traffic flows that considers the converging nature of arrival trajectories and may ultimately be extensible to real-time applications.

Gap 2

A quantitative analysis of the relationship between spatial and energy anomalies detected in arriving aircraft ADS-B data.

Gap 3

An airspace-level method to identify operational patterns, characterize operational states (such as those that are nominal, transitional, or anomalous in nature), and predict operational patterns considering the operations of arriving aircraft.

Considering the three gaps, an **Overarching Research Objective** was formulated as:

Overarching Research Objective

Contribute to the development of an offline data-driven methodology to be applied to ADS-B data for arriving aircraft that enables and performs:

- The reliable identification of air traffic flows
- A quantitative investigation of the relationship between spatial and energy anomalies
- At the airspace level, the identification of operational patterns, characterization of operational states, and prediction of operational patterns

Three sets of research questions and associated hypotheses were formulated related to filling the three gaps and fulfilling the **Overarching Research Objective**. To answer the research questions an experimental approach was designed to test the associated hypothesis/hypotheses such that they were accepted and each gap was filled. Inductively, the filling of each gap to fulfill the **Overarching Research Objective** enabled the acceptance of the **Overarching Hypothesis**.

Research Question 1 was posed due to the identification of **Gap 1** related to the **Air Traffic Flow Identification** step of the proposed methodology. **Gap 1** specified the need for a more reliable method of identifying air traffic flows of arriving aircraft such that the converging nature of arriving aircraft trajectories is considered and the method could be extensible for use in real-time in the future. Therefore, **Research Question 1** inquired how an existing air traffic flow identification could be modified to meet the aforementioned requirements. After evaluating the existing air traffic flow identification methods, the implementation of HDBSCAN with the ED was determined to be the best-suited, where the ED is limited when applied to trajectories that are converging in nature, such as arriving aircraft trajectories, in that there exists an uneven distribution of distances closer to both the airport location and the defined terminal airspace border, which may skew the classifi-

cation of trajectories. HDBSCAN was selected as the clustering algorithm to perform air traffic flow identification due the following three properties: (i) requires no a priori specification of the number of clusters to be identified, (ii) is capable of identifying outliers, and (iii) is capable of identifying clusters of varying densities. However, to overcome the limitation of the use of the ED, the use of the WED was proposed, where **Hypothesis 1** was formulated accordingly. Specifically, it was proposed to use a WED weighting scheme that weight trajectory points less closest to the airport location and closest to the defined terminal airspace border. **Experiment 1** was designed to test whether implementation of HDBSCAN with the ED or the implementation of HDBSCAN with the WED identified air traffic flows "more reliably", where a quantitative metric was derived to evaluate reliability. Results of the implementation of **Experiment 1** lead to **Hypothesis 1** being accepted.

Research Question 2 was posed due to the identification of **Gap 2** related to the **Anomaly Detection** step of the proposed methodology. **Gap 2** specified the need for a quantitative analysis to be performed to determine the relationship, if any, between spatial and energy anomalies detected in arriving aircraft ADS-B data. Related to **Gap 2** was the prior novel distinction between spatial and energy anomalies detected in ADS-B data. However, while air traffic flow identification (spatial anomaly detection) appeared to occur in some instances prior to energy anomaly detection, the manner in which the spatial anomalies were considered was unclear. Therefore, **Research Question 2** inquired what the relationship was between spatial and energy anomalies detected in ADS-B data. Three hypotheses were formulated related to the relationship between spatial and energy anomalies. **Hypothesis 2.1** stated that if a trajectory is detected as being a spatial anomaly, then it is more likely to be detected as an energy anomaly. **Hypothesis 2.2** stated that if only energy-nominal trajectories are considered, then trajectories that have been detected only as spatial anomalies (not energy anomalies) have a higher "degree of anomalousness" in their energy metrics than those trajectories not detected as spatial anomalies. Finally, **Hypothesis 2.3** stated that if only energy-anomalous trajectories are considered then, when

separated into groups based on whether the trajectory has previously been detected as a spatial anomaly or not, the distributions of a measure of the "degree of anomalousness" in energy metrics for the two groups are dissimilar (upon evaluation of statistical properties). **Experiment 2** was designed to test the three hypotheses, which required a procedure to perform spatial anomaly detection, then energy anomaly detection, and, finally, compute an anomaly score as a measure of the "degree of anomalousness" of a trajectory's energy profile such that a statistical analysis of the results could be performed. Spatial anomaly detection was performed by the more reliable method of identifying air traffic flows considering **Hypothesis 1**. A clustering algorithm was similarly identified to be most appropriate for energy anomaly detection. However, the requirements of the clustering algorithm differ compared with the requirements for spatial anomaly detection. Namely, it is not necessary for an algorithm to have the capability to identify clusters of varying densities; therefore, DBSCAN was selected to perform energy anomaly detection. Results of the implementation of **Experiment 2** lead to **Hypothesis 2.1** being accepted, **Hypothesis 2.2** being accepted, and **Hypothesis 2.3** being accepted.

Research Question 3.1 and **Research Question 3.2** were posed due to the identification of **Gap 3** related to the **Airspace-Level Analysis** step of the proposed methodology. **Gap 3** specified the need for an airspace-level method to identify operational patterns, characterize whether the operational state is nominal, anomalous, or transitional, and predict operational patterns considering arriving aircraft operating during specified time intervals. Specifically, **Research Question 3.1** inquired how a data set of operational state representations could be analyzed such that operational patterns are identified and operational states are characterized as either nominal, anomalous, or transitional. The small number of existing methods were observed to have a number of limitations, where several of the methods were not truly "airspace-level". However, it was determined that a clustering algorithm capable of identifying outliers and requiring no a priori specification of the number of clusters is most appropriate. Due to meeting the aforementioned criteria and consider-

