Flight Testing of In-Time Safety Assurance Technologies for UAS Operations

Steven D. Young[1], Ersin Ancel[2], Evan Dill[3], Andrew Moore[4], Cuong Quach[5], Kyle Smalling[6], and Kyle Ellis[7]
NASA Langley Research Center, Hampton, VA, 23681, USA

Ongoing research at NASA is driven by a strategic plan defined by the Aeronautics Research Mission Directorate and a vision for future In-Time Aviation Safety Management Systems (IASMS) as described by the National Academies. In both visions, system safety awareness and provision are expanded through increased access to relevant data; integrated analysis and predictive capabilities; improved real-time detection and alerting of domain-specific hazards; decision support, and in some cases, automated risk mitigation strategies. One primary research focus is to develop means by which more timely (i.e., "in-time") actions may be taken to mitigate precursors, anomalies, or trends that are observed during operations. In this paper, we describe such means as a collection of Services, Functions, and Capabilities (SFCs) that are supported by an underlying information system. For example, an integrated risk assessment capability is envisioned that continuously monitors safety-related metrics and margins and recommends timely operational changes. Assessment functions and/or services can be based on data analytics and predictive models derived from heterogeneous data sets that span relevant indicator metrics and their time histories. Likewise, on-board functions can identify and reduce susceptibility to precursor conditions that have led (and can lead) to aircraft loss-of-control or out-of-control accidents. This paper summarizes development and testing of such an information system tailored to hazards anticipated for future highly autonomous flight missions near and over densely populated areas. Testing is accomplished via simulation and by using small, unmanned aircraft operating over a test range at NASA's Langley Research Center. Flight plans and test scenarios are defined to emulate several use-cases, including package delivery; reconnaissance; fire management; and urban air taxi vertiport operations. Two test phases are summarized with Phase 1 occurring in (2019-2020) and Phase 2 ongoing (2021-present). Results focus on SFC performance, technology readiness level assessment, and requirements discovery/validation. Companion papers are cited throughout for additional details on the recent testing.

---

[1] Research Engineer, Safety-Critical Avionics Branch, AIAA Fellow.
[2] Aerospace Engineer, Aeronautics System Analysis Branch, AIAA Member.
[3] Research Engineer, Safety-Critical Avionics Branch, AIAA Member.
[4] Research Engineer, Flight Dynamics and Controls Branch, AIAA Member.
[5] Research Engineer, Safety-Critical Avionics Branch, AIAA Member.
[6] Research Engineer, Safety-Critical Avionics Branch, AIAA Member.
[7] Deputy Project Manager, Aeronautics Research Directorate, AIAA Member.

I. Introduction

The System-Wide Safety (SWS) project within NASA's Aeronautics Research Mission Directorate (ARMD) seeks to develop highly assured methods or approaches to proactively mitigate safety risks to future aviation operations. The ongoing research is driven by a vision described in ARMD's Strategic Plan [1] and the concept of In-Time Aviation Safety Management Systems (IASMS) as described by the National Academies [2]. In both visions, system safety awareness and provision are expanded through increased access to relevant data; integrated analysis and predictive capabilities; improved real-time detection and alerting of domain-specific hazards; decision support, and in some cases, automated risk mitigation strategies.

As described in Ref. [1] (under Thrust 5) and as IASMS in Ref. [2], one primary research focus is to develop means by which more timely (i.e., "in-time") actions may be taken to mitigate precursors, anomalies, or trends that are observed during operations. In this paper, we describe such means as a collection of Services, Functions, and Capabilities (SFCs) that are supported by an underlying information system [3, 4]. For example, an integrated risk assessment capability is envisioned that continuously monitors safety-related metrics and margins and recommends timely operational changes. Assessment functions and/or services can be based on data analytics and predictive models derived from heterogeneous data sets that span a number of relevant indicator metrics and their time histories. Likewise, on-board functions can identify and reduce susceptibility to precursor conditions that have led (and can lead) to aircraft loss-of-control or out-of-control accidents.

The SWS project began research in this area in 2018, leading to a concept of operations as well as initial architecture and information requirements [3] for two emerging application domains. To show proof-of-concept and expose requirements, research products and a series of tests target SFCs intended to have the greatest potential benefit for safety risk mitigation and cost-effectiveness for each domain. The first domain comprises highly autonomous small UAS operating at low altitudes near and over moderate to densely populated areas; these UAS may perform a variety of public-use missions such as for surveillance/monitoring (e.g., supporting emergency response) and small package delivery (e.g., medical supplies or specimens). Mission times are not longer than 20 minutes and are executed as continuous autopilot-based flights with supervisory oversight from a remote location (e.g., an operations center or ground control station). This location may not always be in visual line-of-sight with the vehicle. A UAS traffic management (UTM) system coordinates airspace access and supports separation assurance. The second domain comprises highly autonomous occupied Vertical Takeoff and Landing aircraft (such as would be used for Advanced Air Mobility [5] or Urban Air Mobility missions [6]) conducting takeoff/climb-out and approach/landing near and over vertiports in moderate to densely populated areas. These segments are executed as continuous autopilot-based flights with supervisory oversight from a remote location and an on-board operator who may have minimal training regarding intervention in off-nominal situations. Air traffic management combines both UTM-like and traditional services and procedures and may involve coordinated operations across a set of vertiports (i.e., a vertiplex). There may be multiple vehicles operating in close proximity to each other in either domain. This paper describes an initial set of IASMS SFCs that have been developed and evaluated to support primarily the sUAS domain. However, this paper does point out where there may be cross-over applicability.

Within the project context, SFCs are defined as follows:

- Service – Information generated at a remote site and available for use during an operation.
    - Presumes the information is needed or useful to the operation
    - Requires connection to a remote server and service provider (preflight, inflight, and/or postflight)
    - Provides information via 'request-reply', 'publish-subscribe', and/or 'broadcast'
    - Produces safety-relevant information
- Function – One or more actions or other means of translating a set of inputs to a desired set of outputs.
    - Resides in all elements of the system (e.g., servers, ground control stations, onboard)
    - Produces safety-relevant outputs or data needed to compute safety-relevant metrics
- Capability – The ability to perform or achieve certain actions or outcomes.
    - May require service(s), function(s), an underlying system, and/or human involvement
    - Includes procedures, training, and an interface for any required human involvement [See note]
    - Provides safety-relevant benefits

[Note: Services and functions may also require human involvement; but for simplicity, this involvement is captured here within the overarching capability that employs the service(s) or function(s).]

Information services are not new to aviation – with some existing for decades and on a global scale (e.g., the Aeronautical Information Service [7], Meteorological Information Service [8], Flight Information Services, Traffic Information Service, Air Traffic Services, and Air Navigation Services). The IASMS envisioned here builds upon these traditional services and the service-oriented architecture that has evolved to provide the high level of safety achieved in the current commercial aviation domain. For the emerging domains, many of these services can and will provide safety benefit. However, new services are also required as has been recognized in architectural concepts (e.g., by Supplemental Data Service Providers (SDSP) and UTM Service Suppliers (USS)).

For the ongoing research, more than 20 SFCs are being investigated. These are organized below according to the research themes given in [1][2] – namely the monitor, assess, and mitigate themes.

- Monitor (predictive)
    - Services: Navigation quality; Radio Frequency (RF) spectrum; Proximity to threats; Battery/engine prognostics; Dynamic traffic density; Urban wind; Flight performance
    - Functions: Constraints; Traffic; Health of battery, engine, link, navigation system, autopilot
- Assess
    - Services: Nonparticipant casualty risk; Obstacle collision risk
    - Functions: Onboard risk assessment; Diagnostic reasoners for anomaly prediction
- Mitigate
    - Function: Onboard deterministic contingency selection logic
    - Capability: Onboard automated risk assessment and contingency execution
    - Capability: IASMS-informed preflight planning, and in-flight oversight (via operator stations)

Some of these SFCs are available to some degree in existing commercial products and are leveraged as appropriate to test the overarching system concept. Others are at a low Technology Readiness Level (TRL); with advancing TRL as one goal of the ongoing research. Many were tested in previous test series conducted by the project (e.g., as described in Ref. [3]). Selected results of prior tests will be discussed for context as they helped to define the scope and objectives of the current test series. Flight testing has the following objectives:

- Advance the TRL for IASMS SFCs; this includes collecting data for verification and validation of underlying models, software, and assumptions
- Expose requirements for future IASMS systems and their design characteristics
- Inform decisions regarding future research – Should elements or aspects be reconsidered or added?
- Advance infrastructure capability to support future testing

Key SFCs under evaluation during the current test series (and that were not tested previously) include:

- RF environment/interference (RFE/RFI) service; with server at NASA's Langley Research Center (LaRC)
- GPS quality forecast services (2); with server at an external partner and at NASA's Johnson Space Center
- Obstacle collision risk assessment service; with server at NASA's Ames Research Center (ARC)
- On-board risk assessment, contingency selection, and contingency execution functions
- On-board diagnostic reasoner and anomaly detection functions
- On-board autopilot (AP) state monitoring function applying run-time assurance techniques
- Operator interface capabilities for IASMS-informed pre-flight planning as well as monitoring of SFC outputs and supporting supervised risk mitigation during flight

There are many additional existing COTS, FAA, and high TRL SFCs that may be applied as part of an IASMS; the above list focuses on addressing gaps, evaluating alternate approaches that may be more cost-effective, and exposing or validating IASMS requirements for the target domains.

Testing is accomplished using both simulation and aircraft in flight. For flight tests, small UAS (i.e., < 55 lbs.) operating over test ranges at NASA's LaRC are used. Independent variables are: Mission/flight plan; System configuration (e.g., the software build number); SFC settings (e.g., alert thresholds); Off-nominal condition (e.g., wind or system failure conditions); Traffic scenario (for multivehicle tests); and the Participants (e.g., ground station operator). Dependent variables are (1) the aircraft, system, and SFC behaviors; and (2) participant behavior/actions. Flight plans and test scenarios are defined to emulate several use-cases, including package delivery, reconnaissance, fire management, and UAM vertiport approaches and departures.

In the following sections we highlight accomplishments during two flight test periods: (2019 – 2020) and (2021 – Present). Complementary papers provide additional details and related research findings by others at NASA and its partners. Some of these are part of special session(s) at the 2022 AIAA Aviation Forum.

II. System Description and Testing (2019-2020)

Although an initial design and some preliminary results are provided in Ref. [3], a more comprehensive set of data was collected during the 2019-2020 timeframe. This data helped to further mature envisioned capabilities and to evaluate the overarching design concept of an IASMS and enabling technologies. Three use-cases were tested along with three information services interacting with Ground Control Station (GCS) functions and on-board functions. Exchange protocols between and among functions and services were evaluated and found to be effective; with areas of improvement identified with respect to extendibility and scalability. Based on the assessment and the data collected, development was initiated for additional SFCs. Examples include: a navigation system quality prediction service, a radio frequency interference predictive capability, an electric powertrain performance prediction service, and on-board functions that can perform in-time, integrated, and automated assessments of risk, contingency management, and maneuvering/re-routing.

