Usability of an Updated Version of the Supplemental Data Service Provider-Consolidated Dashboard for Supporting Uncrewed Aircraft System Traffic Management

Gregory Costedoat[1]
San José State University, San Jose, CA 95192

Paul U. Lee[2], Jolene M. Feldman[3], Vimmy Gujral[4], Lilly Spirkovska[5]
NASA Ames Research Center, Moffett Field, CA 94035

Charles M. Walter[6]
ASRC Federal Data Solutions, LLC, Beltsville, MD 20705

The Supplemental Data Service Provider-Consolidated Dashboard (SDSP-CD) is a preflight planning user interface (UI) that serves to aid operators when drafting routes for small uncrewed aircraft systems (sUASs). The primary function of the SDSP-CD is to identify hazards that an sUAS may encounter along a proposed flight path and assess the severity of these risks. A usability study was conducted on an updated version of the SDSP-CD to determine if the most recent iterations made to the system improved objective performance and subjective user experience. There are two main components of the SDSP-CD interface: (1) the dashboard and (2) the interactive map. The dashboard provides users with hazard and vehicle limitations for each sUAS in their fleet while the map contains a graphical representation of each vehicle's route, hazard details, and geographic information. A series of preflight risk-assessment questions and tasks were developed to examine how participants interact with the updated version of the SDSP-CD. Additionally, a new service that measures vertiport congestion was developed and included as one of the services that was tested. In the present study, participants were trained to use the SDSP-CD and then completed two simulated scenarios during which they performed a variety of tasks, responded to questions, and completed surveys. The two scenarios developed for the present study were the Package Delivery and Hurricane Preparation scenarios. The Package Delivery scenario involved a fleet of four sUASs delivering low-stakes items (e.g., lunches and snacks) to people in a fictitious city. The Hurricane Preparation scenario involved a fleet of 11 sUASs delivering a range of supplies (from medicine to boardgames) to employees stranded at an office park due to road closures caused by an impending hurricane. Participants assumed the role of a fleet manager during both scenarios and were responsible for managing the sUASs in their fleet. Questions included those with objectively correct responses, open-ended strategy responses, and subjective user experience feedback. It was found that participants were largely successful at using the SDSP-CD interface to answer questions with objectively correct responses. Additionally, participants were able to use reasoning and logic based on the information available in the SDSP-CD to determine the cause of various risks and what actions they would consider taking. Finally, although participants reported that there were elements of the UI that could be improved, overall feedback pertaining to user experience suggested that the SDSP-CD concept is viable.

---

[1] Research Associate, Human Systems Integration Division.
[2] Research Engineer, Human Systems Integration Division.
[3] Research Psychologist, Human Systems Integration Division.
[4] Research Engineer, Human Systems Integration Division.
[5] Computer Scientist, Intelligent Systems Division.
[6] Senior Software Engineer, Human Systems Integration Division.

I. Introduction

The uncrewed and remotely piloted vehicle market is currently experiencing a period of monumental growth [1], with applications in domains ranging from research to public services, and commercial interests. Organizing and managing many small uncrewed aircraft systems (sUASs) operating in low altitude airspace requires the paradigm shift embodied by the UAS Traffic Management (UTM) system [2]. UTM relies on a federated set of services provided by competitive open-market service suppliers to enable cooperative management of operations. The notional architecture calls out two types of service suppliers: UAS Supplemental Data Service Providers (SDSP) that provide enhanced services including necessary data and constraint information and the UAS Service Supplier (USS) that supports operational requirements which enable safe and efficient use of airspace.

At NASA, the System-Wide Safety (SWS) project aims to develop a suite of hazard and risk assessment services for UTM operations [3, 4]. These services and tools collect and aggregate data from aviation systems and other data sources that operators can use for identifying hazards and risks when planning sUAS flight routes. These services are expected to function in the role of an SDSP and could include warnings about hazardous weather and wind avoidance, terrain and vertical obstruction warnings, poor Global Positioning System (GPS) reception areas, network interference areas, airspace conflicts, and population density indicators, among others.