ing the effectiveness of DBSCAN in other aviation applications, DBSCAN was selected. However, the presence of transitional operational states within the data set may result in skewed results. Therefore, a recursive DBSCAN procedure was proposed such that transitional operational states are removed from the data set and the clustering repeats until no new transitional operational states are characterized, and **Hypothesis 3.1** is formulated accordingly. **Experiment 3.1** was designed to test **Hypothesis 3.1**. Though, to perform **Experiment 3.1**, the operational state representations were required. Therefore, grid-based matrices representing the density of trajectories within the terminal airspace during specified time intervals were generated. After assembling the entire set of matrices (airspace density matrices), the matrices were flattened and their dimension was reduced (leveraging UMAP) to provide a suitable input to apply DBSCAN. Results of the implementation of **Experiment 3.1** indicated that distinct operational patterns that aligned with distinct official operational patterns were identified and it appeared as though noticeable transitional and anomalous operational states were characterized, which lead to **Hypothesis 3.1** being accepted. Considering the need for the capability to predict operational patterns specified in **Gap 3**, **Research Question 3.2** inquired how operational patterns could be predicted. The few existing methods related to prediction of what may be considered operational patterns leveraged features derived from recorded weather measurements. Further, these methods tended to investigate the performance of a handful of classification algorithms. Therefore, it was proposed to train a set of classification algorithms with features derived from recorded weather measurements and select the best performing model with respect to accuracy, where **Hypothesis 3.2** was formulated accordingly. **Experiment 3.2** was designed to test **Hypothesis 3.2** in which artifical neural network, gradient-boosted decision tree (XGBoost), and SVM algorithms were trained considering weather metrics as input features. Results of the implementation of **Experiment 3.2** indicated that the XGBoost model performed the best with a reasonable accuracy, which lead to **Hypothesis 3.2** being accepted.

KSFO was selected as the terminal airspace for which data was extracted, cleaned, and augmented to enable the testing of the hypotheses. ADS-B data was extracted for the full year of 2019 from the OpenSky Network historical database [135] and ASOS weather data was extracted from the Iowa Environmemtal Mesonet [140]. In testing the hypotheses, a demonstration implementation of the steps proposed methodology was simultaneously performed. The complete offline data-driven methodology to be applied to ADS-B data for arriving aircraft to perform analysis at both the flight level and the airspace level is presented in Figure 6.1.

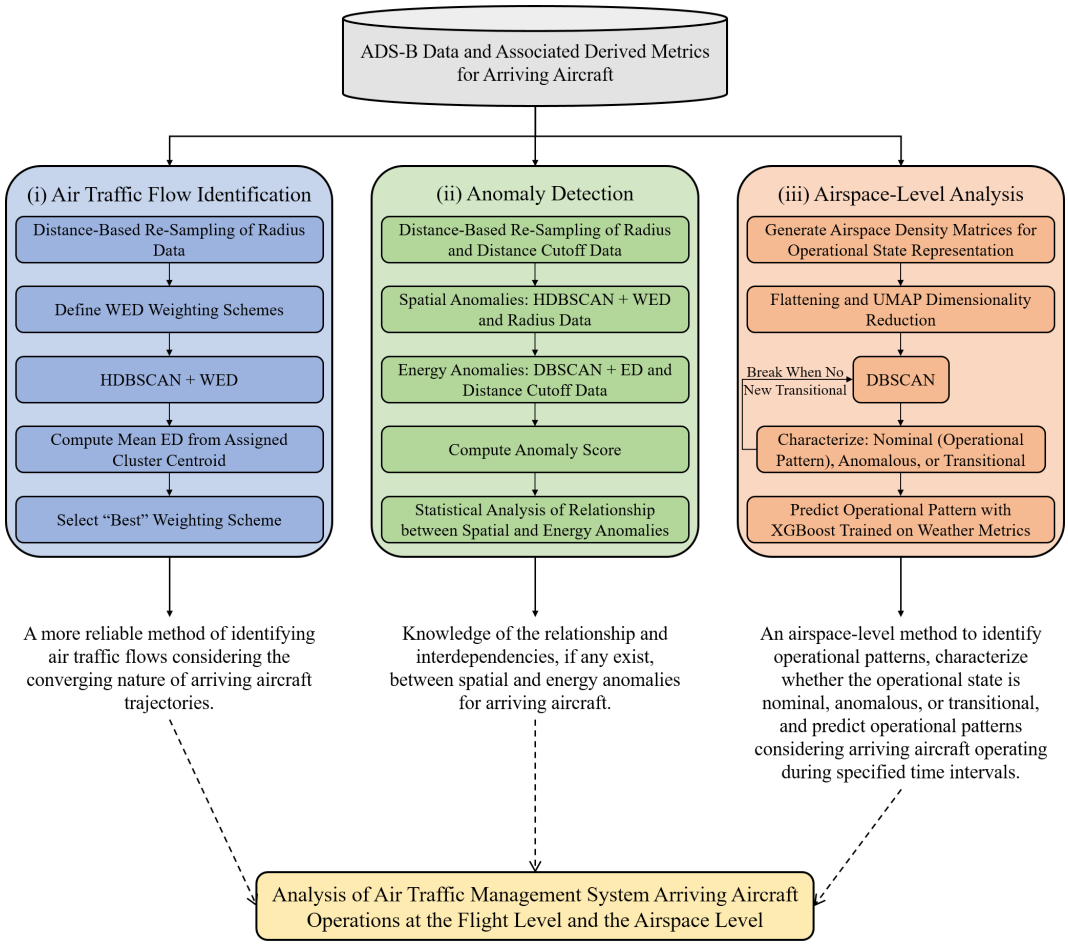

Figure 6.1: A Data-Driven Methodology to Analyze Air Traffic Management System Operations within the Terminal Airspace

6.1 Contributions

The filling of each of the three gaps produces significant contributions to the state-of-the-art terminal airspace ATM system analysis methods within the aviation literature. Further, the uniting of the three steps in a comprehensive methodology to analyze ATM system arriving aircraft operations at both the flight level and the airspace level exists as a contribution. The proposed methodology is intended to be applied prior to exploring new ATM system operational concepts and to support future capacity, efficiency, and safety gains. It is important to understand the current status of ATM system operations within a terminal airspace prior to making changes and also to inform the changes that are made.

