

Initial Performance Evaluation of Flight Path Management Onboard Automation

Terique L. Barney

Crew Systems and Aviation Operations Branch
NASA Langley Research Center Hampton, VA, USA
terique.l.barney@nasa.gov

Bryan A. Barrows

Autonomous Integrated Systems Research Branch
NASA Langley Research Center Hampton, VA, USA
bryan.barrows@nasa.gov

Abstract—Significant developments in automation are necessary to achieve safe and efficient operations in advanced aerial mobility related concepts. Urban Air Mobility (UAM) is a rapidly growing, emerging field that poses a challenging use-case with a tighter scale of operations compared to the traditional commercial transport paradigm. A large part of the challenge is the uncharted territory; as of this paper, no set of operational standards or guidelines for UAM operations have been established and automated en route operations for UAM level 4 (UML-4) have not been studied. Flight Path Management (FPM) automation provides a set of capabilities that are critical toward enabling airborne vehicles to achieve mission success while maintaining operational safety. An initial performance evaluation of FPM automation was conducted using a UAM-adapted version of the Autonomous Operations Planner (AOP), an onboard trajectory management capability developed over years of research targeting commercial transport operations, as its reference implementation. This paper describes the evaluation, including the approach and methodology for simulating FPM automation in UML-4, key results, conclusions, and future work.

Keywords—*flight path management, simulation, automation, urban air mobility, advanced air mobility, separation assurance, schedule contingency, UML-4*

I. INTRODUCTION

Advanced aerial mobility (AAM) [1] is a vision that enables “safe, sustainable, affordable, and accessible aviation for transformational local and intraregional missions.” These missions include applications for which the Urban Air Mobility (UAM) [2] concept poses a challenging use-case and opportunity to apply state-of-the-art automation technology toward an emerging operational concept. Airspace operations in the UAM operational concept are of a tighter scale than traditional commercial transport operations in Federal Aviation Administration (FAA) Code of Federal Regulations (CFR) Part 121 [3], and therefore require technical capabilities that enable the comparatively higher complexity, faster paced, and operationally demanding environment. No set of operational standards or guidelines for UAM have been established yet and automated en route operations for UML-4 have not been researched in depth through modeling and simulation.

The National Aeronautics and Space Administration (NASA) UAM Vision Concept of Operations (ConOps) describes factors impacting the evolution of UAM vehicles, airspace operations, and the surrounding community expected to mature through a series of UAM maturity levels (UMLs) [4]. UML-4 is characterized as “medium density and complexity operations with collaborative and responsible automated systems” [5]. Flight Path Management (FPM) automation technology [6] – defined as a suite of functions that create, monitor, and revise the flight path in response to evolving mission objectives, constraints, and external stimuli – is among the capabilities considered necessary for transitioning toward UML-4 and enabling higher density/complexity AAM operations. FPM ensures that the vehicle flight path meets five objectives: it is feasible to fly to mission completion, is deconflicted from airspace hazards, is coordinated with other traffic, is flexible to accommodate future disturbances, and is optimized to meet business objectives [7]. Over 15 years of strategic airborne trajectory management research targeting commercial transport applications led toward the development of the Autonomous Operations Planner (AOP), an onboard trajectory management capability that provides four-dimensional trajectory (4-DT) resolutions for conflicts on the active route while considering airspace system-level constraints [8]. A version of AOP with functionality that targets the UML-4 operational environment was used as the reference implementation of FPM automation during an initial performance evaluation of an onboard configuration of FPM automation. An overview of the AOP functionality applied in the evaluation will be provided.

The initial performance evaluation of FPM automation, named “FPM-1,” included a batch simulation of UML-4 operations. Objectives of this evaluation included understanding the cause-effect relationships between FPM automation performance and UAM operational safety and mission success; identification of factors impacting FPM automation performance; and baselining performance metrics and use-case scenarios for future research activities. Functional performance of FPM in the UML-4 use-case was measured by achievement of mission success while maintaining operational safety. Like the emerging field of UAM, the simulation itself is in its infant stages; it is a work in progress that yielded insights into the feasibility and suitability of UML-4 implementation

details that previously did not exist for the UAM use-case, including applicability of candidate operational standards and assumptions for FPM in UML-4. The evaluation provided a foundation for addressing challenges of FPM automation application in the UML-4 use-case as the UAM concept rapidly develops. This document describes the initial performance evaluation of FPM automation in UML-4, including FPM's role in the UML-4 operational context, an overview of the FPM automation reference system, and the FPM nominal simulation use-case. An overview of the batch simulation, including scenario design, experiment design, and results are also discussed. The document ends with discussion of future related work and concluding thoughts for the performance evaluation.

II. UML-4 AIRBORNE OPERATIONS

UML-4 operations include hundreds of simultaneous operations across expanded networks of closely spaced, high throughput aerodromes in a single metropolitan area. Aircraft operations are assumed to be conducted in non-visual conditions, higher density than can be accommodated by air traffic control (ATC) and dedicated corridor operations, and simplified (e.g., point-to-point). Providers of service to UAM (PSU) will be available as well as automated systems requiring no human oversight. At UML-4, operators are assumed to be directly responsible for traffic separation, and their use of airborne FPM automation to detect and resolve conflicts is the mechanism by which it is done. Vehicle-to-vehicle cooperative surveillance is assumed, and trajectory intent information on all vehicles is assumed to be available to the automation, potentially relayed automatically through the PSU. These aspects of UML-4 highlight the need for an enabling FPM automation system.

Knowledge accumulated over more than two decades of operations research on the commercial transport application is referenced to start addressing UML-4. However, part of the challenge of transitioning applications are the gaps between commercial transport and UML-4 operational paradigms. Table I provides a summary of these differences and some of the values applied in the UML-4 simulation environment supporting the evaluation of FPM automation. The impacts of some of the differences between the operational paradigms are candidates for exploration in UML-4 research.