In 2018, the initial set of SFCs were tested for a single vehicle sUAS performing a reconnaissance mission near (and over) populated areas. In 2019-2020, additional use-cases were tested and considered. These include multi-stop sUAS package delivery profiles, multi-vehicle sUAS merging and spacing operations, and (through simulation) the applicability to larger vehicles such as may be used to achieve advanced air mobility (AAM) goals and objectives. The initial services include monitoring and predicting (a) the risk to 3rd-party non-participants near the operational area, (b) a vehicle's proximity to high-risk areas (e.g., perimeters of vertical structures along or near the flight path), and (c) the battery's remaining useful life (e.g., in terms of time or range limits until a user-defined minimum voltage threshold will be reached). Application to the AAM use-case requires relevant models for the larger vehicles such as aerodynamic, battery, and electric power train models and these are under investigation/development. One challenge to enable extension of IASMS SFCs to this use-case is the fact that the vehicles and operations themselves are still undergoing development. Furthermore, for this class of vehicles, baselines have not been established that would enable online/in-flight characterization of safety-critical components and systems (e.g., electric powertrain components). To support the assessment, initial testing of this use-case was completed as part of a joint simulation study with the ATM-X project, and additional joint tests are planned with NASA's ATM-X and AAM projects. By participating in such tests, the IASMS and SFC designs can evolve in concert with the design concepts of AAM vehicles and operations such that safety assurance barriers may be overcome somewhat proactively. Selected findings from the 2019-2020 testing follow. Updates for some of the SFCs that are part of 2021-2022 testing are discussed in Section IV.

A. Battery Prognostics (BP)

Battery Prognostics (BP) provides State of Charge (SoC) estimates and remaining-useful-life (RUL) predictions to help prevent in-flight power exhaustion hazards [9-11]. SoC is the percentage of current charge available in the battery while RUL is the amount of remaining flying time assuming previous consumption epoch. Testing during this timeframe (2019-2020) focused on deploying BP as a service at NASA's Ames Research Center. For this implementation, a flight plan was submitted to the service prior to takeoff by the operator (at NASA LaRC), and in return the service provided estimates of SoC for each waypoint in the flight plan. This allowed the operator to determine if the battery can supply adequate power for the planned flight. In addition, during flight, BP provided 1 Hz updates to SoC estimates as well as waypoint SoC projections based on changing operating conditions. Flight tests also allowed for comparing the performance of the equivalent circuit model (ECM), with a more complex electro-chemistry model (EChM) [10]. Accuracy of SoC estimates is assessed by comparing model voltages at landing with measured battery voltage after a 1-hour recovery period. Testing during this timeframe resulted in the following selected observations:

- Assessing the sensitivity of the RUL estimates to battery calibration parameter variations remains challenging due to intrinsic nonlinear mathematical complexity.
- A Monte Carlo based simulation method for estimating RUL variations can also provide information regarding energy consumption under different operational conditions. This information can enable the algorithm(s) to be tuned to the specific loading conditions; which should result in better End of Discharge (EOD) estimates based on the expected flight profile and operational conditions.
- Model instabilities can occur if inputs are not screened/filtered (e.g., to avoid negative current values, large positive/negative values, and zero current values). A constraint was added to the model implementation to prevent the positive charge states from becoming negative as a work-around.

- Data collected during experimental flights helps fine-tune the battery models and track the aging of the calibration parameters (which may need to be updated). Future work focuses on automated processes for feeding back data for updating model and calibration parameters.

Results also indicated the need for gathering more flight data with measurement of truth values so as to extend the models and their applicability to a larger, more diverse set of operating conditions, types of vehicles and battery packs. One challenge to establishing a theoretical basis for estimating the RUL confidence intervals is variability in the calibration parameters. This stems primarily from the modeled voltage instability for certain combinations of conditions. For example, in some cases, battery voltage levels were seen to vary based on temperature variation. Since the battery models rely on state observability, conditions under which states are observable need to be identified with results reflected back into the models.

B. Non-Participant Casualty Risk Assessment (NPCRA)

The NPCRA service was tested as two functions on two separate platforms: (1) a server-based pre-flight risk assessment function, and (2) an onboard in-flight risk assessment function. Below are findings and an assessment of these services as tested. For additional details see [12].

Server-based pre-flight risk assessment function: the core NPCRA functionality was implemented on a public-facing website hosted by a virtual machine or server within LaRC's Office of Chief Information Officer (OCIO). Using the website, users can upload a flight plan, aircraft specific data, and environmental information. Based on this information, users then receive non-participant casualty risk estimates for the proposed trajectory in graphical and numeric formats. A summary assessment of this functionality follows.

- *Response time*: User requests for NPCRA estimates were returned nominally in less than 90 seconds with an additional 5-10 seconds required to visualize results. Variability in response time was largely due to data availability and flight segment lengths. As part of testing conducted jointly with an ATM-X flight simulation study, the flight path parsing function was modified to accommodate potentially longer UAM flights, but response time remained under 90 seconds. As a pre-flight risk assessment tool, this response time was found to be adequate by internal users. Independent confirmation by external users remains to be done.
- *Visualization*: Graphical visualization included the entire flight trajectory, sampled impact locations, and a population density heatmap layer for a 24-hour window of population activity using the time slider tool. This representation was found to be adequate by internal users. In addition, feedback from both Zipline and the FAA indicated that the availability of the population density heatmaps for hours before and after planned flights was very informative in determining a suitable flight hour.
- *Trajectory and severity assessment models*: The risk assessment service was originally designed for multirotor vehicles; as such, trajectory and severity/consequence assessment models are tuned to these vehicle types. Based on interactions with Zipline, a fixed wing platform that can deploy a parachute in off-nominal conditions was added to the NPCRA. In both cases, the update rate for model predictions (i.e., consequences are estimated assuming a simulated failure occurs every 10 meters along the trajectory) and the population density data resolution (i.e., a 10m x 10m grid at one-hour intervals) were found to be adequate; noting that this adopts a rather conservative approach within the severity model.

Onboard in-flight risk assessment function: An instance of the core NPCRA functionality was also flown onboard the test aircraft to provide a real-time risk assessment capability (such as described in Refs. [12, 13]. In addition to the casualty/severity estimation capability described above to support pre-flight planning, the real-time version also includes a dynamic Bayesian Belief Network (BBN) which estimates the likelihood of off-nominal conditions based on inputs from onboard systems and sensors. During flight testing, this data was provided by the core Flight System (cFS) software bus. A summary assessment of this functionality follows.

- *Response time*: In the current version, the onboard NPCRA software monitors a set of aircraft state and health parameters and generates outputs based on available data. Once the software is initialized, the risk assessment takes ~50ms to execute one iteration using a simple point-mass model for the vehicle trajectory or 500-700ms when using a more complex 6-DoF trajectory model. For the system as tested, the aircraft state and health parameters are broadcast in the range of ~0.5 Hz to ~2 Hz (every ~500ms to ~2000ms). These correspond well and show NPCRA can provide risk estimates based on the most up-to-date and available onboard data.
- *Uncertainty management and quantification*: A polynomial chaos expansion (PCE) method was applied as the underlying uncertainty management framework to the model variables within the onboard NPCRA tool. This approach allows for uncertainty estimates to be generated without modification to the underlying NPCRA source code. In addition, estimates for sensitivities and confidence intervals can be found

analytically without significant increase in computational costs. The framework also allows fine-tuning the characteristics of the 3- and 6-DoF trajectory models by investigating input parameter sensitivity values.

Informed by the findings above, a third configuration of NPCRA has been developed and is part of the ongoing testing (to be discussed in section IV). This function is designed to support multiple pre-flight and in-flight aircraft and leverage the computational capabilities of the server while attempting to minimize the communication requirements (e.g., deployment of a dedicated database (SQLite or similar) to handle exchange of large population density files to speed up the server response time to accommodate real-time flight operations). As part of the future research, minimum response time requirements will be identified to support in-flight operations effectively.

C. Navigation System Performance Quality (NavQ)

The NavQ service and supporting functions remain in development. The concept envisions functions that provide (1) predictions of Global Positioning System (GPS) and Inertial Measurement Unit (IMU) performance, (2) access to real-time positioning information from available non-GPS based backup systems, and (3) access to information from available GPS augmentation systems [3]. Multiple assessment-related accomplishments occurred in support of implementing the overarching concept for future testing and evaluation. With respect to access to backup systems, a significant accomplishment was via an inter-agency agreement (IA) with the Department of Transportation (DoT) to test, evaluate, and demonstrate Alternate Positioning Navigation and Timing (APNT) systems as requested through the National Defense Authorization Act for Fiscal Year 2018. APNT systems are systems that operate independently of GPS, or any other Global Navigation Satellite System (GNSS). The Congressional request sought information regarding the efficacy of current APNT systems that may be used as a backup or compliment to GPS. In response, DoT led a multi-agency effort to evaluate eleven different systems. NASA LaRC was the test site for six of the systems. Testing occurred over several weeks in which personnel from NASA, DoT, MITRE, Zeta, and APNT vendors participated in a pre-defined set of trials. During the trials, both ground-based and airborne data were collected using an equipped test van and an unmanned aircraft, respectively.

Testing culminated on March 13, 2020, with a briefing and demonstration at LaRC. Results obtained in these tests were used as part of a final report to Congress regarding the potential backup systems for GPS and recommended next steps [14, 15]. An example of generated results can be seen in Fig. 1 for one of the APNT systems. Each of these plots represent data captured during a single UAS flight. The two plots on the left side of the figure compare the post-processed and validated reference path (in red) and data collected from the APNT system (in blue). In these plots, each data point for the APNT system is color coded to show the magnitude of positioning error (darker blue represents larger error). The plot on the top right shows cumulative 3D position error distribution for the full flight; while the lower right shows error magnitude for all points in the flight. As expected, the magnitude of the 3D positioning error is strongly but not completely correlated with horizontal positioning error. Similar data assessments were done for all APNT systems and flights.

Additional assessments are ongoing based on a set of quantitative performance measures as well as other considerations (e.g., cost-to-implement and maintain, resilience to interference, and TRL). Performance measures include positioning accuracy (2D and 3D), timing accuracy, update rate, availability, continuity, integrity, and reliability. Due to the importance of high-integrity PNT data for safe flight operations, combined with the scarcity of adequate solutions for low altitude urban environments, it is envisioned that access to information from available APNT systems will play a crucial role in AAM operations.