The primary contribution related to filling **Gap 1** is a more reliable method of identifying air traffic flows that considers the converging nature of arriving aircraft trajectories and is possible to extend to a real-time application. Air traffic flow identification is relevant both as a stand-alone analysis and as a data processing step for other ATM system analysis methods. Therefore, a more reliable method for the identification of air traffic flows has far-reaching implications. Completion of the air traffic flow identification step in the proposed methodology enables ATM system operators, planners, and decision-makers to gain a baseline understanding of the dominant patterns of air traffic, i.e. where, spatially, aircraft tend to operate and with what density. Further, in the context of developing trajectory prediction algorithms, developing algorithms with knowledge of standard air traffic flow operations is common such that the air traffic flow information is leveraged to make better predictions.

The primary contribution related to filling **Gap 2** is a procedure to explore the relationship between spatial and energy anomalies. Further, the novel distinction between spatial anomalies and energy anomalies detected in ADS-B data is important. Overall, the results stemming from the completion of the **Anomaly Detection** step in the proposed methodology may be of great interest to ATM system operators, decision-makers, and planners,

specifically as ATM systems are modernized, new technologies are deployed, a new concept of operations is implemented. It is necessary to understand the current dynamics in the spatial and energy dimensions, and their potential interdependencies, such that these may be considered when evaluating the deployment of a new technology or the implementation of a new concept of operation. Moreover, if completion of the **Anomaly Detection** step reveals that, generally, if a trajectory is detected as a spatial anomaly, it is more likely to also be detected as an energy anomaly, then actionable insights may be derived. For instance, it would be advantageous to alert pilots that an increased emphasis should be placed on management the aircraft's energy state, whether a trajectory is experiencing off-nominal spatial operations due to pilot actions or instructions from ATC. This could aid in prevention of potentially risky or unsafe situations associated with poor energy management, such as unstable approaches, runway excursions, etc.

The primary contribution related to filling **Gap 3** is a method to identify operational patterns, characterize whether an operational state is nominal, anomalous, or transitional, and predict operational patterns for arriving aircraft operating during specified time intervals. Notably, an airspace-level analysis method had not previously been presented within the literature that analyzes the aggregation of the trajectory time-series data of all arriving aircraft operating during specified time intervals. The characterization of operational states as being transitional enables additional insights to be obtained regarding the terminal airspace structure. For instance, knowledge of the time intervals in which the operational pattern transitions from one to another may be valuable for ATM system planners and decision-makers as new operational concepts are developed and implemented. Further, the conformance of flight-level operations to standard operations may be evaluated for transitional operational states. The characterization of operational states as being transitional has many benefits over considering these transitional operational states to be a distinct operational pattern or grouped in with an existing identified pattern, as may occur applying existing methods. Finally, while the ability to predict is not especially relevant in the context of

an offline analysis to be performed to support decision-making, the "decision boundaries" that may be derived from the prediction model are relevant. This derived output from the prediction models may be leveraged to further understand the circumstances related to the observation of any of the distinct operational patterns. Therefore, when considering future planning for implementation of a new concept of operations, these circumstances may be considered.

6.1.1 Observations: San Francisco International Airport

Application of the proposed methodology to analyze ATM system arriving aircraft operations at KSFO in 2019 resulted in several KSFO-specific observations being made. For instance, the identification of air traffic flows revealed that there typically exists very little air traffic towards the west side of the terminal airspace within 20 nautical miles of the airport location. This may be important to be aware of as plans to integrate AAM operations into the NAS are developed.

Further, the exploration of the relationship between spatial and energy anomalies detected in ADS-B data revealed that, generally, if a trajectory was detected as a spatial anomaly, it was more likely to also be detected as an energy anomaly. Additionally, the average anomaly score for trajectories detected as only spatial anomalies was generally higher than the average anomaly score for nominal trajectories. ATM system operators, such as ATC, may benefit from this observation, as it may factor into decision-making in instances where there is the potential to instruct an aircraft to spatially deviate from standard operations. Moreover, it was revealed that there appeared to be underlying differences in energy metrics sequences of trajectories detected only as energy anomalies and trajectories detected as both spatial and energy anomalies, which should be considered when making decisions on actions that air traffic controllers or pilots should take as well as in development of other analysis methods.

With respect to an airspace-level analysis of ATM system arriving aircraft operations,

it was revealed that there existed three distinct operational patterns that were observed in 2019. However, the official SFO operational plans only indicate two distinct plans: the official SFO west plan and the official SFO south east plan. The identification of operational patterns reveals that the official SFO west plan actually manifests as two distinct sets of air traffic flows that are observed. This may be of interest to ATM system operators. Further, the typical time intervals in which the operational patterns transition between one another may be of interest.

6.1.2 Publications

At the time of writing, two published works accompany this thesis:

- **Corrado, S.J.**, Puranik, T.G., Pinon, O.J., & Mavris, D.N. (2021). A clustering-based quantitative analysis of the interdependent relationship between spatial and energy anomalies in ADS-B trajectory data. Transportation Research Part C: Emerging Technologies, 131. https://doi.org/10.1016/j.trc.2021.103331.

- **Corrado, S.J.**, Puranik, T.G., Pinon, O.J., & Mavris, D.N. (2020). Trajectory Clustering within the Terminal Airspace Utilizing a Weighted Distance Function. Proceedings, 59(1), 7. MDPI A, https://doi.org/10.3390/proceedings2020059007

It is noted these published works are associated with the first two steps in the proposed methodology (**Air Traffic Flow Identification** and **Anomaly Detection**). Two published works associated with the final step in the proposed methodology (**Airspace-Level Analysis**) are forthcoming, though are not completed at the time of writing.

Undoubtedly, an additional major contribution of this thesis is the curation of a cleaned, processed, and augmented ADS-B trajectory data set of arriving aircraft within the KSFO terminal airspace in 2019. Additional published work leveraging the cleaned, processed, and augmented ADS-B trajectory data set of arriving aircraft within the KSFO terminal airspace in 2019 includes:

- Kumar, S.G., **Corrado, S.J.**, Puranik, T.G., Mavris, D.N.. Classification and Analysis of Go-Arounds in Commercial Aviation Using ADS-B Data. Aerospace. 2021; 8(10):291. https://doi.org/10.3390/aerospace8100291

- **Samantha J. Corrado**, Tejas G. Puranik, Olivia J. Pinon-Fischer, Dimitri Mavris, Rodrigo Rose, Jesse Williams and Roohollah Heidary. "Deep Autoencoder for Anomaly Detection in Terminal Airspace Operations," AIAA 2021-2405. AIAA AVIATION 2021 FORUM. August 2021. https://doi.org/10.2514/6.2021-2405

6.2 Limitations and Recommendations for Future Work

The potential limitations of the proposed methodology are discussed. Addressing some of these potential limitations offer opportunities for future work. Thus, recommendations for future work are presented.