The operational concepts of commercial transport and UML-4 have several fundamental differences. In commercial transport operations, flight durations are on the order of hours, 4-DT compliance is typically only required at controlled times of arrival such as waypoints and is on the order of tens of seconds, and large wind prediction errors impact trajectories considerably. Conversely, in UML-4, flight durations are on the order of tens of minutes, 4-DT compliance is potentially continuous (with required precision of less than 10 seconds) and large wind prediction errors may impact trajectories significantly. More granular and frequent wind prediction updates may be required to maintain continuous 4-DT compliance in UML-4. Additionally, UAM operations will eventually compete with mixed operations that commercial transport cruise operations typically do not need to consider, such as general aviation aircraft operation and operations that

TABLE I. DIFFERENCES BETWEEN OPERATIONAL PARADIGMS

Traditional Commercial Transport Operations	UML-4
Concept of Operations	
Longer flight durations, intermittent 4-DT compliance, low to moderate impact from wind prediction error.	Shorter flight durations, continuous 4-DT compliance, high impact from wind prediction error.
Separation Minima	
Class A Airspace: 5 nm laterally, 1000 ft vertically	Assumed minima: 0.25 nm laterally, 250 ft vertically
Conflict Detection and Resolution	
Longer strategic look-ahead time horizons and time constraints. Strategic and tactical CD&R functions are well separated in time.	Shorter strategic look-ahead time horizons and time constraints. Strategic, tactical and hazard avoidance CD&R functions not well separated in time. May cause interoperability challenges.
Vehicles	
Jet transports Jet propulsion – slow control response. Revenue operations value high flight efficiency.	Future UAM vehicles VTOL capability, fast control response. High sensitivity to energy use, potentially requiring reduced margins for contingencies
Aircraft speeds and altitudes	
Mach 0.75 to 0.85, 10,000 to 40,000 ft MSL	0 to 120 kts, 0 to 2000 ft AGL

require constant air traffic control involvement (e.g., departures and arrivals at major airports near aerodromes).

Separation standards for UML-4 have yet to be established, but differences between these standards and criteria for near mid-air collision (NMAC) are assumed to be considerable smaller than standards used by ATC to enable high traffic density operations. Approaches to conflict detection and resolution (CD&R) are also on different levels: for commercial transport operations, strategic look-ahead horizons are on the order of 10 minutes and time constraints are on the order of minutes, but for UML-4, strategic look-ahead horizons are on the order of just a few minutes and time constraints are on the order of tens of seconds (or seconds). Additionally, transitions between strategic deconfliction, detect and avoid, and collision avoidance applications are assumed to be more fluid in UML-4 compared to commercial transport operations, potentially creating interoperability challenges that warrant further examination.

Aircraft in UML-4 operate within significantly lower altitude and speed ranges, but with a wider range of speed changes in normal operations that accommodates a more dynamic operating environment. Compared to commercial transport vehicles, UAM vehicles have increased maneuverability and faster control responses to avoid ground obstacles, resolve conflicts, comply with required time of arrival (RTA) changes received en route, and adjust to wind gusts and adverse weather. UAM vehicles such as electric vertical takeoff and landing (e-VTOL) model types also have a higher sensitivity to energy use, potentially requiring reduced margins for contingencies.

A subset of the challenges of transitioning to UML-4 were examined as scenario conditions during the evaluation of FPM automation in UML-4 batch simulation, including differences in wind error, traffic density, and RTA delays issued en route. Additional conditions potentially impacting FPM automation

performance, such as levels of FPM capability and limited maneuver degrees of freedom in strategic resolutions, were also examined.

III. OVERVIEW OF FLIGHT PATH MANAGEMENT AUTOMATION SYSTEM

A. FPM automation

FPM ensures that a safe and operationally acceptable flight path for a vehicle is available to the vehicle operator and flight control system for the life of the flight. A fully capable FPM automation system does this by creating a flight path that meets all five objectives introduced in section I (feasible, deconflicted, coordinated, flexible, and optimal), monitors it continuously, and revises it automatically in response to evolving mission objectives, system constraints, and other stimuli. In this study, a separate instance of FPM automation is hosted onboard each aircraft in the simulation and focused exclusively on ensuring the ownship meets the goals of FPM.

FPM automation generates a 4-DT flight path, the associated operational intent, and energy consumption profile in response to changing mission circumstances. It also provides users with contextual information, flight conformance status, contingency response options, and system health information. Live inputs to the FPM automation system enable the generated flight path to account for mission objectives, pilot preferences, ownship performance capabilities, ownship and traffic state and intent information, current and predicted atmospheric conditions, relevant operational rules, obstacles and terrain constraints, and airspace configuration.

An FPM automation system hosted onboard the ownship benefits from direct access to onboard real-time data, the shortest loop-closure time, the fleet operator's proprietary business goals (via the pilot) and the vehicle's proprietary performance model. These attributes of computed flight paths result in less conservative planning, reduced airspace buffer requirements, tighter separation standards, and improved responsiveness, all of which contribute to increased system capacity and efficiency. This configuration of the FPM automation system is capable of computing flight paths with greater accuracy and tighter conformance than a system shared or used by a third party without access to proprietary data.

B. Reference implementation of FPM automation in UML-4

AOP is a research-prototype automation system "designed to enable vehicle operational autonomy within the research concept of Autonomous Flight Rules (AFR)" [9]. The original design approach for AOP was to develop an integrated suite of strategic and tactical airborne trajectory management functionality compatible with commercial transport avionics architectures and industry standards. Development of AOP over years of research led toward a suite of functions not only applicable to commercial transport operational objectives, but capable of meeting FPM objectives in the UAM use-case as well [10].

The functions selected for FPM-1 included four-dimensional trajectory prediction; flight path change evaluation; automated flight path revision; conflict detection,

prevention, and resolution; coordination mechanisms; constraint conformance, contingency response generation; energy consumption prediction, and energy reserves monitoring. The lack of a reference case for automated en route operations in UML-4 introduced an additional challenge: proper adjustment of AOP function parameters and simulation settings to match the UML-4 use-case application. For FPM-1, tuning of these settings was focused on enabling operational feasibility, safety, and mission completion given the operational assumptions and scope selected for the implementation. FPM-1 therefore provided an additional benefit of generating baseline data for FPM in UML-4 operations.

IV. APPROACH TO PERFORMANCE EVALUATION

A. Overview of UAM-Simulation Environment

NASA's airspace & traffic operations simulation (ATOS) capability was developed to study the feasibility of distributed air/ground traffic management concept elements in en route and terminal airspace [11], [12]. ATOS provides a robust, mid-fidelity, real-time distributed simulation of aircraft, their communication, navigation, and surveillance environment, and airspace conditions such as traffic, atmospheric conditions, and schedule requirement updates. ATOS components relevant to FPM-1 include a simulation manager, aircraft simulators for traffic operations research (ASTOR), a traffic generator, scenario generation tool, and high-level architecture. ATOS is capable of simulating hundreds of aircraft using multiple simulation platforms and various aircraft performance models. Each ASTOR may be equipped with research decision support tools such as AOP. ATOS has been developed and extended for many years at NASA Langley [13] to support experiments of multiple batch and piloted simulation studies [14] that allow research of advanced distributed air-ground concepts and airborne decision support tools. The core capabilities of ATOS were adapted in recent years to create the UAM simulation environment (UAM-Sim).