D. Geo-spatial Feature Models and Databases

Accurate geospatial feature data (e.g., 3D boundaries of buildings, power lines, and trees) are required for several envisioned services. For example, as previously discussed, feature data for the NASA LaRC flight test range (derived from a 2018 high-resolution LiDAR-based survey) was used for both the NavQ and Proximity to Threat (PtT) services and function development. The PtT service is further discussed in Section III and in Ref. [16].

In addition to traditional LiDAR-based survey data, researchers began to assess commercially available high-volume satellite-based survey data [17]. A satellite-based survey of a 250 km^2 region containing an entire NASA National Campaign-Developmental Test Airspace [18, 19], including downtown Dallas and the DFW airport, was acquired and is under evaluation as input to the higher-level SFCs under development. Example images of these two data sets are shown in Fig. 2., where the LiDAR-based data for is shown on the left and the satellite-based data is shown on the right. Both the native survey representation (point cloud) and a more compact polyhedral representation (shape file) are being evaluated. All buildings in Fig. 2 are shown in compact form, but the trees in the left image are shown in point cloud form.

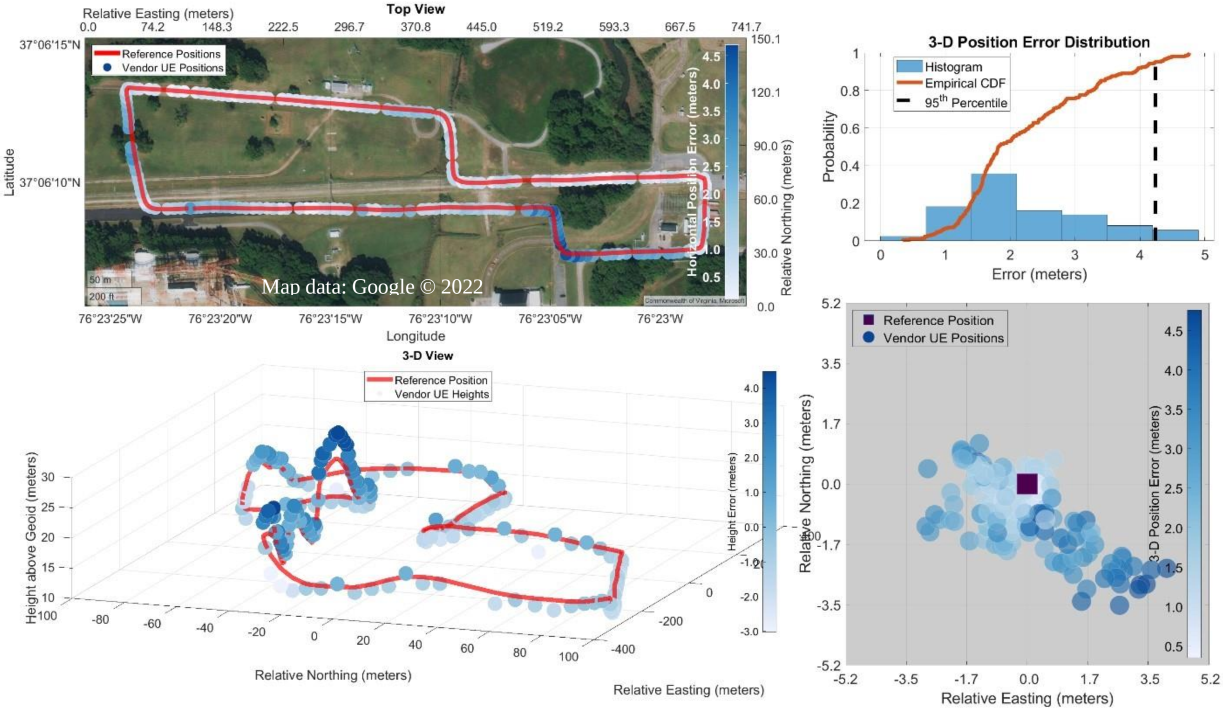


Fig. 1 Example results of APNT system performance assessment (one test flight).

Fig. 2 Geospatial databases used to support SFC research and development.

E. Assessment of Information Exchange Protocols

IASMS architectures (such as presented in Ref. [3]) rely on timely information exchange between the ground-based systems/services and the flight vehicle through a ground station (GS) system or operations center. This communication may occur over multiple communications links with potentially different protocols. For the developmental system under test, the link between the vehicle and the GS is defined as a binary message encoding sent over either a dedicated serial link or standard TCP/IP link. For the links between the GS and SDSPs, web-based information exchange protocols using JSON and XML encoding over TCP/IP are adopted to allow flexibility during development and integration. The use of web-based information exchange protocols allows leveraging well-defined session management and data security protocols when passing raw measurements and safety metrics. Because JSON and XML are text-based encodings, they are easily extensible to add data types or accommodate changes in the underlying data model.

Assessment of protocol and bandwidth: Figure 3 shows an example of battery voltage and current draw measurements sent from the vehicle to the BP SDSP during one of the test flights. This data was exchanged during execution of package delivery flight mission profile (Flight 109). In this case, measurements and reports are exchanged between the GS and BP at 2 Hz allowing timely reporting of battery safety metrics to the UAS operators. This shows that JSON- or XML-based internet protocols can be effectively used to send time-sensitive data. Battery measurements were sent from the vehicle to the GS over a dedicated wireless link, with relevant information then forwarded from the GS to the BP service over the ground network infrastructure. For this representative flight, the vehicle and its GS exchanged 4.9 MB of data at an average rate of 28.7 kb/s; while the GS to BP link exchanged 2.53 MB of data at an average rate of 14.5 kb/s.

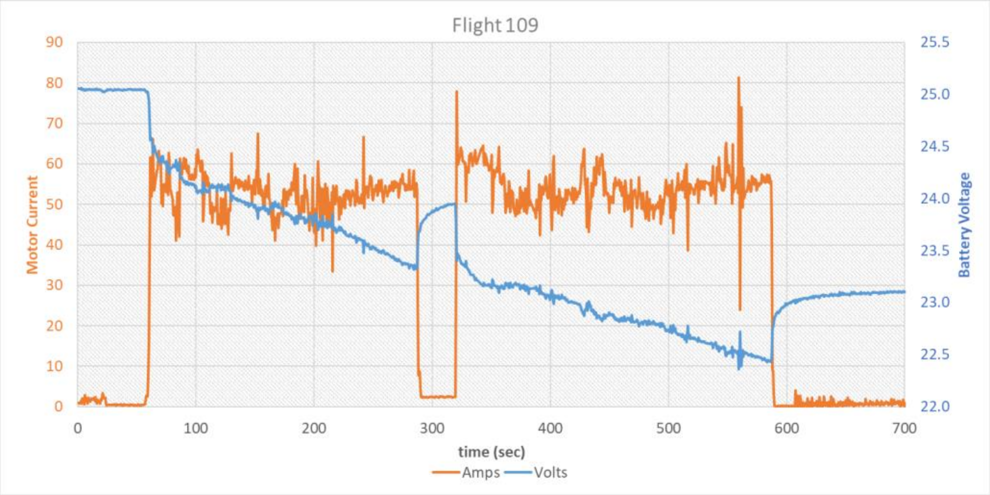


Fig. 3 Battery voltage and current draw values as sent to the BP service during Flight 109.

To further investigate the impact of exchange protocol selection on the scalability to enterprise level operations, a playback simulation was constructed to examine the response time as vehicle count increases and to collect data in order to model bandwidth requirements for enterprise-scale scenarios that may have many request/reply demands. At the enterprise operations level over an urban area, it is likely that the vehicle to GS link will be implemented on the cell infrastructure using internet TCP/IP links. Based on this assumption and for the worst case from a bandwidth load perspective, each vehicle is managed by its individual GS and the GS is also communicating on the cell network. This implies the exchanges between the vehicle and its GS may incur twice the cell bandwidth – one exchange for the vehicle to the cell infrastructure and a second exchange from the cell infrastructure to the GS. This test configuration creates the worst case for the GS to SDSP links as it adds the network traffic on the cell infrastructure once more.

Due to limited access to equipment, assessment testing used flight log playback scenarios that were setup to model the worst-case scenario over the public mostly-wired network. While cell infrastructure is not available in these tests, it is still useful to measure the bandwidth and transaction rates over the land network to collect data for modeling and bandwidth projections. Fig. 4 shows a sample playback run of four vehicles “flying” the same package delivery flight plan with cascaded start times. Each vehicle was “operated” by a pair of computers – one representing the vehicle and one the GS. The same computer types were used as in the flight testing. Software executing on the GS is identical to that used on the flight. Likewise, software executing on the vehicle is similar except for the playback components that retrieve raw data from log files recorded during previously conducted flights. In total, four similarly-configured pairs of computers are used to communicate with the same SDSP which supplied the previously-described battery safety metrics. Network statistics and transaction rates were collected to project bandwidth needs at enterprise scale.

There are 3000+ transactions in each playback log. Fig. 5 shows transaction time which includes time from the message leaving the GS to the time the server reply is received. An average of 3.2 transactions per second (tps) is observed at the start and peaks at about 12 tps when all four flights are active. Even at 12 tps, a vast majority of the response times are in the 150 ms to 380 ms range with a few taking multiple seconds. A deeper look in Fig. 5 shows that 500-byte-sized transactions take about 175 ms longer to complete than 200-byte-sized transactions. Unfortunately, the relative contributions of message size and server processing time is not discernable in this measurement. Keep in mind this measurement is done using land infrastructure and the delay is likely to increase when conducted over cell service. The criticality of a 100 ms response time can be considered in the context of a free fall incident, where a vehicle operating at 400 ft will free fall to impact in about five seconds.

By overlapping and executing multiple playback flights using this tool, preliminary estimates regarding scalability of the architecture and information exchange bandwidth can be computed. For example, Fig. 6 projects the bandwidth needs out to 100 vehicles interacting with 10 SDSPs and their services. Since one of the premises for urban flight operation is availability of cell service for communications links, the published bandwidth of 3G and 3G-HSPA cell sites are shown as benchmarks in the graph. When the projections are put in the context of a single cell site, the approximate system capacity can be deduced based on bandwidth limits. As indicated in the graph, a scenario of 50 vehicles interacting with five SDSPs would consume most of the bandwidth of a 3G cell site. The 10 vehicles to 10 SDSP case is probably a more realistic saturation point because there are other users of the bandwidth in that site and the available bandwidth may flex due to environment conditions. For the 10x10 case, it takes up about 2.17 Mb/s of the available data rate. Ten vehicles in a 3-5 mile cell gives a lower, and in several other respects more manageable, airspace density than the former case of 50 vehicles in a cell area.