6.2.1 Limitations

A few limitations related to the application of the proposed methodology may exist. For instance, no inherent capability exists to take into account potential dynamics and interdependencies related to metroplex operations, i.e. the operations of aircraft within the overlapping terminal airspace areas associated with multiple airports. For certain complex and high-traffic metroplex systems, such as the New York City metroplex, application of the proposed methodology separately to each airport's ATM system arriving aircraft operations may be insufficient to provide meaningful insights to ATM system operators, planners, and decision-makers. Additionally, limitations related to the specific application of each of the steps in the proposed methodology are discussed.

Air Traffic Flow Identification

Considering the **Air Traffic Flow Identification** step, despite the more reliable identification of air traffic flows by implementing HDBSCAN with the WED, the WED does

suffer from one primary limitation in common with the ED. This limitation is the potential for "misalignment" of trajectory points, i.e. trajectory points may not be ordered after re-sampling such that each successive trajectory point is closest to its corresponding trajectory point with respect to all other trajectory points. The impact of the misalignment may become more pronounced when considering air traffic flows that are more "spread out".

Anomaly Detection

The **Anomaly Detection** step suffers primarily from the limitation of the quality/quantity of information available within the ADS-B data. For instance, data errors, especially in open-source ADS-B data, are common and may impact results for specific trajectories. Though, more relevant is that the energy metrics computed do not take into account aircraft weight, which very significant in the context of energy management. However, ADS-B data does not include weight information nor precise information from which weight information may be reliably derived. Instantaneous metrics that either directly or indirectly provides aircraft weight information are often readily available in FOQA data. Further, availability of this information would enable stabilized approach criteria to be fully assessed, which would enable some "labels" to be assigned to certain trajectories.

Airspace-Level Analysis

The **Airspace-Level Analysis** step suffers from the limitation that the airspace density matrix may not be the "best" method of representing the aggregation of the time-series trajectory data for all arriving aircraft operating during a specified time interval. For instance, a threshold is required to be set such that time intervals that do not contain at least a specified number of trajectories are discarded. This may result in loss of information, Specifically, removal of certain time intervals either before or after true transitional operational states may result in the transitional operational states not being able to be characterized.

6.2.2 Recommendations for Future Work

Several avenues for future work related to the proposed offline data-driven methodology to be applied to ADS-B data to analyze ATM system arriving aircraft operations exist. For instance, considering the potential limitation of the proposed methodology to be applied to complex and high-capacity metroplex systems, the proposed methodology may be expanded to consider metroplex systems. Specifically, the most significant modification would occur to the **Airspace-Level Analysis** step of the proposed methodology to enable an expansion to metroplex systems.

In the context of future applications, while the current applications of the proposed methodology do include analysis to support integration of AAM concepts, further specification of the proposed methodology to include returning an analysis of the most appropriate heliport location(s) within a terminal airspace may be beneficial. Moreover, the proposed methodology may be augmented with a step to provide an analysis of airspace availability for AAM operations within specified time intervals.

With respect to enhancing the **Anomaly Detection** step of the proposed methodology, it is of interest to analyze the anomalies detected (both spatial and energy) in the context of a reduced set of stabilized approach criteria. As mentioned, the evaluation of the full set of stabilized approach criteria is not possible without aircraft weight information. However, a reduced set may be evaluated with respect to the ADS-B data metrics, which may provide some sort of guidance related to the assessment of the spatial and energy anomalies detected. It would also be advantageous to augment the evaluation of the anomalies detected with an assessment of ATC recordings. Recently, the OpenSky Network [135] has announced plans to provide ATC recordings to researchers for analysis [137].

In addition, it may be desirable to explore the re-ordering of the proposed methodology steps such that the **Airspace-Level Analysis** step occurs before and informs all other steps. For instance, air traffic flows may be identified "within" operational patterns, analogous to how energy anomalies are detected "within" air traffic flows. Performing subsequent

analyses based upon the operational pattern observed for arriving aircraft operating during a specified time interval may shed more light on potential airspace-level behavior and flight-level behavior interdependencies.

Finally, with respect to application of the proposed methodology, this thesis research focused solely on arriving aircraft with the KSFO terminal airspace in 2019. An avenue for extension and further testing of the proposed methodology includes the application to more diverse terminal airspace areas. For instance, the KSFO terminal airspace is actually rather unique in the relatively dominant use of the official SFO west plan as well as its relatively constant weather patterns/variation throughout the year. Therefore, applying the proposed methodology to data extracted for one or more different terminal airspace is recommended.

REFERENCES

[1] International Civil Aviation Organization, *Future of Aviation*, https://www.icao.int/Meetings/FutureOfAviation/.

[2] Air Transport Bureau, "Effects of novel coronavirus (covid-19) on civil aviation: Economic impact analysis," ICAO, Monteal, Canada, Oct. 2021.

[3] Federal Aviation Administration, *FAA Aerospace Forecast Fiscal Years 2021-2041*, https://www.faa.gov/data_research/aviation/aerospace_forecasts/media/FY2021-41_FAA_Aerospace_Forecast.pdf, 2021.

[4] B. Pearce, "Covid-19 outlook for air traval in the next 5 years," IATA, May 2020.

[5] E. Gulbas, "Covid-19 airline industry outlook," IATA, Oct. 2021.

[6] IATA, "Economic performance of the airline industry: 2021 end-year report," IATA, Oct. 2021.

[7] Federal Aviation Administration, "FAA Strategic Plan FY 2019-2022," FAA, 2019.

[8] International Civil Aviation Organization, *World Aviation and the World Economy*, https://en.wikipedia.org/wiki/Map_projection.

[9] "Doc 4444: Procedures for Air Navigation Services Air Traffic Management," ICAO, 2016.