UAM vehicle simulators required several changes to support UAM operations. Trajectory profiles required new modeling in the simulators and in CD&R algorithms to include vertical, transition and wing-borne segments for e-VTOL models. This includes unique capabilities like segments for climbing while accelerating and descending while decelerating, which typically could be safely ignored in ATOS simulations due to the performance limitations of commercial transport. Additionally, although UAM en route operations are the primary focus of this paper, UAM simulation includes a terminal phase for takeoff, landing, and eventually pre-departure that will require additional attention in future research to properly model. UAM-Sim also included two e-VTOL performance model options, lift + cruise and quad-rotor vehicles, that had differing performance envelopes. The capability of UAM-Sim vehicles receiving RTA changes (delays) mid-flight was also simulated as widespread. During FPM-1, all UAM-Sim vehicles were equipped with the onboard configuration of FPM automation (AOP).

B. Scenario design and simulation runs

Each simulation scenario used throughout the FPM-1 batch simulation represent medium-density and complexity UML-4 operations across approximately 1,800 square nautical miles of the Dallas/Fort Worth (DFW) metropolitan. Between 150 and 200 simultaneous operations across 55 vertiports were simulated. Scenario runs gradually increased traffic counts from 0 to a minimum of 150 simultaneous flights, maintained the traffic volume for at least 30 minutes, then decreased traffic volume back to 0 by ceasing departures and allowing remaining airborne vehicles to land at their destination. Each UAM vehicle was represented by one of two e-VTOL aircraft performance models equipped with AOP. Operations were all below a 3,000-foot mean sea level (MSL) altitude ceiling, and vertiports were assigned the MSL altitudes of their approximate real-world above ground level (AGL) altitudes at their coordinate location, to approximate altitude differences along changing terrain between vertiports. Vehicles were to maintain a 0.25 nautical mile radius of lateral separation, and 250-foot vertical separation, from other traffic – values based on terminal area separation distances for detect and avoid (DAA) well-clear volumes. This provided a proxy for simulated operations in the absence of established UAM separation standards. Static, three-dimensional wind grids between altitudes of 0 and 3,000-foot MSL, and 0 to 10 knots, were also included.

During scenario generation, each vehicle was assigned flight plans with direct routing between a randomly selected pair of departure and arrival vertiports based on traffic demand. If an intruding airspace restriction, referred to here as air traffic management (ATM) environment zones, was present along the point-to-point routing, the flight plan was instead generated using non-direct routing around the airspace restriction. ATM environment zones represent areas that are likely to require UAM vehicle interaction with ATC prior to traversal [15], which is assumed to be an operational limitation in UML-4 and an additional burden to ATC. ATM environment zones spanned from 0 MSL to infinite altitude ceilings in simulation. The route structure implemented for navigation around and between ATM environment zones included two-way routing near DFW airport (see rectangular box in Fig. 1), where higher traffic flow volumes and complexity is expected, as well as a handful of narrow passages that loosely emulated the maneuver limitation effects of corridor operations (see red oval in Fig. 1).

Fig. 1 is a snapshot from the batch simulation. Included in the image are the flown trajectories of over 150 eVTOL vehicles (blue airplane-shaped icons, with yellow flight trails) simultaneously flying between vertiport locations (red and black bubble icons) in the DFW area. Also included in the image are ATM environment zones (yellow polygons), which all flights avoided while navigating between departure and arrival vertiports, maintaining traffic separation, and complying with schedule.

Vehicles other than e-VTOL model types, such as general aviation vehicles and commercial transports, were excluded from FPM-1. Additionally, as a batch simulation, FPM-1 did not include any Human-in-the-Loop operations. Each trajectory resolution generated by AOP was automatically accepted and

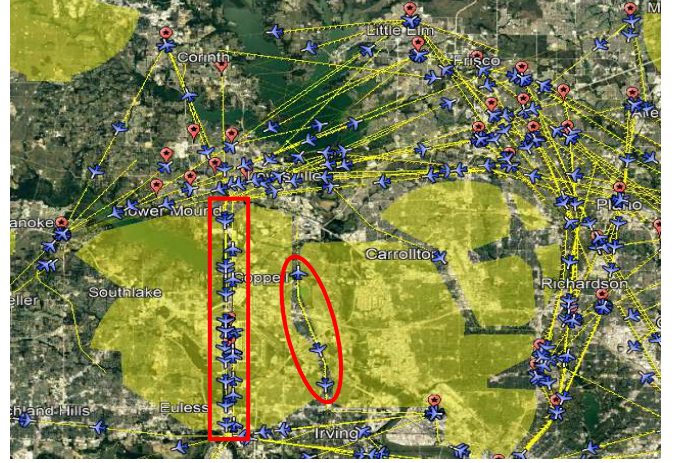


Fig. 1 Snapshot of UML-4 operations simulation in DFW region

executed by the vehicle flight control system in the UAM models. Specific DFW metropolitan region constraints and rules were not simulated.

C. Normal FPM operations in UML-4 simulation

Normal flight operations in FPM-1 simulations focused on the feasibility, deconfliction and coordination objectives for FPM. At scenario launch, each UAM vehicle took off at its scheduled time and proceeded on its assigned flight plan to the destination vertiport. AOP onboard each vehicle was granted the ability to alter the current route during the flight whenever an updated 4-DT was generated to resolve a conflict, meet an RTA delay constraint, or both. In lieu of a pilot in command for the batch simulation, the UAM vehicle flight control system automatically uploaded and executed alternative routes provided by AOP.

Several operational assumptions and simulation conditions were applied to FPM-1. No preflight deconfliction services or en route deconfliction services besides FPM automation were included in the simulation. No two-way interchanges with a PSU or ATC were simulated. Traffic state data was always available to vehicles. Vertiport/vertipad capacity limits were not modeled; all flights were programmed to land at their original destination with no redirection. Among other typical airspace features not modeled in FPM-1 were other ATM technologies external to AOP, vertiport and terminal area operations, communications infrastructure (e.g., network latencies, interference, and range effects), dynamic convective weather, special use airspace, and temporary flight restrictions. The effects of a subset of presumed PSU services (i.e., transmission of 4-D traffic intent information and vertiport RTA changes to airborne vehicles) were emulated but not explicitly modeled.