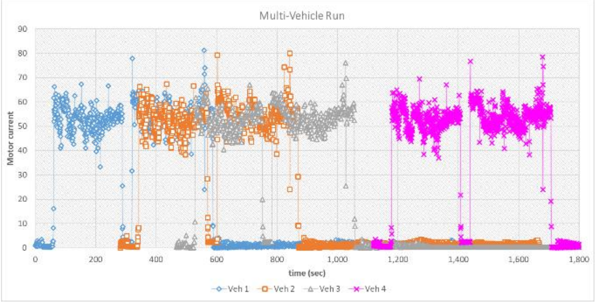


Fig. 4 Playback simulation test – four vehicles sending motor battery data to the BP service provider.

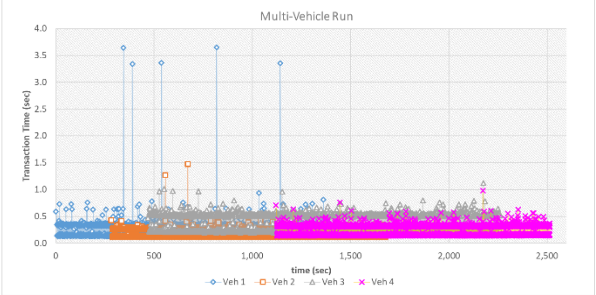
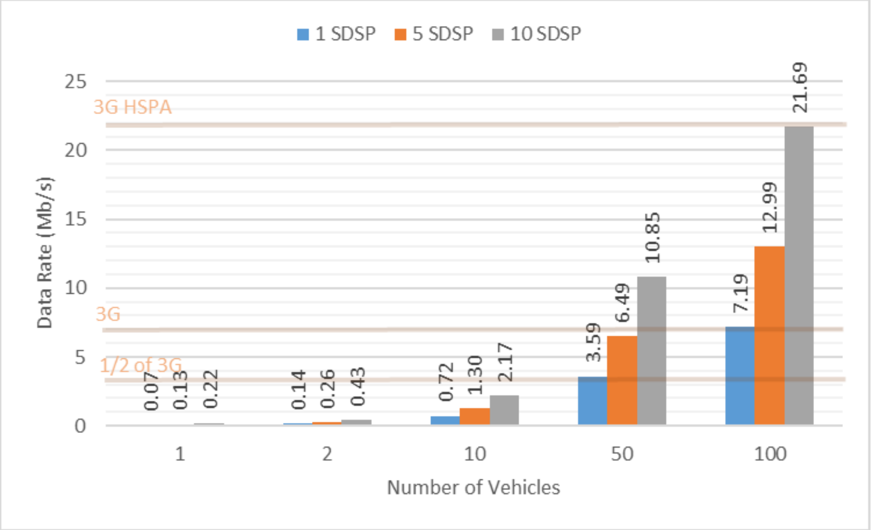


Fig. 5 Transaction response times for overlapping flights.



-

Fig. 6 Bandwidth projections for information exchanges at enterprise scale.

Findings and observations. (1) The JSON message encoding used in the developmental system can be effective in passing time-sensitive data. Also, JSON encoding can be easily transformed to support industry accepted models such as the Aeronautical Information Exchange Model (AIXM). JSON- and XML-based message encoding are, however, not the most efficient means of encoding if bandwidth limits are a concern. (2) Message size relates to response time which can impact safety (i.e., late responses may not be useful if the context has changed since the request was issued). There is a measurable difference in transaction time between 200- and 500-byte sized transactions. Using shorter transaction data payloads should improve overall system response which is critical when monitoring safety parameters. (3) Operating GS on a ground-based (land) network could free up cell network bandwidth to increase information exchange capacity. (4) Future work may consider using CNPC-based data links in addition to or in place of cell links for vehicle to GS exchanges. (5) Future work will also consider proxy services between GS and SDSPs to reduce network load if the GS is operated from the cell network.

III. System Description and Testing (2021-Present)

For the current phase of testing, the system-under-test has been updated and expanded based on the previous findings, industry developments, and maturation of the IASMS Concept of Operations [20]. More than 20 SFCs have been implemented and are at various stages of technical maturity. The revised system architecture is shown in Fig. 7. Several additional acronyms are used in Fig. 7 to represent elements of the system not described in Section II.

- OS&N – Observation Stations and Network (e.g., for Weather (Wx) and RF spectrum measurements)
- APPDAT – Application Platform, Packaged Deployment and Analytics Technologies [21]. APPDAT is a platform hosted at NASA Johnson Space Center (JSC) for developing cloud-native applications and making services discoverable and accessible by both NASA and non-NASA users.
- BP – Battery Prognostics [3, 9-11], a service that tracks and predicts state-of-charge and remaining useful life of onboard power source(s); outputs include estimated time when end of discharge (EOD) will be reached (or remaining flight time); and probability of reaching EOD before end of mission. See also Section II-A.
- PtT – Proximity to Threat [3, 16], a service that tracks and predicts proximity to high-risk areas near the flight path (e.g., the perimeter of vertical structures); outputs include portions of the vehicle trajectory that violate proximity thresholds (including start/end points); nearest approach point; distance to nearest approach point; and severity of violation.
- NavQ CAP – Corridor Assessment of Positioning [22], a NavQ service that provides estimates of Nav-related performance measures along a user-specified flight corridor and time window (described further below).
- NavQ GAP – Geometric Assessment of Positioning [23], a NavQ service that provides estimates of Nav-related performance measures over a user-specified coverage and time window (described further below).
- RFE/RFI – RF Environment and RF Interference [3], a service that provides estimates of RF-related performance measures over a user-specified coverage area and forecast period (described further below).
- USS/PSU – UAS Traffic Management Service Supplier [24] and Provider of Services for UAM [6]. These emerging capabilities provide air traffic management services.
- FIMS – Flight Information Management System [24]; provides access to FAA-provided information services within UTM ecosystems.
- DD – Dynamic Density [25], a service supporting air traffic management safety by tracking (and forecasting) metrics associated with air traffic density for selected airspace volumes. This service is not used during the sUAS flight tests; however it is part of flight simulation studies being conducted by NASA's ATM-X project.

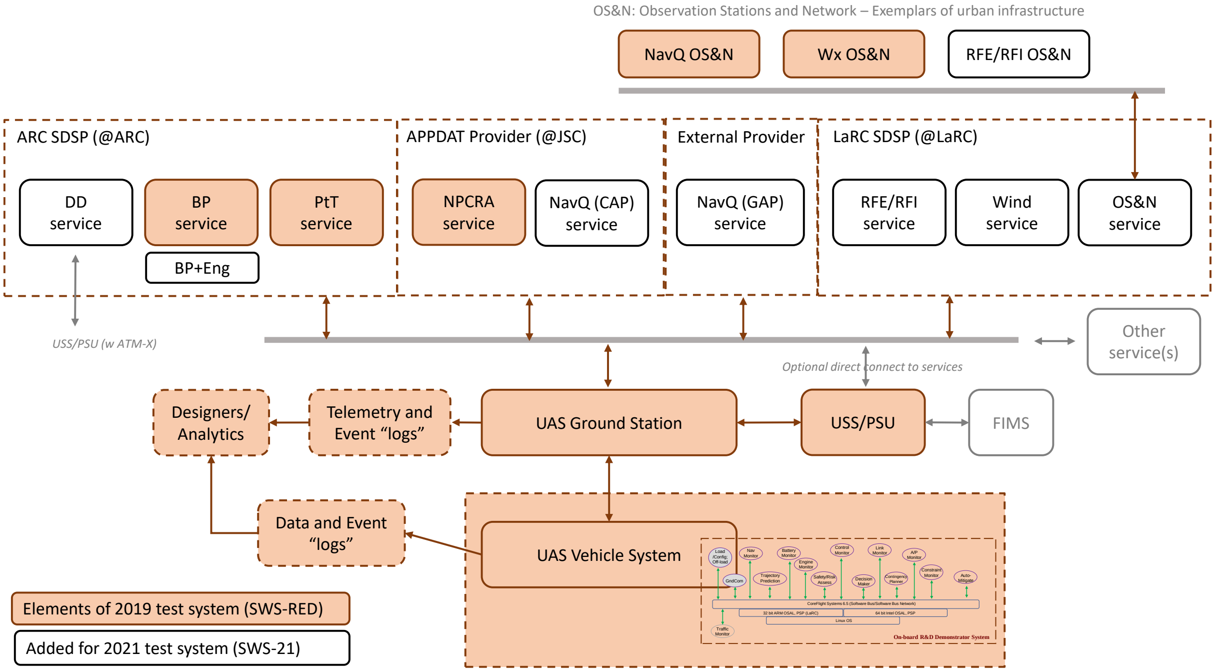


Fig. 7 IASMS-based system-under-test.

A. Key Off-board SFC Developments (post-2020)

GPS performance *prediction capabilities and functions*. As part of the broader NavQ concept, two methods were explored to develop a tool or function to predict the quality of satellite-based positioning information (e.g., from GPS) – in particular for flight segments at low altitudes near vertical obstructions such as buildings or trees. Future work may combine these methods into a single service or function. The first method, Corridor Assessment of Positioning Systems (CAPS) service, produces a functional navigation fidelity map for a specified flight corridor. This map contains predictions of blockage of orbital satellite signals by geospatial features; predictions are computed by ray-casting from positions in the flight range to each satellite through high-resolution models of the physical environment [22]. The resulting function has advanced from a prototype to a beta-quality web-based mapping tool with performance suitable for pre-flight planning. Development has begun to package the map values into a UAV-compatible data stream to enable a function that may be used in-flight. The interactive map and its output data stream have been ported to the NASA-hosted APPDAT-based cloud server [21] for early adopter evaluation.

A second method, the Geometric Assessment of Positioning Systems (GAPS) service, is being developed in collaboration with the MITRE corporation [23]. This prognostic service leverages available real-time satellite state information (e.g., ephemeris) and a digital surface model of an operational area to generate a discretized time-varying 4D volume of estimates of a set of quality metrics associated with GPS (e.g., the Horizontal Dilution of Precision (HDOP)). This volume may be generated across a user-selected flight environment and time period. High Performance Computing (HPC) methods and optimizations are employed to increase the computational efficiency of this process, allowing for real-time deployment in certain scenarios. Under development is a capability to integrate real-time measurements from networked GPS receivers in the operational area to continuously refine the predictions and underlying models. Results of this process have been generated in a developed prototype, along with a visualization tool. Predictions are generated based on a user specified region, timeframe, and performance measure of interest. An example is shown in Fig. 8 for the HDOP measure. In this example, the yellow area in the center is a building with a height of ~150m. Comparing the two images, the red X marks an area southwest of the building where HDOP is predicted to change significantly within nine minutes. On the left, a flight could maintain good HDOP down to the surface at this location. However, on the right (9 minutes later), good HDOP can only be maintained down to ~80m altitude. This indicates poor navigation performance can be expected below 80m here (e.g. during takeoff and landing).