[10] "laying down the framework for the creation of the single European sky (the framework regulation)," Journal of the European Union, Mar. 2004.

[11] Mayara Conde Rocha Murca, "Data-Driven Modeling of Air Traffic Flows for Advanced Air Traffic Management," Ph.D. dissertation, Massachusetts Institute of Technology, 2018.

[12] Federal Aviation Administration, *Modernization of the U.S. Airspace*, https://www.faa.gov/nextgen/, 2020.

[13] Single European Sky (SES), *Single European Sky*, https://www.sesarju.ed/background-ses, 2020.

[14] International Civil Aviation Organization, "The implementation of the aviation system block upgrade in china," in *Thirteeth Air Navigation Conference*, ICAO, Oct. 2018.

[15] Civil Aviation Bureau, Japan, *CARATS: Collaborative Actions for Renovation of Air Traffic Systems*, https://www.icao.int/APAC/Meetings/2010/atfm_sg1/.

[16] Departamento de Controle de Espaco Aereo, *SIRIUS: What is it?* https://www.decea.gov.br/sirius/.

[17] Federal Aviation Administration, "NextGen Annual Report: A Report on the History, Current Status, and Future of National Airspace System Modernization," FAA, 2020.

[18] Federal Aviation Administration, "Concept of Operations for the Next Generation Air Transportation System. Version 3.2," FAA, 2011.

[19] International Civil Aviation Organization, "Global Air Traffic Management Operational Concept," ICAO, 2005.

[20] International Civil Aviation Organization, *Strategic Objectives*, https://www.icao.int/about-icao/Council/Pages/Strategic-Objectives.aspx.

[21] Federal Aviation Administration, "Vision for Trajectory Based Operations, Version 2.0," FAA, Sep. 2017.

[22] International Civil Aviation Organization, *Doc 9859 safety management manual*, 4th ed., ICAO, 2018, 192 pp.

[23] D. R. Insua, C. Alfaro, J. Gomez, P. Hernandez-Coronado, and F. Bernal, "A framework for risk management decisions in aviation safety at state level," *Reliability Engineering & System Safety*, vol. 179, no. 71, pp. 74–82, Nov. 2018.

[24] T. J. Logan, "Error prevention as developed in airlines," *International Journal of Radiation Oncology, Biology, Physics*, vol. 71, no. 1, S178–S181, 2008.

[25] Puranik, T.G., "A Methodology for Quantitative Data-driven Safety Assessment for General Aviation," Ph.D. dissertation, Georgia Institute of Technology, 2018.

[26] Federal Aviation Administration, *Aviation Safety Action Program*, https://www.faa.gov/initiatives/asap, Aug. 2017.

[27] National Aeronautics and Space Administration, *Aviation Safety Reporting System*, https://asrs.arc.nasa.gov/overview/summary.html.

[28] Federal Aviation Administration, *Advisory Circular 120-82 Flight Operational Quality Assurance*, Washington DC, United States, 2004.

[29] Federal Aviation Administration, *Aviation Safety Information Analysis and Sharing Program*, https://www.faa.gov/news/fact_sheets/news_story.cfm?newsId=18195, Apr. 2016.

[30] Federal Aviation Administration, *System Approach for Safety Oversight Program*, https://www.faa.gov/about/initiatives/saso/, Oct. 2019.

[31] Federal Aviation Administration, *Safety Assurance Systsem*, https://www.faa.gov/about/initiatives/sas/, Oct. 2019.

[32] National Aeronautics and Space Administration, *Strategic Implementation Plan 2019 Update*. 2019.

[33] National Academies of Sciences, Engineering, and Medicine, *In-Time Aviation Safety Management: Challenges and Research for an Evolving Aviation System (2018)*. Washington, DC, United States: The National Academies Press, 2018.

[34] D. Rinehart, H. Jimenez, and M. Blake, "Development of real-time system-wide safety assurance definitions and concept fundamentals," in *16th AIAA Aviation Technology, Integration, and Operations Conference*, AIAA, Washington, DC, United States, Jun. 2016, pp. 1–12.

[35] L. Cary, *Capacity: Factors to Consider*, https://www.icao.int/Meetings/AMC/MA/2006/atfm/pres12.pdf, Mar. 2006.

[36] International Civil Aviation Organization, *Doc 9613 performance-based navigation manual*, 4th ed., ICAO, 379 pp.

[37] Federal Aviation Administration, *The Future of the NAS*. 2016.

[38] ICAO Air Traffic Management Requirements and Performance Panel (ATMRPP), "Global TBO Concept: Version 0.11," ICAE.

[39] Federal Aviation Administration, *Trajectory Based Operations (TBO)*, https://www.faa.gov/air_traffic/technology/tbo/, Feb. 2021.

[40] J. Kim, "Emergence: Core ideas and issues," *Synthese*, vol. 151, pp. 547–559, Aug. 2006.

[41] Wikipedia, *Emergence*, https://en.wikipedia.org/wiki/Emergence.

[42] "Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities," INCOSE, 2015.

[43] M. Sharma, K. Das, M. Bilgic, B. Matthews, D. Nielsen, and N. Oza, "Active learning with rationales for identifying operationally significant anomalies in aviation," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Sep. 2016, pp. 209–225.

[44] T. Puranik and D. N. Mavris, "Identification of instantaneous anomalies in general aviation operations using energy metrics," *Journal of Aerospace Information Systems*, vol. 17, no. 1, pp. 51–65, 2020.

[45] L. Li, J. Hansman, R. Palacios, and R. Welsch, "Anomaly detection via a gaussian mixture model for flight operation and safety monitoring," *Reliability Engineering & System Safety*, vol. 193, Jan. 2020.

[46] B. Matthews, S. Das, K. Bhaduri, K. Das, R. Martin, and N. Oza, "Discovering anomalous aviation safety events using scalable data mining algorithms," *Journal of Aerospace Information Systems*, vol. 10, no. 10, Oct. 2013.

[47] K. Sheridan, T. Puranik, E. Mangortey, O. J. Pinon, M. Kirby, and D. N. Mavris, "An application of dbscan clustering for flight anomaly detection during the approach phase," in *2020 AIAA SciTech Forum*, AIAA, Jan. 2020, pp. 1–20.