D. Experiment Design

FPM performance in FPM-1 was defined as a measure of FPM automation's ability to meet mission objectives while maintaining operational safety. This was evaluated through measures of mission success and safety achievement. The study sought to understand the impacts of different flight path deconfliction approaches, vehicle maneuver options, and

schedule contingencies on FPM performance in UAM operations. Five parameters were identified as independent variables (IVs) manipulated across the FPM-1 experiment matrix: traffic demand (2 levels), wind error (2 levels), RTA delay impact (2 levels), CD&R capability (3 levels), and maneuver degrees of freedom available to FPM for execution (3 levels).

Table II shows the eight different types of scenario configurations used in the experiment. These scenario configurations were embedded in the nine different test cases shown in Table III, each test case varying in FPM CD&R capability and maneuver option set, for a total of 72 tests in a 2x2x2x3x3 full factorial matrix.

TABLE II. SCENARIO CONFIGURATION COMBINATIONS

Scenario Configuration	RTA Delay Impact	Traffic Demand	Wind Error
1	Mild	Peak	Error
2			No Error
3		Lull	Error
4			No Error
5	Moderate to Significant	Peak	Error
6			No Error
7		Lull	Error
8			No Error

TABLE III. TEST MATRIX FOR FPM INITIAL PERFORMANCE EVALUATION

Test Matrix		CD&R Capability		
		Full Capability	No Trajectory Buffers	No Traffic Intent
Maneuver Option Sets Available	Full	Scenarios 1-8	Scenarios 1-8	Scenarios 1-8
	Lateral Only	Scenarios 1-8	Scenarios 1-8	Scenarios 1-8
	Vertical Only	Scenarios 1-8	Scenarios 1-8	Scenarios 1-8

The traffic demand and wind error IVs enable observation of potential impacts of different environmental conditions on AOP performance. Traffic demand is representative of average flows between vertiports during “peak” and “lull” demand segments of a typical traffic day [16]. Wind error indicates the existence of a difference between predicted (forecasted) wind data utilized by FPM automation and truth (actual) winds experienced by the simulated aircraft. The RTA delay impact and maneuver options IVs enable observation of the impact of system-level constraints on AOP performance. RTA delay impact is implemented via three components for each flight: probability of receiving an RTA change/delay (which varied by vertiport), magnitude of delay, and period of flight within which a delay may be issued. For the flights that received a delay, RTA delay was observed at two levels of expected impact to AOP: “mild” delay (20% of time remaining in the flight, issued anytime between 2 minutes after takeoff and 8 minutes before the estimated time of arrival), and “moderate to significant” delay (35% to 40% of time remaining in the flight, issued anytime between 4 minutes after takeoff and 6 minutes before the estimated time of arrival). The maneuver options IV setting limited FPM automation’s CD&R capabilities and RTA conformance resolution capabilities according to one of three levels of degrees of freedom within the maneuver resolution option sets: lateral maneuvers only; vertical maneuvers only; and no maneuver limitations. Hybrid resolutions that have

unique lateral and vertical maneuver components combined were only available in the “no maneuver limitations” level. The final IV, CD&R capability, enabled various features of CD&R functionality to be available to the vehicle. These included trajectory prediction uncertainty bounds (TPUBs – essentially extra buffering to the separation criteria) [17], availability of traffic intent input data, and whether conflict resolution was enabled.

Impact on FPM performance was measured by several metrics. This paper discusses two of those key metrics: total losses of separation (LOS), and conformance to RTA. These two metrics were used as proxies for operational safety and mission success, respectively.

E. Data Collection

For all simulation scenarios, data collection was performed across a virtual central processing unit (CPU) cluster that distributes dedicated scenario runtime and processing of simulation vehicles. Subsets of UAM vehicle simulators were randomly assigned a CPU machine on the cluster during scenario initialization. Each UAM vehicle’s performance flight data and scenario event logging of simulation data was collected locally during runtime. The data within the local repositories of each CPU machine was aggregated into a final data repository upon completion of each scenario run. Summary files containing AOP recorded ownship data across the entire set of scenarios, including descriptive flight characteristics and details for conflict occurrences, were generated. These data summaries enabled key metrics to be analyzed from one or two output source files. Criteria for identifying simulation errors was defined *a priori* to data exploration and analysis as a necessary measure for ensuring data quality, then used to clean flight data stored within all summary files of outliers.

V. RESULTS

Two key metrics for FPM-1, LOS count and RTA conformance, were explored as dependent variables (DVs). LOS count incremented with the number of “significant” LOS events each flight experienced, characterized as operational errors, i.e., significant penetrations of separation volumes. Thresholds for counting LOS events were configured to penetrations exceeding five percent of the lateral separation standard, and at least 20% of the vertical separation standard (see Table I for separation values). RTA conformance was measured for each flight as the actual arrival time to the destination (or AOP’s estimated arrival time if the flight was still operating at scenario termination) minus the last RTA assignment received, in seconds.

Analysis of the experimental data began with DV summary statistics and exploration of distributions across the full matrix. An aggregated dataset of the 72 scenario runs contained 50,963 unique observations (flights) at the completion of outlier removal during data post-processing. Pivot tables were generated for LOS count and RTA conformance results. LOS count results are presented as percentages of flights across the dataset that had at least one LOS. The LOS count data fit a Poisson distribution, therefore a generalized linear model (GLM) Poisson regression test with a logarithmic link was

applied. The model only included main effect and second-degree (2-term) interaction effects. RTA conformance pivot table data are presented as mean and median values in seconds, calculated for flights that did not receive an RTA delay assignment separately from flights that did. Further investigation of RTA conformance and alternative testing is discussed in the future work section.

A. LOS Count

Recall that one scenario batch run was executed for each cell in the 72-cell matrix. The data in Table IV represent the percentages of flights that experienced at least one LOS out of all flights in the run/cell indicated in the table. Each cell is also on a color gradient that indicates which scenarios held the lowest to the highest percentages across the matrix. The lowest and highest percentages of LOSes are indicated by white and dark red coloring, respectively.