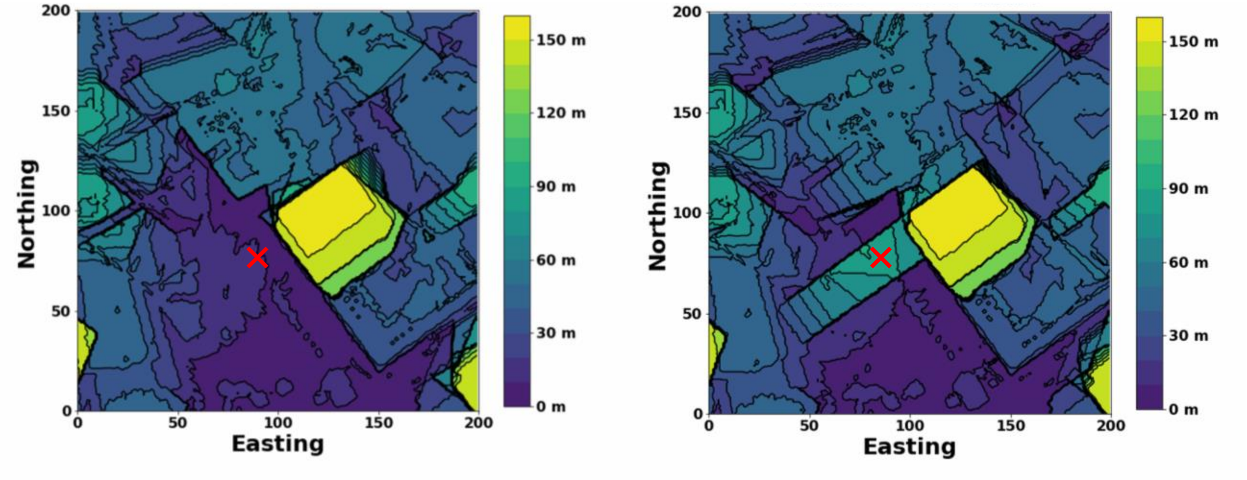


Fig. 8 Example result of predicted GPS HDOP performance measure.
(Test region is a 200m x 200m segment of downtown Boston, MA)
(Colorbar represents altitude at which HDOP remains < 5)
(Left – at T+21 min; Right – at T+30 min)

Current work is focused on uncertainty quantification and validation, analyzing various techniques for infusing real time data from GPS receivers located in the area, improving computational speed, and moving to both SDSP (server-based) and onboard versions. The in-situ receiver data can help to both quantify the uncertainty in the predictions and update/refine/validate the computational methods. Testing is ongoing for the downtown area of Corpus Christi, TX, in collaboration with partners at the Lone Star UAS Test Center. Results will be published in the future.

NPCRA service capability update. The NPCRA concept discussed in Section II.B and Ref. [3] was significantly updated and implemented as a NASA-hosted cloud service [21]. Following the migration, testing showed the response of the service during pre-flight was considerably improved (<10 seconds). Additionally, publicly available building footprint files were ingested to account for sheltering effects on non-participant populations. As part of the 2021-2022 testing, use of an API-compatible version during flight has been implemented to provide a source of updated casualty probability (Pc) estimates. Latency is being assessed associated with use of computation/data provided by cloud-based services to support in-time risk mitigation capabilities. In such tests, the GCS computer uses downlinked telemetry data (or alternative trajectories as part of re-routing in case of emergency) and queries the NPCRA server for Pc values. Once complete, NPCRA relays the results for the current and/or proposed trajectory to the GCS. The assessment result is either evaluated by the GCS operator or uplinked to the aircraft to support automated mitigation actions. This type of testing is ongoing with results to be published in the future.

RF environment hazards monitoring. RF interference has a known role in compromising safety of flight operations. They include both coupling within and across on-board equipment and degraded or failed communication links. These hazards can have particularly critical manifestations in highly automated UAS and UAM flights; where they can lead to (1) unintended AP mode switching on the vehicle, (2) loss of operator situation awareness when intervention may be most needed, and/or (3) the inability to recall a vehicle entering restricted air space (e.g., due to an out-of-control condition). Adding to performance uncertainty, RF energy levels vary in time and are influenced by weather, local terrain/structures, and human activity. There is a higher likelihood that these factors will co-exist during low altitude urban flight operations, which suggests a critical need for effective RF environment monitoring.

To help mitigate this risk, an RF risk (or safety) monitoring and assessment service is envisioned (e.g., see Fig. 9 and Refs. [3, 26]). To develop and evaluate the feasibility of this concept, a set of COTS observations stations have been installed and an RF propagation model has been developed for the NASA LaRC test range [26]. The model is based on locations and types of known emitters, and local-area digital surface models (DSMs) such as described previously. The resulting services and functions will be evaluated in the future according to project milestones. The recent testing has focused on a (1) evaluation of the COTS spectrum monitors (which are often used near commercial airports and space vehicle launch sites); and (2) model validation using data collected by a ground and air vehicle equipped with a mobile RF spectrum monitor to cross-check model-based estimates (see Fig. 9). Some research also investigates using the model to calculate anticipated energy levels along the flight path to provide operators (or automation) with awareness of the potential for RF hazards in-trail.

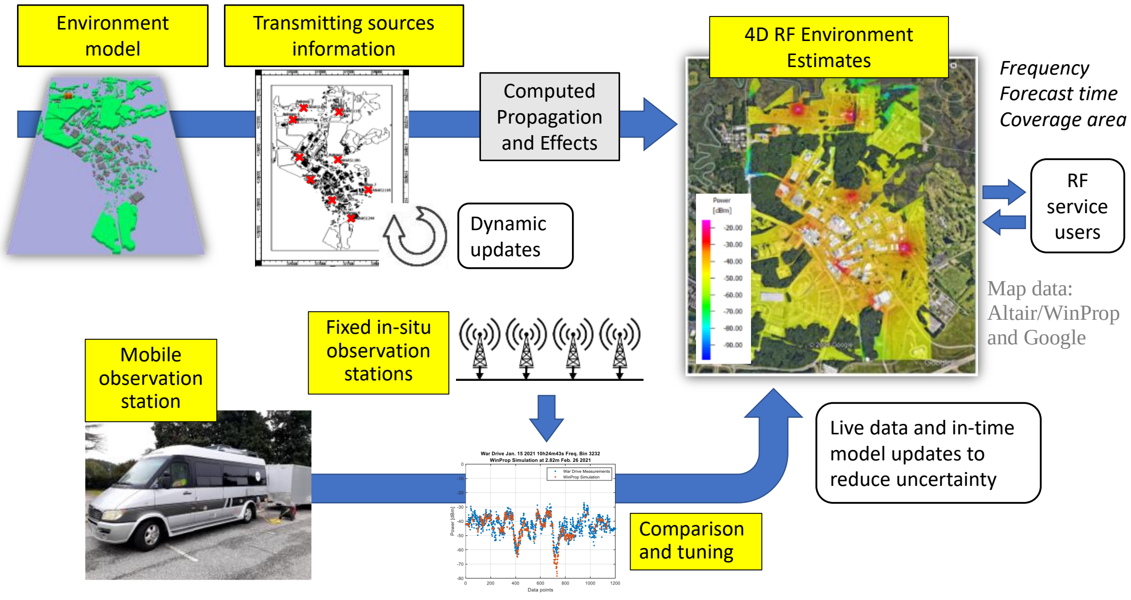


Fig. 9 RFE/RFI service concept as tested [3, 26].

Wind estimation and modeling. This research endeavors to provide improved estimates of urban wind speed and direction along a flight route submitted to a service by end-users or functions. The current wind measurement and forecasting infrastructure is well-suited to traditional aviation operating near airports or at higher altitudes. However, when flying at low altitudes in areas with complex terrain (including buildings and vegetation) and not near airports, available wind forecasts must be treated as coarse indicators only. As it is impractical to fully instrument even modest-sized urban areas, methods are needed to interpolate from low-resolution wind measurements and/or forecasts along possible UAM and sUAS flight routes. Ideally, such an interpolation function would also provide forecast uncertainty bounds that reflect its inferred or measured limitations. Further, if flight is anticipated near vertical structures, air flow disturbances around such structures may need to be considered. Two activities are underway in this area:

- Computational methods are being developed and tested to project wind effects from kilometer-scale resolutions onto meter-scale resolutions. Wind simulation software that includes a terrain model and a prevailing wind speed is being tested for performance and accuracy. As this point, the service can provide computed wind fields represented (after thresholding) as hazardous regions for interactive display, or as geo-tagged values in an SDSP data stream.
- Both publicly available wind observations/forecasts and data products from a commercial wind data service provider are under evaluation [27, 28]. Computed winds fields as well as these data are compared with actual readings from a network of NASA-hosted weather stations at LaRC. Optimization methods are being investigated to reduce computation latency and adjust to near real-time inputs from available wind stations along planned flight paths. This activity complements ongoing collaborations with external partners via NASA's Small Business Innovative Research program.

Human/machine interfaces. Flight research and system requirements analyses have been aided by three user interface development activities; and over the course of development, insights about In-Time Aviation Safety Management procedural requirements have in turn been applied. Although there is some overlap in intended function and requirements, each interface targets a different primary objective in operations. These include:

- SDSP Dashboard [29] – primarily intended to support users of safety-related cloud-based services, particularly users who may be using several information services as part of pre-flight planning or in-flight monitoring for safety management.
- GCS Operator Interface [30, 31] – primarily intended to support GS operators as they create and load flight plans; and then track flights in progress. For example, they may preview a flight and determine whether safety metric thresholds may be exceeded; they may observe in-flight trends across the safety metrics and send commands to the UAS autopilot; and they may review/approve suggested autopilot mode changes such as may be generated by on-board automation and sent to the display via telemetry. This interface is an extension to an existing COTS product [31].
- Operations Planning Tool – primarily intended to support researchers and aviators as they set up scenarios/flight plans and configure the system for testing (e.g., setting appropriate thresholds for monitoring functions and defining coordinated flight plans for multiple vehicle tests). Details on the design and evaluation of this tool will be published in the future.

B. Key On-board SFC Developments (post-2020)

The current onboard system architecture is an extension of the cFS-based system previously tested [3, 32-33]. The updated and current system architecture is shown in Fig. 10. The primary focus of the update is the addition of functions (and a capability) that allow for automated safety risk mitigation based upon:

- Additional run-time assurance (RTA) techniques (e.g., Refs. [34, 35])
- Consideration of multiple hazards by an integrated risk assessment function [13]
- A function designed to safely bound behavior against a set of constraints [36]
- Tracking of AP states and context-specific limitations

Note: The risk mitigation capability is described in detail in a companion paper at the 2022 AIAA Aviation Forum [37].