[48] S. Pozzi, C. Valbonesi, V. Beato, R. Volpini, F. M. Giustizieri, F. Lieutaud, and A. Lieu, "Safety monitoring in the age of big data," in *Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)*, 2011, pp. 1–9.

[49] L. Basora and T. D. Xavier Olive, "Recent advances in anomaly detection methods applied to aviation," *Aerospace*, vol. 6, no. 117, 2019.

[50] Raj Deshmukh, "Data-Driven Anomaly and Precursor Detection in Metroplex Airspace Operations," Ph.D. dissertation, Purdue University, 2020.

[51] B. Sridhar, "Applications of machine learning techniques to aviation operations: Promises and challenges," in *2020 International Conference on Artificial Intelligence and Data Analytics for Air Transportation (AIDA-AT)*, 2020, pp. 1–12.

[52] T. Hoyland, C. Spafford, and A. Medland, *MRO Big Data - A Lion or a Lamb?* 2016.

[53] M. C. R. Murca, R. J. Hansman, L. Li, and P. Ren, "Flight trajectory data analytics for characterization f air traffic flows: A comparative analysis of terminal area operations between new york, hong kong and sao paulo," *Transportation Research Part C*, vol. 97, pp. 324–347, 2018.

[54] X. Olive and J. Morio, "Trajectory clustering of air traffic flows around airports," *Aerospace Science and Technology*, vol. 84, pp. 776–781, 2019.

[55] A. Gavrilovski, H. Jimenez, D. N. Mavris, A. H. Rao, S. Shin, I. Hwang, and K. Marais, "Challenges and opportunities in flight data mining: A review of the state of the art," in *AIAA Infotech@Aerospace*, San Diego, California, United States, Jan. 2016, pp. 1–18.

[56] B. Matthews, "Automatic anomaly detection with machine learning," in *World Aviation Festival*, 2019.

[57] A. Kallio and J. Tuimala, "Data mining," *Encyclopedia of Systems Biology*, 2013.

[58] M. C. R. Murca, R. J. Hansman, and H. Balakrishnan, "Trajectory clustering and classification of air traffic flows," in *16th AIAA Aviation Technology, Integration, and Operations Conference*, AIAA, Washington, DC, United States, 2016.

[59] F. Netjasov, M. Janic, and V. Tosic, "Developing a generic metric of terminal airspace traffic complexity," *Transportmetrica*, vol. 7, no. 5, pp. 369–394, 2011.

[60] Metroplex Environmental, http://metroplexenvironmental.com/oapm.html.

[61] Airbus, *A Statistical Analysis of Commercial Aviation Accidents 1958-2020*, 2020.

[62] X. Olive and P. Bieber, "Quantitative assessments of runway excursion precursors using mode s data," in *ICRAT - International Conference for Research in Air Transportation*, Jun. 2018.

[63] M. Gariel, A. N. Srivastava, and E. Feron, "Trajectory clustering and an application to airspace monitoring," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1511–1524, 2011.

[64] R. Deshmukh and I. Hwang, "Incremental-learning-based unsupervsied anomaly detection algorithm for terminal airspace operations," *Journal of Aerospace Information Systems*, vol. 19, no. 9, pp. 362–384, 2019.

[65] X. Olive, J. Grignard, T. Dubot, and J. Saint-Lot, "Detecting controllers' actions in past mode s data by autoencoder-based anomaly detection," in *SESAR Innovation Days*, 2019, pp. 1–8.

[66] X. Olive and L. Basora, "Identifying anomalies in past en-route trajectories with clustering and anomaly detection methods," in *Thirteenth USA/Europe Air Traffic Management Research and Development Seminar*, 2019, pp. 1–10.

[67] X. Olive, L. Basora, B. Viry, and R. Alliger, "Deep trajectory clustering with autoencoders," in *International Conference on Research in Air Transportation 2020*, ICRAT, Tampa, Florida, United States, 2020, pp. 1–8.

[68] X. Olive and L. Basora, "Detection and identification of significant events in historical aircraft trajectory data," *Transportation Research Part C: Emerging Technologies*, vol. 119, 2020.

[69] A. Tanner and M. Strohmeier, "Anomalies in the sky: Experiments with traffic densities and airport runway use," in *7th OpenSky Workshop 2019*, 2019, pp. 1–12.

[70] S. Madar, T. G. Puranik, and D. N. Mavris, "Application of trajectory clustering for aircraft conflict detection," in *2021 IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)*, IEEE, 2021, pp. 1–9.

[71] A. Behere, L. Isakson, T. G. Puranik, Y. Li, M. Kirby, and D. Mavris, "Aircraft landing and takeoff operations clustering for efficient environmental impact assessment," in *AIAA Aviation 2020 Forum*, 2020, p. 2583.

[72] A. Behere, J. Bhanpato, T. G. Puranik, M. Kirby, and D. N. Mavris, "Data-driven approach to environmental impact assessment of real-world operations," in *AIAA Scitech 2021 Forum*, 2021, p. 0008.

[73] L. Basora, J. Morio, and C. Mailhot, "A trajectory clustering framework to analyse air traffic flows," in *7th SESAR Innovation Days*, Belgrade, Serbia, 2017, pp. 1–8.

[74] S. Kaushik, *An introduction to clustering and different methods of clustering*, https://www.analyticsvidhya.com/blog/2016/11/an-introduction-to-clustering-and-different-methods-of-clustering/, 2016.

[75] M. Enriquez, "Identifying temporally persistent flows in the terminal airspace via spectral clustering," in *Air Traffic Management R&D Seminar*, 2013, pp. 1–8.

[76] A. Eckstein, "Automated flight track taxonomy for measuring benefits from performance based navigation," in *2009 Integrated Communications, Navigation and Surveillance Conference*, 2009, pp. 1–12.

[77] K. Pearson, "Liii. on lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 6, pp. 559–572, 1901.

[78] S. P. Lloyd, "Least squares optimization in pcm," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, 1982.

[79] F. Rehm, "Clustering of flight tracks," in *AIAA Infotech@Aerospace*, 2010, pp. 1–9.

[80] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Second International*

Conference on Knowledge Discovery and Data Mining, Portland, Oregon: AAAI Press, 1996, pp. 226–231.