TABLE IV. LOS COUNT PIVOT TABLE

% of Vehicles with One or More LOSes	Column Labels			
	Lull		Peak	
Row Labels	No Wind Error	Wind Error	No Wind Error	Wind Error
Mild RTA Delay Impact				
	Full Maneuver set			
Full CDR	3.79%	1.20%	9.91%	9.80%
No Trajectory Buffers	2.79%	1.40%	5.90%	6.52%
No Traffic Intent	7.58%	7.34%	18.85%	18.46%
	Lateral Maneuvers only			
Full CDR	2.45%	1.84%	4.72%	4.42%
No Trajectory Buffers	2.02%	2.42%	3.77%	4.78%
No Traffic Intent	7.38%	8.69%	17.54%	18.82%
	Vertical Maneuvers only			
Full CDR	3.64%	4.67%	7.52%	8.33%
No Trajectory Buffers	3.42%	3.21%	7.73%	7.29%
No Traffic Intent	6.96%	7.34%	18.95%	17.82%
Moderate to Significant RTA Delay Impact				
	Full Maneuver set			
Full CDR	2.01%	3.23%	10.19%	9.99%
No Trajectory Buffers	2.00%	0.80%	9.76%	10.11%
No Traffic Intent	11.35%	11.63%	23.02%	20.99%
	Lateral Maneuvers only			
Full CDR	4.65%	3.46%	5.89%	5.93%
No Trajectory Buffers	2.21%	1.62%	5.91%	4.86%
No Traffic Intent	10.97%	10.04%	21.86%	22.81%
	Vertical Maneuvers only			
Full CDR	4.46%	4.25%	11.91%	12.49%
No Trajectory Buffers	4.44%	3.64%	10.92%	11.03%
No Traffic Intent	10.95%	12.79%	23.46%	23.50%

Two main trends can be observed. Scenarios with “peak” traffic demands had higher percentages of flights that experienced LOS events than those with “lull” demand, as expected due to the difference in traffic density. Additionally, the percentage of vehicles in scenarios that have no traffic intent data available experience at least double – and in many cases, at least triple - the amount of LOS events compared to the other two capability levels that include traffic intent data. This implies that FPM automation access to traffic intent data

might be significant in maintaining adequate FPM performance and minimizing LOS events. The table also indicates a subtle difference in LOS counts between RTA delay impact levels, but it is not as pronounced as the first two IVs. There appears to be very little impact by the presence of wind error in the simulation.

The Poisson GLM was used to estimate cause-and-effect relationships between each IV and the LOS count metric more directly than the pivot table. The GLM also included flights that had no LOS events in the analysis, whereas Table IV only observed flights with at least one LOS. For the statistical significance level, alpha, the standard value of 0.05 was used. After the test was run, insignificant main and interaction effects were incrementally removed from the effect summary result tables until only significant effects remained. This regression test was run three different ways: once across all flights in the experiment matrix, once across all flights that never received an RTA delay assignment (~78.2% of all flights), and once across all flights that received an RTA delay assignment (~21.8% of all flights). The latter two tests were done to observe any influence that RTA delays may have had on the effects impacting LOS count. The results for these tests are provided in Tables V, VI and VII.

TABLE V. LOS COUNT REGRESSION EFFECT SUMMARY – ALL FLIGHTS

Source	Log Worth	P Value
CD&R Capability	250.272	<.0001
Traffic Demand	138.79	<.0001
RTA Delay Impact	25.078	<.0001
Maneuvers	19.817	<.0001
Maneuvers * CD&R Capability	17.518	<.0001
Traffic Demand * CD&R Capability	3.597	0.00025
Traffic Demand * Maneuvers	1.881	0.01316

TABLE VI. LOS COUNT REGRESSION EFFECT SUMMARY – NO RTA DELAY (39,867 FLIGHTS)

Source	Log Worth	P Value
CD&R Capability	251.983	<.0001
Traffic Demand	141.711	<.0001
Maneuvers * CD&R Capability	14.455	<.0001
Maneuvers	9.614	<.0001
RTA Delay Impact	9.084	<.0001
Traffic Demand * Maneuvers	2.269	0.00538

TABLE VII. LOS COUNT REGRESSION EFFECT SUMMARY – ONE RTA DELAY (11,096 FLIGHTS)

Source	Log Worth	P Value
CD&R Capability	109.437	<.0001
Traffic Demand	50.204	<.0001
RTA Delay Impact	22.079	<.0001
Maneuvers	12.101	<.0001
Traffic Demand * CD&R Capability	3.38	0.00042
Maneuvers * CD&R Capability	3.366	0.00043
RTA Delay Impact * CD&R Capability	1.814	0.01534

Results across Tables V, VI and VII largely support the previous observations that CD&R capability levels and traffic demand significantly impact LOS count, and that wind error is not a significant factor (p value > 0.05). The RTA delay impact and maneuver options IVs, however, are less consistent; both of these IVs are among the four main effects that have p values extremely close to 0 consistently across Tables V, VI and VII. However, the Log Worth transform value for RTA delay impact - used to transform very small p -values into usable numbers - is smaller for runs with no RTA delay assignment (Table VI) than the other two datasets (Tables V and VII). Higher Log Worth values generally indicate higher dependence relationships. It may be necessary to further investigate the nature of how the presence of RTA delay assignments (rather than the varying magnitudes and assignment periods when issued) impacts LOSes across the airspace system.

B. RTA Conformance

The data in Table VIII represent the mean RTA conformance values, in seconds, across all flights in the run/cell indicated in the table. Each cell is also on a color gradient that indicates which scenarios had mean values closest to zero (white) and which had values furthest away (dark blue is more negative; dark orange is more positive). Negative values indicate early arrival of flights to their destination relative to the final RTA received, and positive values indicate later arrival.

Overall, the largest average non-conformance was less than 2.5 seconds, indicating generally good RTA conformance even at UML-4 density/complexity conditions. Initial observations of Table VIII indicate that RTA delay impact level and maneuver option availability significantly impact RTA conformance. Flights at the significant RTA delay impact level tended to land earlier than flights at the mild RTA delay impact level. Additionally, all of the mean values for flights at the significant RTA delay impact and vertical maneuvers only levels were negative (early). This may suggest a significant interaction effect between these two IV levels in particular, which makes sense in operations such as what is expected in UML-4. The availability of lateral maneuvers and trajectory changes such as flight path stretches are often critical toward generating additional flight time without causing additional conflicts. This may be of particular importance if RTA delays are issued during corridor operations, especially when issued RTA delays are relatively high in magnitude across the airspace. Wind error, traffic demand, and CD&R level appear to have less significant impacts.