Fig. 10 illustrates how the cFS-based onboard system operates independent of the COTS onboard system (e.g., the AP). This independent system includes several cFS applications that monitor various safety-critical system states, assess risk, and support risk mitigation (automated or supervised). Independent monitoring is a tenet of RTA approaches (i.e., a highly-assured function that monitors the behavior of a less-assured system).

Outputs of the cFS-based system are provided to two GCS platforms to support supervised risk mitigation – a COTS platform used by the GS operator, and a custom platform used by researchers. A separate output goes to the AP; this output is only activated for tests of the automated safety risk mitigation capability as described in Ref. [37].

Several additional acronyms and labels are used in Fig. 10 to describe elements of the on-board research system. These include:

- GroundCom – A function that converts cFS bus exchanges for downlink to researchers during flight; the function also supports pre-flight uploading of configuration files and post-flight off-loading of recorded data.
- Contingency selector – A function that chooses and initiates AP mode changes from available alternatives (as determined by the AP state monitor application); this function is further discussed in Ref. [37].
- Safety metric monitoring – A collection of functions that monitor parameters associated with battery health, engine/motor performance, proximity to threats, traffic, and link performance [3, 38]; this also includes diagnostic reasoning and anomaly detection functions such as have been used on space platforms [39]. Extensions to previous research on battery health monitoring on-board were developed to support auto-mitigation of battery-related hazards. For this work, algorithm and models originally designed for a fixed-wing UAV are adapted for multi-rotor sUAS and 1 Hz SoC and RUL estimates are reported to the cFS software bus to support risk assessment and contingency selection [37].
- Risk assessment – A function that uses a Bayesian belief network technique to assess safety risk in terms of likelihood and severity; likelihood estimates may be for various user-selected outcomes (e.g., loss of control due to loss of power), while severity estimates are based on casualty risk to people on the ground (i.e., based on potential impact areas in the event of loss-of-control) [12, 13, 37]. See also Sec II-B.
- AP state monitor – A function that monitors the state and behavior of the COTS autopilot; this includes monitoring the AP mode as well as modes that may be safely transitioned to (from the current state); because of the role this function plays in auto-mitigation, an RTA/RTV technique has been applied to improve the level of assurance [34, 37].
- Safeguard (SG) constraint monitor [36] – A function that warns of impending violations of geo-spatial constraints (e.g., stay-in and stay-out areas) as well as other constraints (e.g., range limits, path deviations, and altitude/airspeed restrictions); SG operates on isolated hardware and is qualified at the NASA Class B safety-critical level; the assurance level required of safety-critical software on uncrewed NASA spacecraft.
- NUCs – Two Intel-based single-board computers (Primary and Secondary Node (PN/SN)); each node hosts the same applications, with the only difference being that the SN is not connected to the AP.
- Bridge applications – Convert information passing between serial buses and the cFS bus. These are for communicating with the COTS AP (MavBridge), COTS embedded systems and sensors (CANBridge), and the Safeguard unit (SGBridge).

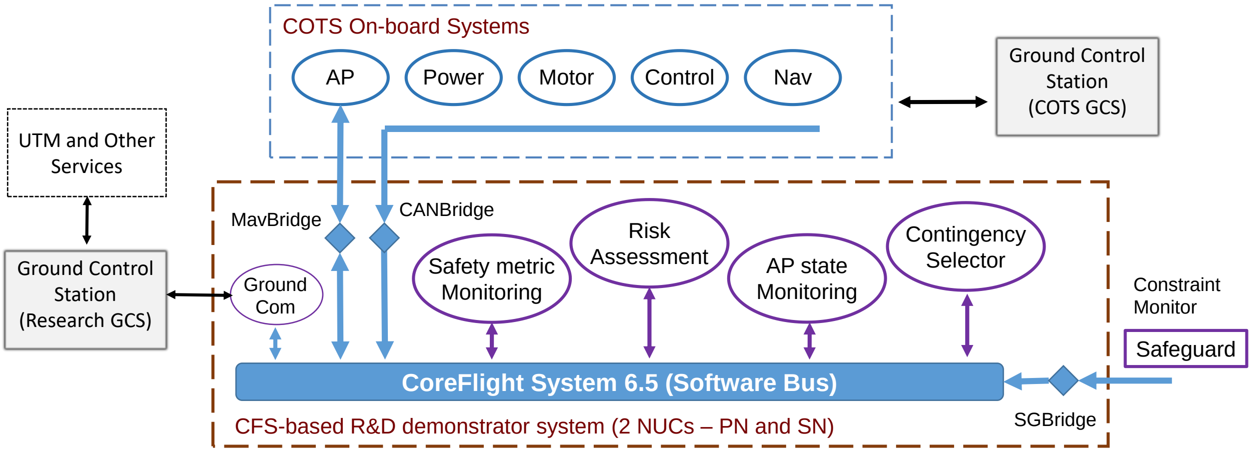


Fig. 10 UAS on-board system as tested.

C. Test Vehicle and Platform

The configuration of the vehicle remains similar to what was used during the 2019-2020 flight testing; however, three important changes were implemented. The airframe was changed from a DJI-based octocopter to a Tarot-based octocopter. The motors used are from KDE Direct and were selected for their ability to provide real-time telematics to an Arducopter-based autopilot. Finally, the research computer system was expanded to two computers to allow research software with different classification levels to be flown at the same time. The modified Tarot is shown in Fig. 11 as equipped for the recent testing. As shown, the COTS autopilot is on top, with two payload trays underneath (one for the research hardware and one for the batteries). Underneath the research hardware tray, a large white omni-directional RF spectrum monitor antenna is shown as mounted for some flights. The GPS receivers and most of the communications antennas reside on top of the airframe. The take-off weight in research configuration is about 28 lbs.

Stator Dimensions	OD: 42mm Height: 13mm
Max Thrust	10 lbs
Efficiency at Hover	17 g/W
ESC Current Rating	40 A
ESC Operating Voltage	6 S LiPo
Propeller Size	18”

Fig. 11 Test vehicle.

D. Test Ranges

As with the 2019-2020 testing, two areas within the NASA LaRC test range are used. Range 1 is mostly open field; although there are some geo-features of interest (e.g., an extended linear paved feature akin to a runway or roadway). Range 2 includes several buildings and streets (i.e., indicative of sub-urban/urban regions), with occasional surface traffic and pedestrians. Imagery of these ranges is given for each scenario in a later sub-section.

E. Test Objectives

As described in Section I, test objectives are to: (1) Advance/assess the TRL of the system concept as well as for the IASMS-based SFCs; Collect data for V&V of underlying models, software, and assumptions; (2) Expose requirements for future IASMS systems; and design characteristics requiring refinement; (3) Inform decisions regarding future project research objectives; and (4) Advance test infrastructure capability to support testing with additional partners in the future (e.g. install/evaluate a COTS RF spectrum monitoring system). In terms of objective (1), TRL is assessed based on three criteria (for TRLs of five or greater):

a) tests span conditions that are representative of what may be encountered in the application domain(s)
b) technologies perform their intended functions with no unexplainable behaviors or unintended (negative) consequences
c) usability and acceptability ratings of “good” or better are achieved for any human interface elements

Consideration for these criteria also helps to achieve objectives (2) and (3) (i.e., exposing/validating requirements and planning future work).

F. Scenarios

To achieve the project’s R&D goal, TRL criteria 1 is an important consideration when defining scenarios. Namely, “tests span conditions representative of what may be encountered in the application domain(s).” In order to define scenarios that can cover this span, a set of use-case types are selected. Then, at least one nominal or reference scenario is defined for each use-case type; and one or more off-nominals are defined that expose/utilize the SFCs (and the system concept) under test.

Four use-case types are selected: Package Delivery, Reconnaissance, Vertiport Approach/Departure, and Fire Management. The first two were also used during the 2019-2020 testing. The latter two allowed alignment with other projects underway at NASA aimed at UAM vertiport operations [40] and traffic management for emergency response [41, 42]. A common set of 15-20 off-nominals were defined, primarily based on situations that could expose the output state space of the automated risk mitigation capability. Example off-nominals include low battery power, C2 link loss, and close proximity to people, structures, or airspace constraints. Flight plans for three of the nominal scenarios are shown in Fig. 12 to Fig. 14. Simulated traffic can be added to any scenario and is one of the key independent variables for the testing. Some traffic examples are shown in the figures.

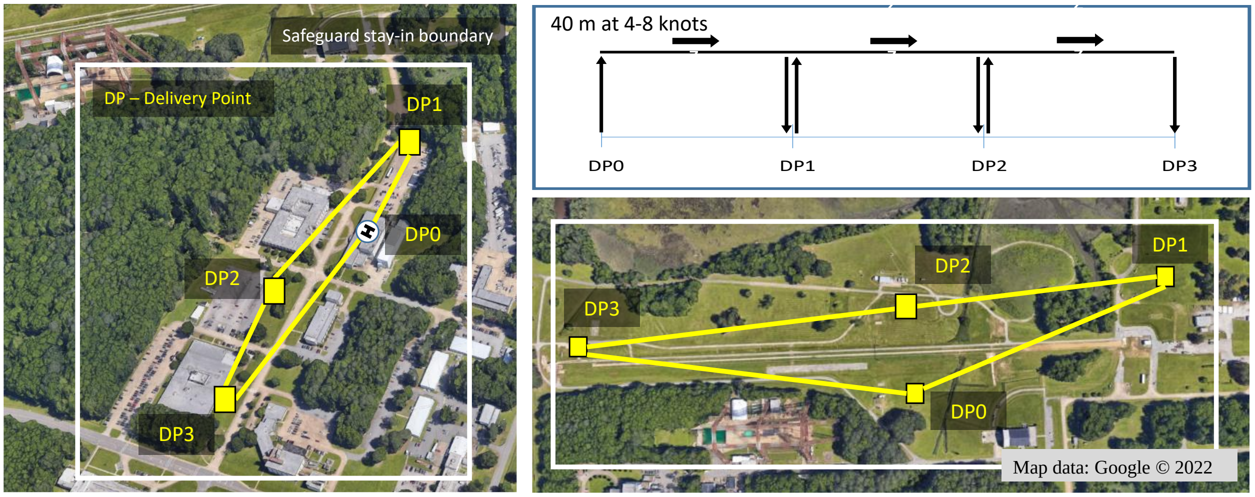


Fig. 12 Nominal package delivery scenarios.
(right – range 1; left – range 2)