[81] M. Enriquez and C. Kurcz, "A simple and robust flow detection algorithm based on spectral clustering," in *International Conference on Research in Air Transportation*, Berkley, California, United States, 2012, pp. 1–6.

[82] A. Marzuoli, M. Gariel, A. Vela, and E. Feron, "Data-based modeling and optimization of en route traffic," *Journal of Guidance, Control, and Dynamics*, vol. 37, no. 6, pp. 1930–1945, 2014.

[83] M. C. R. Murca and R. J. Hansman, "Identification, characterization, and prediction of traffic flow patterns in multi-airport systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 5, pp. 1683–1696, 2018.

[84] D. Delahaye, S. Puechmorel, S. Alam, and E. Feron, "Trajectory mathematical distance applied to airspace major flows extraction," in *EIWAC 2017: Air Traffic Management and Systems III*, 2017, pp. 51–66.

[85] R. J. G. B. Campello, D. Moulavi, and J. Sander, "Density-based clustering based on hierarchical estimates," in *Advances in Knowledge Discovery and Data Mining*, AAAI Press, 2013, pp. 160–172.

[86] D. H. Douglas and T. K. Peucker, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature," *Cartographica: The International Journal for Geographic Information and Geovisualization*, vol. 10, pp. 112–122, 1973.

[87] X. Olive and L. Basora, "A python toolbox for processing air traffic data: A use case with trajectory clustering," in *7th OpenSky Workshop 2019*, 2019, pp. 1–12.

[88] L. van der Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of Machine Learning Research*, vol. 9, no. 86, pp. 2579–2605, 2008.

[89] P. Besse, B. Guillouet, J.-M. Loubes, and F. Royer, "Review & perspective for distance based clustering of vehicle trajectories," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 11, pp. 3306–3317, 2016.

[90] Merriam-Webster, *anomaly*, https://www.merriam-webster.com/dictionary/anomaly.

[91] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, no. 3, Jul. 2009.

[92] D. B. Araya, K. Grolinger, H. F. ElYamany, M. A. M. Capretz, and G. T. Bitsuamlak, "Collective contextual anomaly detection framework for smart buildings," *2016 International Joint Conference on Neural Networks (IJCNN)*, pp. 511–518, 2016.

[93] R. Deshmukh and I. Hwang, "Anomaly detection using temporal logic based learning for terminal airspace operations," in *AIAA SciTech 2019 SciTech Forum*, AIAA, 2019, pp. 1–11.

[94] T. G. Puranik and D. N. Mavris, "Identifying instantaneous anomalies in general aviation operations," in *AIAA Aviation Forum*, AIAA, Denver, Colorado, United States, 2017, pp. 1–15.

[95] T. G. Puranik and D. N. Mavris, "Anomaly detection in general aviation operations using energy metrics and flight-data records," *Journal of Aerospace Information Systems*, vol. 15, no. 1, pp. 22–35, 2018.

[96] MathWorks, *Unsupervised Learning*, https://www.mathworks.com/discovery/unsupervised-learning.html.

[97] G. Jarry, D. Delahaye, F. Nicol, and E. Feron, "Aircraft atypical approach detection using functional principal component analysis," *Journal of Air Transport Management*, vol. 84, 2020.

[98] T. Puranik, H. Jimenez, and D. Mavris, "Energy-based metrics for safety analysis of general aviation operations," *Journal of Aircraft*, vol. 54, no. 6, pp. 2285–2297, 2017.

[99] S. J. Corrado, T. G. Puranik, O. P. Fischer, and D. N. Mavris, "A clustering-based quantitative analysis of the interdependent relationship between spatial and energy anomalies in ads-b trajectory data," *Transportation Research Part C: Emerging Technologies*, vol. 131, p. 103 331, 2021.

[100] S. Das, B. L. Matthews, A. N. Srivastava, and N. C. Oza, "Multiple kernel learning for heterogeneous anomaly detection: Algorithm and aviation safety case study," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '10, New York, NY, USA: Association for Computing Machinery, 2010, pp. 47–56.

[101] L. Li, M. Gariel, R. J. Hansman, and R. Palacios, "Anomaly detection in onboard-recorded flight data using cluster analysis," in *2011 IEEE/AIAA 30th Digital Avionics Systems Conference*, 2011, 4A4-1-4A4–11.

[102] L. Li, S. Das, J. R. Hansman, R. Palacios, and A. N. Srivastava, "Analysis of flight data using clustering techniques for detecting abnormal operations," *Journal of Aerospace Information Systems*, vol. 12, no. 9, pp. 587–598, Jun. 2015.

[103] T. G. Puranik, H. Jimenez, and D. N. Mavris, "Utilizing energy metrics and clustering techniques to identify anomalous general aviation operations," in *AIAA SciTech Forum*, AIAA, Grapevine, Texas, United States, 2017, pp. 1–19.

[104] K. Kim and I. Hwang, "Terminal airspace anomaly detection using temporal logic learning," in *International Conference on Research in Air Traffic 2018*, ICRAT, 2018, pp. 1–4.

[105] G. Jarry, D. Delahaye, and E. Feron, "Trajectory approach analysis: A post-operational aircraft approach analysis tool," in *9th SESAR Innovation Days*, Athens, Greexe, 2019, pp. 1–8.

[106] S. Corrado, T. Puranik, O. Pinon-Fischer, D. Mavris, R. Rose, J. Williams, and R. Heidary, "Deep autoencoder for anomaly detection in terminal airspace operations," in *AIAA Aviation 2021*, 2021.

[107] Federal Aviation Administration, *FAA Operations & Performance Data*, https://aspm.faa.gov/aspmhelp/index/Main_Page.html.

[108] S. R. Proud, "Go-around detection using crowd-sourced ads-b position data," *Aerospace*, vol. 7, no. 16, 2019.

[109] G. DiGravio, M. Mancini, R. Patriarca, and F. Costantino, "Overall safety performance of air traffic management system: Indicators and analysis," *Journal of Air Transport Management*, vol. 44-45, pp. 65–69, 2015.