Data in Table IX displays similar information as Table VIII, except with the use of median values instead of means. Median values in this table allow us to observe the data without influence from outliers, which likely played a noteworthy role in the establishment of the mean values on Table VIII. The color gradient in Table IX ranges from white (lowest value in the table) to dark purple (highest value in the table). Data in Table IX suggests two considerations worth mentioning: traffic demand may indeed have a larger impact on RTA conformance than Table VIII suggests, and that median values for all flights are within a very tight range (0.66 and 0.79 seconds) for RTA conformance. The former finding makes sense operationally:

TABLE VIII. MEAN RTA CONFORMANCE IN SECONDS, ALL FLIGHTS

Mean RTA Conformance Values	Column Labels			
	Lull		Peak	
Row Labels	No Wind Error	Wind Error	No Wind Error	Wind Error
Mild RTA Delay Impact				
	Full Maneuver set			
Full CDR	0.49	0.51	0.36	0.38
No Trajectory Buffers	0.37	0.38	0.41	0.54
No Traffic Intent	0.61	0.59	0.66	0.50
	Lateral Maneuvers only			
Full CDR	0.63	0.66	0.68	0.70
No Trajectory Buffers	0.59	0.61	0.68	0.71
No Traffic Intent	0.57	0.56	0.50	0.65
	Vertical Maneuvers only			
Full CDR	0.57	0.58	0.65	0.66
No Trajectory Buffers	0.56	0.56	0.54	0.39
No Traffic Intent	0.60	0.57	0.55	0.65
Moderate to Significant RTA Delay Impact				
	Full Maneuver set			
Full CDR	0.35	0.35	-0.40	-0.44
No Trajectory Buffers	0.30	0.31	0.16	-0.37
No Traffic Intent	0.42	0.40	0.36	0.13
	Lateral Maneuvers only			
Full CDR	-0.10	0.47	0.20	0.28
No Trajectory Buffers	0.31	0.38	0.31	0.36
No Traffic Intent	-0.06	0.37	0.28	0.30
	Vertical Maneuvers only			
Full CDR	-1.66	-1.81	-2.24	-2.20
No Trajectory Buffers	-1.83	-1.73	-1.98	-1.86
No Traffic Intent	-1.78	-1.84	-1.54	-1.69

FPM automation is expected to face an increase in difficulty conforming to RTA delays amidst higher traffic densities and are likely to arrive later than planned. The latter finding suggests that the multitude of outliers may need further investigation, considering the large difference between the mean values and median values.

It is possible that the large number of outliers in RTA conformance data contributed to the lack of fit to any known distributions. The data was separated into two subgroups: flights that received no RTA delay (subgroup A) and flights that did (subgroup B), similarly to what was done in Tables VI and VII in order to observe any differences in their distributions. Fig. 2 displays the final RTA conformance values (in seconds) for each of these two subgroups. Several key differences were observed: outliers in subgroup A (39,867 in total) were mostly between -5 (early) and 15 seconds (late), and the handful of flights outside of that range arrived late by up to 38 seconds. Conversely, subgroup B (11,096 in total) had a much wider spread of outliers ranging from 226 seconds early to 12 seconds late, most of which clustered on the side of early arrival.

The implemented ranges of possible RTA magnitudes and assignment periods for the mild and significant RTA delay impact groups potentially caused greater differences in non-conformance between the two subgroups in Fig. 2 than

TABLE IX. MEDIAN RTA CONFORMANCE IN SECONDS, ALL FLIGHTS

Median RTA Conformance Values	Column Labels			
	Lull		Peak	
Row Labels	No Wind Error	Wind Error	No Wind Error	Wind Error
Mild RTA Delay Impact				
	Full Maneuver set			
Full CDR	0.72	0.71	0.76	0.76
No Trajectory Buffers	0.72	0.71	0.76	0.77
No Traffic Intent	0.72	0.72	0.76	0.77
	Lateral Maneuvers only			
Full CDR	0.73	0.73	0.79	0.79
No Trajectory Buffers	0.72	0.73	0.78	0.77
No Traffic Intent	0.72	0.72	0.76	0.77
	Vertical Maneuvers only			
Full CDR	0.72	0.73	0.77	0.77
No Trajectory Buffers	0.72	0.72	0.78	0.77
No Traffic Intent	0.72	0.72	0.77	0.76
Moderate to Significant RTA Delay Impact				
	Full Maneuver set			
Full CDR	0.67	0.67	0.75	0.73
No Trajectory Buffers	0.67	0.66	0.75	0.76
No Traffic Intent	0.70	0.68	0.74	0.76
	Lateral Maneuvers only			
Full CDR	0.66	0.69	0.78	0.76
No Trajectory Buffers	0.68	0.70	0.76	0.76
No Traffic Intent	0.69	0.68	0.74	0.74
	Vertical Maneuvers only			
Full CDR	0.68	0.67	0.76	0.75
No Trajectory Buffers	0.67	0.67	0.76	0.76
No Traffic Intent	0.68	0.69	0.74	0.73

expected in simulation. Therefore, mean values for RTA conformance were reinvestigated in subgroups A and B separately.

As expected, mean results from subgroup A (not shown) displayed very little difference between cells in the experiment matrix. Mean values across all 5 IVs in all 72 cells were between 0.74 and 0.95 seconds. When focusing on minimum values in the pivot table, most values were between -5 and 0 seconds. Flights with only vertical maneuvers in the “lull” scenarios all landed between -1 and 0 seconds. When focusing on maximum values, results were skewed higher (between 2 and 39 seconds late), especially in scenarios that had significant RTA delay impacts, wind error, and only lateral maneuvers available. Overall, results in subgroup A were tightly clustered and favored late arrivals to the destination.

Upon observation of the mean values in Table X, subgroup B appeared heavily impacted by the availability of lateral maneuver options, RTA delay impacts and traffic demand. All of the mean values were below 1 second, and the most extreme of these occurred at the “peak” demand, significant RTA delay impact, and vertical-only (no lateral) maneuver IV levels. It is also worth noting that the majority of the minimum values for all of the “peak” scenario runs at the significant RTA delay impact level (18 scenarios total) in subgroup B were well below -140 seconds. Likewise, the minimum values were