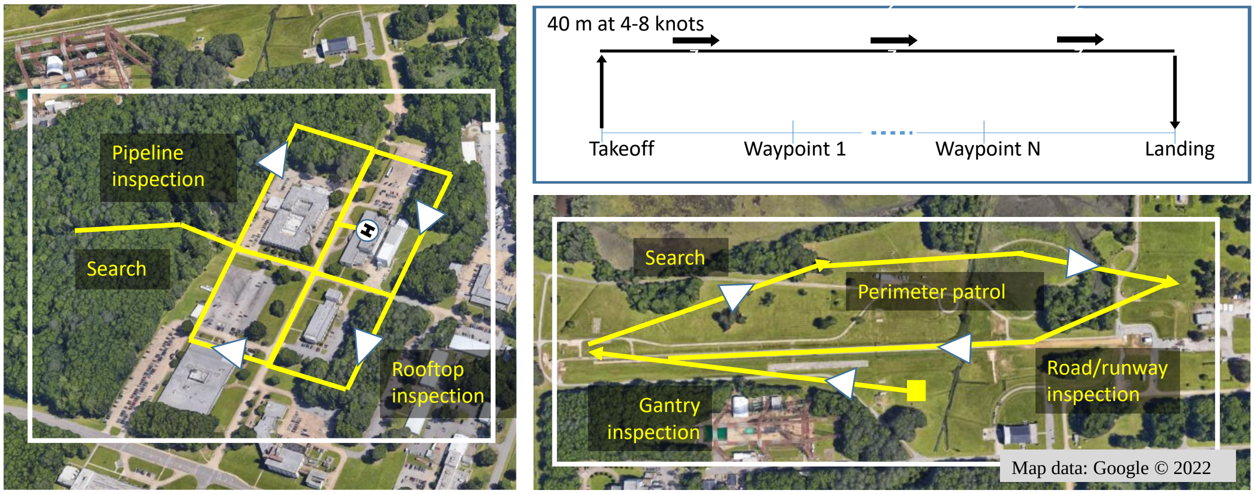
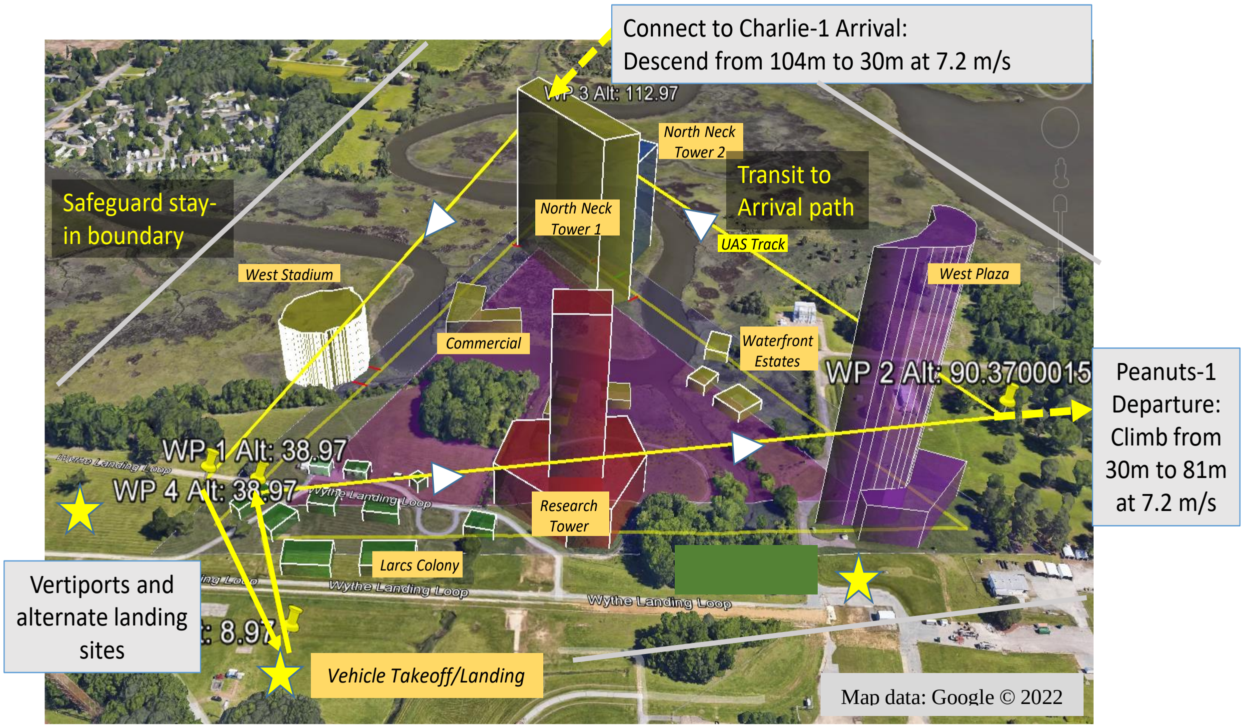


Fig. 13 Nominal reconnaissance scenarios (with traffic).
(right – range 1; left – range 2)



Map data: Google © 2022

Fig. 14 Nominal vertiport approach/departure scenario.

G. IVs/DVs, Measures, and Data Recording

Although the flight testing is more a validation activity than an experimental activity, there are a set of independent variables (IVs). Defining these helps to assure the three TRL criteria may be assessed while also supporting requirement analysis. These IVs are in essence the knobs we turn over the course of testing to see if expected cause-effect relationships occur. Performance and behavior are characterized by a set of dependent variables (DVs). The DVs are the metrics and measures that are recorded for each flight, or results of analytics that make use of the recorded data. IVs for the testing are:

- Mission/flight plan – Selectable from among the scenarios referred to previously (4 use-cases, 2 ranges)
- System configuration – A set of "builds" that constitute use of specific system and software versions
- SFC settings – A set of configurations that include which SFCs are enabled, alerting thresholds, etc.
- Multi-vehicle – Simulated air traffic movements may be enabled/disabled for each mission/flight plan
- Participants – The various crews that may evaluate procedures and human-machine interfaces
- Off-nominal condition – A set of conditions needed to reach states where SFCs perform their intended role in risk mitigation (e.g., moderate wind, rapid battery discharge, reduced engine power, and loss of link)

DVs fall into one of two categories: (1) the aircraft, system, and SFC behaviors; and (2) participant behavior/actions. For (1) a set of more than 200 parameters have been identified and are recorded (at a 1 Hz rate or higher). Data types correspond to the measures collected during the 2019-2020 testing; allowing for more direct comparisons in some cases. For (2), it was decided to use logged data and a playback capability for HF-related evaluations of human interfaces (e.g., displays). However, for the limited set of participants in the actual flight tests, post-flight interviews are planned for some tests later in the series. Audio/video recordings can also provide information on this aspect.

High-level measures of particular interest for the SFCs are: (1) Accuracy/Integrity – the extent to which outputs match simulation results or design expectations (e.g., errors due to bias and noise measurement uncertainty); (2) Timeliness – the extent to which outputs match simulation results or design expectations in terms of update rate, latency, and information loss frequency; and (3) Repeatability of outputs – the extent to which the outputs are consistent across repeated flights along the same trajectory and with similar environment conditions. Also of interest are information exchange measures across and among SFCs. These include: (4) Assurance – the extent to which information has been lost or corrupted during transmission from source to destination; (5) Completeness – the extent to which all data needed by a function is current and available when needed; and (6) Repeatability of exchange – the extent to which information exchange performance is consistent across repeated flights along the same trajectory and with similar environment conditions. Following completion of the testing, these measures will be determined using a set of quantitative metrics computed based on the recorded data.

H. Test Structure and Schedule

Testing is comprised of a series of stages; each building on prior stages and adding complexity. *Stage 0* – Research system integration testing; functional check flights; and procedures checkout (in a test Cage and on Range 1) (Oct-Dec, 2021); *Stage 1* – SFC testing, using nominal and off-nominal scenarios; without the onboard auto-mitigate capability engaged (Range 1 only) (Jan-Apr, 2022); *Stage 2* – Repeat Stage 1 testing with onboard auto-mitigate capability engaged for some flights (Range 1 only) (May, 2022); *Stage 3* – Repeat Stage 2 testing using Range 2 scenarios (on Range 2 only) (June-July, 2022). Note: Other than demonstrations, operator interface/display evaluations occur in parallel utilizing a playback capability of flight logs.

I. Summary of Testing to Date

At the time of submission of this paper, 44 flights have been completed and Stage 1 testing is complete. In general, the system and SFCs have performed as expected across the defined scenarios. To avoid duplication, preliminary results appear in companion papers submitted to the 2022 Forum [16, 29, 30, 37]. A comprehensive set of results will be published in the future once all testing is complete.

IV. Conclusion and Next Steps

Future advanced safety management systems seek to include more timely risk assessment and mitigation capabilities. These “in-time” capabilities may perform on the order of a few days to a few seconds; and may be supervised or fully automated. This paper summarizes the design and testing of such a system tailored to hazards associated with emerging flight domains. Several system SFCs are under investigation to complement or extend state-of-the-practice. Test results indicate feasibility in most cases; with a path exposed to higher TRL in others. Readers are encouraged to see preliminary results of ongoing flight testing in multiple companion papers at the 2022 Forum. Future work will apply the system concept to additional use-cases and platforms, tailoring particularly to highly-autonomous aircraft operating in challenging environments.

Acknowledgments

The authors would like to thank NASA’s Aeronautics Research Mission Directorate for its leadership, support, and sponsorship regarding the subject of this paper. In particular, the Associate Administrator, Robert (Bob) Pearce, Program Director, Akbar Sultan, and Project Manager, Misty Davies. They along with their teams established a strategic vision that recognizes safety assurance as one of the most difficult challenges facing an aerospace community that seeks to fly using increasingly complex, autonomous, and novel designs, but where hazards and risks may emerge in new and unexpected ways. Key contributors to the R&D and testing were Kevin Barnes, Ryan Condotta, Kaveh Darafsheh, Scott Dorsey, Ken Eure, John Foster, Sloan Glover, Alwyn Goodloe, Rafia Haq, Ed Hogge, Bailey Ethridge, Julian Gutierrez, Tom Johnson, Anne MacKenzie, Natasha Neogi, Truong Nguyen, Nick Rymer, Mike Scherner, Chris Thames, Andrew Turner, Sixto Vazquez, Portia Banerjee, Kevin Bradner, Matteo Corbetta, Vimmy Gujral, Chetan Kulkarni, Lynne Martin, John Ossenfort, Nikunj Oza, Lilly Spirkovska, Charles Walter, and Jason Watkins. Thank you all for your dedication, creativity, and can-do spirit.

References

[1] National Aeronautics and Space Administration, “Aeronautics Research Strategic Implementation Plan (2019 Update),” [online], URL: https://www.nasa.gov/aeroresearch/strategy/ , [retrieved 15 Apr 2022].

[2] National Academies of Sciences, Engineering, and Medicine, *In-Time Aviation Safety Management: Challenges and Research for an Evolving Aviation System,* National Academies Press, 2018. doi:10.17226/24962.