[110] G. DiGravio, M. Mancini, R. Patriarca, and F. Costantino, "Overall safety performance of air traffic management system: Forecasting and monitoring," *Safety Science*, vol. 72, 2015.

[111] M. Kovacova, A. Licu, and T. Linter, "Aerospace performance factor and its potential advances," in *LOGI 2018*, 2018.

[112] F. Netjasov and D. Crnogorac, "Assessment of safety performance indicators of future air traffic management system," in *XLIV Symposium on Operational Research*, 2019.

[113] E. Mangortey, T. G. Puranik, O. J. Pinon, and D. N. Mavris, "Classification, analysis, and prediction of the daily operations of airports using machine learning," in *AIAA SciTech 2020 SciTech Forum*, AIAA, Orlando, FL, United States, 2020.

[114] M. C. R. Murça, "Identification and prediction of urban airspace availability for emerging air mobility operations," *Transportation Research Part C: Emerging Technologies*, vol. 131, 2021.

[115] H. Zhong, H. Liu, and G. Qi, "Analysis of terminal area airspace operation status based on trajectory characteristic point clustering," *IEEE Access*, vol. 9, pp. 16 642–16 648, 2021.

[116] D. J. Berdt and J. Clifford, "Using dynamic time warping to find patterns in time series," in *3rd International Conference on Knowledge Discovery and Data Mining*, 1994.

[117] M. Vlachos, G. Kollios, and D. Gunopulos, "Discovering similar multidimensional trajectories," in *18th International Conference on Data Engineering*, Aug. 2002.

[118] L. Chen, M. T. Özsu, and V. Oria, "Robust and fast similarity search for moving object trajectories," in *SIGMOD '05*, 2005.

[119] F. Hausdoff, "Grundz uge der mengenlehre," 1914.

[120] M. M. Fretchet, "Sur quelques points du calcul fonctionnel," *Rendiconti del Circolo Matematico di Palermo*, vol. 22, no. 1, pp. 1–72, 1906.

[121] scikit-learn documentation, *Sklearn.cluster.dbscan*, https://scikit-learn.org/stable/modules/generated/sklearn.cluster.DBSCAN.htm.

[122] A. Fernandez, D. Martinez, P. Hernandez, S. Cristobal, F. Schwaiger, and J. M. Nunez, "Flight data monitoring (fdm) unknown hazards detection during approach phase using clustering techniques and autoencoders," in *SESAR Innovation Days 2019*, Athens, Greece, 2019, pp. 1–8.

[123] C. R. Patlolla, *Understanding the concept of hierarchical clustering technique*, https://towardsdatascience.com/understanding-the-concept-of-hierarchical-clustering-technique-c6e8243758ec.

[124] L. McInnes, J. Healy, and S. Astels, "Hdbscan: Hierarchical density based clustering," *The Journal of Open Source Software*, vol. 2, no. 11, Mar. 2017.

[125] S. G. Kumar, S. J. Corrado, T. G. Puranik, and D. N. Mavris, "Classification and analysis of go-arounds in commercial aviation using ads-b data," *Aerospace*, vol. 8, no. 10, 2021.

[126] B. Figuet, R. Monstein, M. Waltert, and S. Barry, "Predicting airplane go-arounds using machine learning and open-source data," in *Proceedings of the 8th OpenSky Symposium 2020*, Nov. 2020, pp. 1–12.

[127] J. Bro, "Fdm machine learning: An investigation into the utility of neural networks as a predictive analytic tool for go around decision making," *Journal of Applied Sciences and Arts*, vol. 1, 2017.

[128] S. Corrado, T. Puranik, O. Pinon, and D. Mavris, "Trajectory clustering within the terminal airspace utilizing a weighted distance function," in *Proceedings of the 8th OpenSky Symposium 2020*, Nov. 2020, pp. 1–10.

[129] R. Deshmukh, D. Sun, and I. Hwang, "Data-driven precursor detection algorithm for terminal airspace operations," in *Thirteenth USA/Europe Air Traffic Management Research and Development Seminar*, 2019, pp. 1–7.

[130] B. L. Welch, "The generalization of 'student's' problem when several different population variances are involved," *Biometrika*, vol. 34, no. 1/2, pp. 28–35, 1947.

[131] L. McInnes, J. Healy, and J. Melville, *Umap: Uniform manifold approximation and projection for dimension reduction*, 2020. arXiv: 1802.03426 [stat.ML].

[132] T. Chen and C. Guestrin, "Xgboost," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2016.

[133] *(KSFO) San Francisco International Airport*, https://www.aopa.org/destinations/airports/KSFO/details.

[134] *SFO Flight Patterns and Operations*, https://www.flysfo.com/community/noise/sfo-flight-patterns-and-operations.

[135] M. Schäfer, M. Strohmeier, V. Lenders, I. Martinovic, and M. Wilhelm, "Bringing up opensky: A large-scale ads-b sensor network for research," in *13th IEEE/ACM International Symposium on Information Processing in Sensor Networks (IPNS)*, 2014, pp. 38–94.

[136] M. Strohmeier, X. Olive, J. Lübbe, M. Schäfer, and V. Lenders, "Crowdsourced air traffic data from the opensky network 2019–2020," *Earth System Science Data*, vol. 13, no. 2, pp. 357–366, 2021.

[137] *OpenSky Network*, https://opensky-network.org/.

[138] A. Academy, *Heading, Track, Bearing, and Course Explained*, https://airplaneacademy.com/heading-track-bearing-and-course-explained/.

[139] X. Olive, "Traffic, a toolbox for processing and analysing air traffic data," *Journal of Open Source Software*, no. 7553, p. 1518, 2019.

[140] Iowa State University, *Iowa Environmental Mesonet*, https://mesonet.agron.iastate.edu/request/download.phtml, 2021.

[141] *Map Projection*, https://www.usna.edu/Users/oceano/pguth/md_help/html/mapb0iem.htm.

[142] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[143] T. Blajev and W. Curtis, *Go-Around Decision-Making and Execution Project*. Mar. 2017.

[144] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[145] L. McInnes, J. Healy, N. Saul, and L. Grossberger, "Umap: Uniform manifold approximation and projection," *The Journal of Open Source Software*, vol. 3, no. 29, p. 861, 2018.