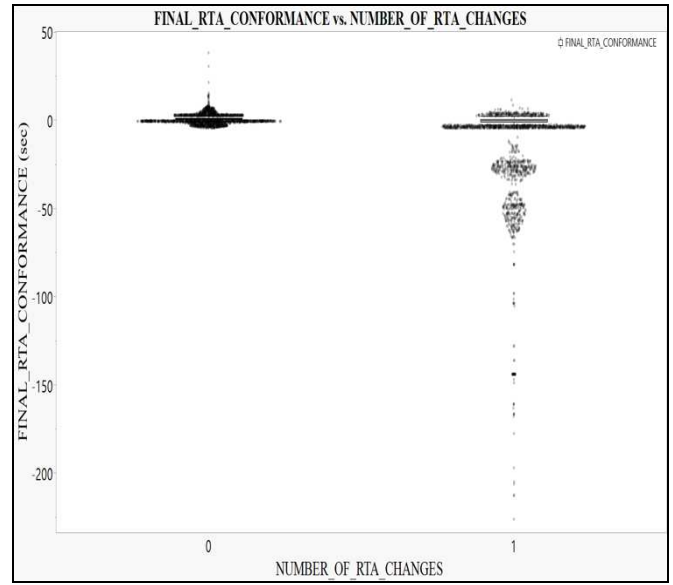


Fig. 2 RTA Conformance box plot with outliers, grouped by RTA delay count

below -100 seconds for most of the “peak” scenario runs at the mild RTA delay impact level. The maximum values all ranged from 1 to 11.5 seconds, with the lowest of these values occurring at the vertical-only maneuver level with mild RTA delay impact.

The findings in Tables VIII, IX, and X indicate that maneuver options, RTA delay impact, and traffic demand were likely significant in influencing RTA conformance – especially at the vertical-only (no lateral maneuvers) and significant RTA delay impact levels in the first two IVs. The significance of traffic demand on RTA conformance appeared to increase for flights that received an RTA delay. These trends were also observable in box plots (not shown) similar to Fig. 2, but with the five IVs as overlays and additional axes. RTA conformance data did not fit any known distributions; therefore, an additional test was not applied at the time of analysis.

It is worth noting that non-parametric test analogs to analysis of variance were considered in addition to the descriptive statistics described above, in order to see if any vastly different results would be obtained and increase confidence in the observations made using descriptive statistics for RTA conformance. In order to conduct these tests with the tools available, however, one-way analyses of RTA conformance by each of the five IVs would need to be run one at a time while holding the IV levels of the other four IVs constant. Results indicated that there was not enough flight data to rely on results obtained using this approach.

VI. POST-EVALUATION TEST AND DEVELOPMENT

Preliminary analyses of FPM-1 data led toward several implementation changes to the UML-4 simulation environment. In UAM-Sim, a trajectory prediction capability that considers wind data was integrated into the traffic pre-departure flight planner and scheduling tool, making original flight plans more realistic and feasible. A preflight CD&R capability was added that enabled flights initialized on the ground to deconflict their route with airborne aircraft (mostly

TABLE X. MEAN RTA CONFORMANCE IN SECONDS, FLIGHTS WITH ASSIGNED RTA DELAY (SUBGROUP B)

Median RTA Conformance Values	Column Labels			
	Lull		Peak	
Row Labels	No Wind Error	Wind Error	No Wind Error	Wind Error
Mild RTA Delay Impact				
	Full Maneuver set			
Full CDR	-0.15	-0.08	-1.23	-1.14
No Trajectory Buffers	-0.76	-0.83	-1.19	-0.51
No Traffic Intent	-0.16	-0.09	-0.04	-0.86
	Lateral Maneuvers only			
Full CDR	-0.02	-0.08	0.07	0.05
No Trajectory Buffers	-0.09	-0.07	-0.01	0.00
No Traffic Intent	-0.17	-0.19	-0.86	-0.03
	Vertical Maneuvers only			
Full CDR	-0.15	-0.15	-0.05	-0.03
No Trajectory Buffers	-0.15	-0.15	-0.59	-1.38
No Traffic Intent	-0.15	-0.18	-0.64	-0.05
Moderate to Significant RTA Delay Impact				
	Full Maneuver set			
Full CDR	-0.87	-0.87	-5.44	-5.54
No Trajectory Buffers	-0.97	-0.91	-2.43	-5.41
No Traffic Intent	-0.73	-0.79	-1.73	-3.12
	Lateral Maneuvers only			
Full CDR	-2.55	-0.78	-2.72	-2.07
No Trajectory Buffers	-1.05	-0.93	-2.00	-2.09
No Traffic Intent	-2.46	-0.93	-2.13	-2.04
	Vertical Maneuvers only			
Full CDR	-8.32	-8.75	-15.54	-15.40
No Trajectory Buffers	-8.67	-8.34	-14.07	-13.69
No Traffic Intent	-8.96	-9.04	-12.24	-13.04

on final approach to the same vertiport) prior to takeoff. When unable to deconflict prior to their original takeoff time, the flights are able to either delay departure time until deconfliction is possible or abort the mission entirely prior to scheduled departure. The separation standards in FPM-1 were effectively demonstrated as feasible during nominal UML-4 operations with minimal impact to separation assurance. As a result, the vertical separation component was increased from 250 feet to 500 feet in order to match more closely with that of DAA well-clear volumes.

Ongoing test and evaluation after FPM-1 showed improvements in system performance after implementation changes were made. The following data was observed in simulation runs where scenario conditions were set to place the highest amount of stress on AOP performance. Since FPM-1, the number of flights that land more than five seconds later than their respective RTA delay assignment decreased by 5.6%. The number of flights that had one LOS event decreased by 12.6%, and the number of flights with more than one LOS event decreased from 1.8% to zero. Flights with LOS durations greater than 5 seconds decreased by 7.7%. The number of flights that came within a 0.25 nautical mile slant-range distance of other vehicles decreased by 36%. Lastly, the implementation of the preflight CD&R capability led to a 71% decrease in the number of LOS events observed across each

scenario. Results from these changes to the UAM-Sim environment provide justification for re-baselining FPM-1 data and scenarios in future simulations.

VII. CONCLUSIONS

As the first research activity focused on automated trajectory management of airborne operations in UML-4, the initial performance evaluation of FPM automation generated foundational information (including baseline scenarios and FPM functional performance data) to build upon. Results of the evaluation have already led toward improvements in the UAM-Sim capability, but this capability is still a work in progress. The challenge of appropriately designing and tweaking UML-4 scenarios and operational conditions, such as giving much greater attention to the preflight phase of UAM missions, was also recognized. Integration of concurrent UAM and AAM research into scenario design and development, such as ATM environment zones unique to expected operational demands for a metropolitan area, may help to begin addressing the challenge of unknown unknowns.