[3] Young, S. D., Ancel, E., Moore, A. J., Dill, E. T., Quach, C. C., Foster, J. V., Darafsheh, K., Smalling, K. M., Vazquez, S. L., Evans, E. T., Okolo, W. A., Corbetta, M., Ossenfort, J. P., Watkins, J., Kulkarni, C. S., and Spirkovska, L., “Architecture and Information Requirements to Assess and Predict Flight Safety Risks During Highly Autonomous Urban Flight Operations,” NASA/TM–2020220440, 2020.

[4] Kirkman, D., Mooberry, J., Reeser, R., Yang, M., Gould, K., Koelling, J., Davies, M., Ellis, K., Prinzel, L., Krois, P., Mah, R., and Infeld, S., "Informing New Concepts for UAS and Autonomous System Safety Management using Disaster Management and First Responder Scenarios," *IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)*, 2021, pp. 1-8, doi: 10.1109/DASC52595.2021.9594356.

[5] NASA Aeronautics Research Institute, "Advanced Air Mobility Project, Ecosystem, and Working Groups," [online], URL: https://nari.arc.nasa.gov/aam. [retrieved 15 Apr 2022].

[6] Federal Aviation Administration, "Urban Air Mobility Concept of Operations," Version 1.0, FAA, June 2020.

[7] International Civil Aviation Organization, Standards and Recommended Practices for Aeronautical Information Services, ICAO Annex 15, 16th edition, July 2018.

[8] International Civil Aviation Organization, Standards and Recommended Practices for Meteorological Services for International Air Navigation, ICAO Annex 3, 20th edition, July 2018.

[9] Hogge, E., Smalling, K., Quach, C., Eure, K., Kulkarni, C., Vazques, S., and Storm, T., "Battery Health Monitoring and Prognostics Framework for Flight Tests," Annual Conference of PHM Society, 2022.

[10] Eure, K., and Hogge, E., "Mathematical Characterization of Battery Models," NASA TM-2020-5008059, 2021.

[11] Watkins, J., Teubert, C., and Ossenfort, J., "Prognostics As-A-Service: A Scalable Cloud Architecture for Prognostics," Annual Conference of the Prognostics and Health Management Society, 2019.

[12] Ancel, E., Capristan, F. M., Foster, J. V., and Condotta, R. C., "In-Time Non-Participant Casualty Risk Assessment to Support Onboard Decision Making for Autonomous Unmanned Aircraft," AIAA Aviation 2019 Forum, American Institute of Aeronautics and Astronautics, Dallas, TX, 2019. doi:10.2514/6.2019-3053.

[13] Ancel, E., Capristan, F., Foster, J. V., and Condotta, R., "Real-time Risk Assessment Framework for Unmanned Aircraft System (UAS) Traffic Management (UTM)," Aviation Technology, Integration, and Operations (ATIO) Conference, American Institute of Aeronautics and Astronautics, Denver, CO, 2017. doi:10.2514/6.2017-3273.

[14] Hansen, A., Mackey, S., Wassaf, H., Shah, V., Wallischeck, E., Scarpone, C., Barzach, M. and Baskerville, E., Complementary PNT and GPS Backup Technologies Demonstration Report, Sections 1 through 10, No. DOT-VNTSC-20-07, John A. Volpe National Transportation Systems Center (US), 2021.

[15] Growing, J.T., "DOT holds first GPS backup technology demonstration," *GPS World*, Vol. 34, No. 4, April, 2020.

[16] Banerjee, P., Corbetta, M., Jarvis, K., Smalling, K., and Turner, A., "Probability of Obstacle Collision for UAVs in Presence of Wind," Aviation Forum 2022, American Institute of Aeronautics and Astronautics, 2022.

[17] Moore, A., Schubert, M., Fang, T., Smith, J., and Rymer, N., "LiDAR-derived navigational geofences for low altitude flight operations," Aviation Forum 2020, American Institute of Aeronautics and Astronautics, 2020.

[18] Verma, S. A., Monheim, S. C., Moolchandani, K. A., Pradeep, P., Cheng, A. W., Thipphavong, D. P., and Wei, B., "Lessons Learned: Using UTM Paradigm for Urban Air Mobility Operations," IEEE/AIAA 39th Digital Avionics Systems Conference (DASC), 2020, pp. 1-10, doi: 10.1109/DASC50938.2020.9256650.

[19] Craven, N., Verma, S., Monheim, S., Cheng, A., Wood, R., Dulchinos, V., Seah, C., Renema, F., Farrahi, A., Zahn, D. and Liddell, D., "X3 Simulation with National Campaign-Developmental Test (NC-DT) Airspace Partners," NASA TM-2021-0011098, 2021.

[20] Ellis, K. K., Krois, P., Koelling, J. H., Prinzel, L. J., Davies, M. D., and Mah, R. W., "A Concept of Operations (ConOps) of an In-time Aviation Safety Management System (IASMS) for Advanced Air Mobility (AAM)," AIAA Scitech 2021 Forum, American Institute of Aeronautics and Astronautics, 2021, doi:10.2514/6.2021-1978.

[21] Shoemaker, P. and Estes, C., "Accessible Telemetry Streams using a Zero Trust Architecture for the Flight Operations Directorate," [online], URL: https://ntrs.nasa.gov/citations/20210010837, [retrieved 8 Apr 2022]

[22] Moore, A., Schubert, M., Rymer, N.H., Villalobos, D., Glover, J.S., Ozturk, D. and Dill, E.T., "Volume Raycasting of GNSS Signals through Ground Structure Lidar for UAV Navigational Guidance and Safety Estimation," AIAA Scitech 2022 Forum, American Institute of Aeronautics and Astronautics, 2022.

[23] Dill, E., Gutierrez, J., Young, S., Moore, A., Scholz, A., Bates, E., Schmitt, K. and Doughty, J., "A Predictive GNSS Performance Monitor for Autonomous Air Vehicles in Urban Environments," In Proceedings of the 34th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2021) (pp. 125-137), 2021.

[24] Federal Aviation Administration, "UTM Concepts of Operations," v2.0, FAA, March 2, 2020.

[25] Spirkovska, L., Kulkarni, C.S., Watkins, J., and Martin, L., "Urban Air Mobility Airspace Dynamic Density," Aviation Forum 2022, American Institute of Aeronautics and Astronautics, 2022.

[26] Mackenzie, A., Nguyen, T., Barnes, K., and Scherner, M., “Simulated Versus Measured UHF Radiated Power at LaRC Street Level,” NASA/TM–20210013542, 2021.

[27] National Oceanic and Atmospheric Administration, National Centers for Environmental Information, “North American Mesoscale Forecast System,” [online], URL: https://www.ncei.noaa.gov/products/weather-climate-models/north-american-mesoscale [retrieved 26 March 2022].

[28] TruWeather Solutions, Inc., “TruFlite Core Specifications and API Instructions,” [online], URL: https://truweathersolutions.com/truflite-core/ [retrieved 26 March 2022].

[29] Feldman, J, Martin, L., Bradley, J., Walter, C., and Gujral, V., “Developing a dashboard interface to display assessment of hazards and risks to sUAS flights,” Aviation Forum 2022, American Institute of Aeronautics and Astronautics, 2022.

[30] Ho, N., Johnson, W., Wakeland, K., Panesar, K., and Ancel, E., “Preliminary Risk Visualization Concepts for NASA In-Time Aviation Safety Management System,” Aviation Forum 2022, American Institute of Aeronautics and Astronautics, 2022.

[31] Ho, N., Johnson, W., Panesar, K., Wakeland, K., Sadler, G., Wilson, N., Nguyen, B., Lachter, J. and Brandt, S., “Application of human-autonomy teaming to an advanced ground station for reduced crew operations,” IEEE/AIAA 36th Digital Avionics Systems Conference (DASC), 2017, doi: 10.1109/DASC.2017.8102124.

[32] National Aeronautics and Space Administration, “Core Flight System Background and Overview,” [online], URL: https://cfs.gsfc.nasa.gov/ and https://cfs.gsfc.nasa.gov/cFS-OviewBGSlideDeck-ExportControl-Final.pdf [retrieved 26 March 2022].

[33] Lowry, M., et. al., "Autonomy Operating System for UAVs: Pilot-in-a-Box", Aviation Forum 2017, American Institute of Aeronautics and Astronautics, 2017.

[34] Goodloe, A., “Challenges in High-Assurance Runtime Verification,” 8th International Symposium on Leveraging Applications of Formal Methods, Verification, and Validation (ISoLA 2016); Corfu, Greece, October 10-14, 2016.

[35] Neogi, N., Young, S., and Dill, E, “Establishing the Assurance Efficacy of Automated Risk Mitigation Strategies,” Aviation Forum 2022, American Institute of Aeronautics and Astronautics, 2022.

[36] Dill, E., Hayhurst, K., Young, S., and Narkawicz, A., “UAS Hazard Mitigation through Assured Compliance with Conformance Criteria,” AIAA Scitech 2018 Forum, American Institute of Aeronautics and Astronautics, 2018.

[37] Ancel, E., Young, S., Quach, C., Haq, R., Darafsheh, K., Smalling, K., Vazquez, S., Dill, E., Condotta, R., Ethridge, B., Teska, L., and Johnson, T., “Design and Testing of an Approach to Automated In-Flight Safety Risk Management for sUAS Operations,” Aviation Forum 2022, American Institute of Aeronautics and Astronautics, 2022.

[38] Consiglio, M., Muñoz, C., Hagen, G., Narkawicz, A., and Balachandran, S., "ICAROUS: Integrated configurable algorithms for reliable operations of unmanned systems," IEEE/AIAA 35th Digital Avionics Systems Conference (DASC), 2016, doi: 10.1109/DASC.2016.7778033.

[39] Mathews, B., Srivastava, A., Iverson, D., Beil, B., and Lane, B., "Space shuttle main propulsion system anomaly detection: A case study," *IEEE Aerospace and Electronic Systems Magazine*, Vol. 26, No. 9, pp. 4-13, September 2011.

[40] Northeast UAS Airspace Integration Research Alliance (NUAIR), “High-Density Automated Vertiport Concept of Operations,” NASA Contractor Report, Doc. ID 20210016168, May 27, 2021.

[41] Mercer, J., McSwain, R., and Ippolito, J., “Scalable Traffic Management for Emergency Response Operations (STEReO),” Aviation Forum 2020, American Institute of Aeronautics and Astronautics, 2020.

[42] Martin, L., Arbab, Y., and Mercer, J., “Initial Exploration of Scalable Traffic Management for Emergency Response Operations – System User Requirements for Safe Integration of Small UAS,” IEEE/AIAA 40th Digital Avionics Systems Conference (DASC), 2021.