Upcoming FPM activities provide an opportunity to revisit investigation of FPM-1 type data with more robust, bug-free simulation enhancements. Data collection should include more replicates to generate larger sample sizes of flight data at UML-4 traffic densities. One or two sets of IV levels representative of nominal operations should be identified and used as baselines during analyses. This will enable clearer understanding of the significance of each IV level independently when observing main and interaction effects and grant more control over comparisons against the nominal operational cases. With larger datasets, similar statistical tests (e.g., GLM) used in FPM-1 should be used while analyzing impacts to the LOS count and RTA conformance DVs as long as the data fits the appropriate statistical assumptions. The addition of more flight data also enables the use of non-parametric tests when parametric tests are not appropriate in future evaluations.

Additional metrics that should be analyzed in follow-on evaluations to characterize mission success and safety achievement include the following five metrics: flight time remaining before LOS when conflicts are resolved, the slant-range distance between vehicles at closest point of approach, vehicle energy consumption (actual vs. predicted), conflict resolution maneuver execution counts, deviation from the flight path (lateral and vertical components), and conflict resolution computation time. Both LOS count and RTA conformance will remain key metrics for characterizing safety achievement and mission success, respectively. Proper evaluation methods for RTA conformance will be investigated further. It may be necessary to perform data transform, determine appropriate outlier removal, and/or fit alternative distribution models to RTA conformance in order to apply more powerful statistical testing. To paint a broader picture and understanding of cause-and-effect relationships impacting FPM performance, parametric tests should be applied to the remaining five key FPM performance metrics that were not discussed in this paper, wherever possible, with emphasis on understanding interaction effects.

VIII. FUTURE WORK

Several research and development activities leading toward a series of flight tests will follow the FPM-1 initial performance evaluation. The AOP functionality for the reference implementation of FPM automation will continue to undergo iterative adjustments (e.g. AOP parameter tuning) based on testing results and evolving operational concepts related to UML-4. The potential benefits of additional FPM automation features, such as tactical intent-based conflict resolution functionality, should be examined. It is also suggested that the UAM-Sim environment be continuously expanded to model more UML-4 features as they are defined over time, such as communication infrastructure components (e.g., PSU) and automation system interoperability (e.g., system cross-check capability). Continuous test and evaluation within the UAM-Sim environment, including performance improvements, will continue to take place leading toward planned research activities.

Another performance evaluation of FPM automation that includes a user interface for AOP and human-in-the-loop operations will take place in the UAM-Sim environment. This activity will also generate a new set of batch simulation performance data using a subset of scenarios utilized during FPM-1. Results will inform and prepare development activities supporting integration of FPM automation into a series of flight tests focused on improving and evaluating FPM automation functionality onboard the aircraft.

REFERENCES

- [1] National Academies of Sciences, Engineering, and Medicine, "Advancing aerial mobility: A National Blueprint," The National Academies Press, Washington DC, 2020.
- [2] Federal Aviation Administration, "Urban air mobility (UAM) concept of operations v1.0," June 2020. URL https://assets.evtol.com/wp-content/uploads/2020/07/UAM_ConOps_v1.0.pdf.
- [3] Federal Aviation Administration, 14 Code of Federal Regulations Part 121. URL: <https://www.ecfr.gov/current/title-14/chapter-I/subchapter-G/part-121?toc=1>
- [4] K. H. Goodrich, and C. R. Theodore, "Description of the NASA Urban Air Mobility Maturity Level (UML) Scale," AIAA Paper 2021-1627, January 2021. <https://doi.org/10.2514/6.2021-1627>.
- [5] B. P. Hill, D. DeCarme, M. Metcalfe, C. Griffin, S. Wiggins, C. Metts, B. Bastedo, M. D. Patterson, N. L. Mendonca, "UAM Vision Concept of Operations (ConOps) UAM Maturity Level (UML) 4," December 2020. URL <https://ntrs.nasa.gov/citations/20205011091>
- [6] D. A. Karr, D. J. Wing, T. L. Barney, V. Sharma, T. J. Etherington, and J. L. Sturdy, "Initial design guidelines for onboard automation of flight path management," AIAA Aviation 2021 Forum (p. 2326), 2021.
- [7] V. Sharma, D. J. Wing, "Flight Path Management Automation Concept for Advanced Air Mobility", NASA/TM-2022-XXXXXXX, Hampton, 2022, unpublished.
- [8] D. A. Karr, R. A. Vivona, D. A. Roscoe, S. M. Depascale, and D. J. Wing, "Autonomous operations planner: A flexible platform for research in flight-deck support for airborne self-separation," AIAA Paper 2012-5417, September 2012. <https://doi.org/10.2514/6.2012-5417>.
- [9] D. J. Wing and W. B. Cotton, "Autonomous Flight Rules: A Concept for Self-Separation in U. S. Domestic Airspace," NASA TP-2011-217174, November 2011.
- [10] R. Vivona, D. A. Karr and D. Roscoe, Pattern-Based Genetic Algorithm for Airborne Conflict Resolution. AIAA-2006-6060, 2006.
- [11] Daniel Finkelsztain, Daniel, et al., "Airspace and Traffic Operations Simulation for Distributed ATM Research and Development", AIAA-2005-6488, AIAA Modeling and Simulation Technologies Conference and Exhibit, San Francisco, CA, 2005.
- [12] M. T. Palmer, M. G. Ballin, "A High-Performance Simulated On-board Architecture to Support Traffic Operations Research", AIAA-2003-5452, AIAA Modeling and Simulation Technologies Conference; Austin, TX, 2003.
- [13] Langley Enhanced IT Services (LITES), "Airspace and Traffic Operations Simulation (ATOS) System Software Project: ATOS System Capabilities Document, Version 5.0," NASA Langley Research Center, Hampton, VA, October 31, 2015.
- [14] T. L. Barney, K. M. Ballard, B. K. Buck, M. C. Underwood and R. C. Chartrand, "Simulation of a Representative Future Trajectory-Based Operations Environment," NASA/TM-2020-5006885, Hampton, VA, 2020.
- [15] J. N. Keeler, S. A. Verma, V. L. Dulchinos, and R. D. Wood, "Development of Near-term Urban Air Mobility Routes and Airspace Integration," NASA/TM-20210024606, 2021.
- [16] M. Rimjha et al., "Demand Forecast Model Development and Scenarios Generation for Urban Air Mobility Concepts," NASA Report-20205005881, 2020. URL: <https://ntrs.nasa.gov/citations/20205005881>.
- [17] D. A. Karr and R. Vivona, Conflict Detection Using Variable Four-Dimensional Uncertainty Bounds to Control Missed Alerts. AIAA-2006-6057, 2006.