Our team has been investigating the user interface (UI) requirements needed to display and organize hazard and risk assessment services in a preflight planning system. To accomplish this, we employed a user-centered design methodology. The intended user is a fleet manager, a person responsible for planning safe routes for a fleet of sUAS. The services have been integrated into a single dashboard called the Supplemental Data Service Provider-Consolidated Dashboard (SDSP-CD) that allows the fleet manager to view the routes of each vehicle in their fleet and assess risks and flight hazards. In previous studies [5, 6], the SDSP-CD interface was evaluated with five services – battery reserve, proximity hazards, population hazards, GPS coverage/availability, and radio frequency interference (RFI) – across different use scenarios. User feedback provided insights on the required information for assessing the potential hazards, suggestions for improvements, as well as response time measures and scores on perceived workload and usability.

In this paper, we describe a usability study of an updated version of the SDSP-CD that incorporates suggested improvements. It also has added a sixth service, vertiport congestion, that predicts potentially hazardous levels of traffic congestion at one or more vertiports based on preflight schedules [7]. The updated SDSP-CD tool with six hazard-related services was evaluated against a nominal, fair-weather scenario and an emergency scenario involving a severe weather prediction. In the following sections, we describe the six hazard assessment services and the SDSP-CD interface (see Section II), the study methodology (see Section III), and the study results (see Section IV). We then discuss our findings, implications for interface improvements, and recommendations for future research (see Section V). Finally, we highlight our main findings and future directions for the SDSP-CD (see Section VI).

II. SDSP-CD: User Interface, Services, and Tool Capabilities

The SDSP-CD services were originally envisioned and developed to provide hazard alerts to UAS pilots during live flights. However, the role of these services was later expanded to include preflight planning. The preflight version of the SDSP-CD tool is designed to be used by a *fleet manager*, an operational role in which one is responsible for planning routes for a fleet of UAS, analogous to the role of an airline dispatcher in commercial aviation. A fleet manager is expected to monitor all vehicles in their fleet and to be aware of environmental conditions that could impact scheduled flights. Additionally, the fleet manager is responsible for mitigating potential risks and hazards prior to departure (e.g., see Part 121 Subpart U regulations [8] for details regarding airline dispatcher responsibilities). The tool is envisioned to enable the fleet manager to quickly identify if any hazards are associated with a proposed route and take actions to mitigate any potential risks prior to taking off. The goal of the UI is to provide users with a comprehensive overview of potential risks and hazards associated with a proposed flight route using visual design strategies that serve to highlight hazards and assess the severity of the risks these hazards may present.

The SDSP-CD interface (see Fig. 1) consists of two main sections: the dashboard located at the top left of the display and an interactive map that makes up the space below the dashboard. It should be noted that the vehicle callsigns in the SDSP-CD interface are labeled as “UAV” (unmanned aerial vehicle) followed by a number (e.g., UAV6). The dashboard (see Fig. 2) provides a comprehensive, color-coded overview of the hazards that are present on each vehicle’s proposed route. Each row on the dashboard represents a vehicle in the user’s fleet and each column represents one of the six different services (with a scrolling window that can be used to accommodate larger fleets). Banded rows (a design strategy for lists in which one of two alternating colors are used for each row) were implemented on the most recent version of the SDSP-CD to aid users in being able to accurately identify the status of the alert icons associated with a particular vehicle. The color-coded alerts indicate the status of each service, with *green* representing that no safety thresholds have been exceeded, *yellow* representing that, although safe in its current state, the vehicle is approaching a safety threshold, and *red* representing that a safety threshold has been surpassed. The dashboard has a user settings menu that allows the user to upload vehicle information, set safety threshold margins from a predefined safe range for the alerts associated with each service, and filter the services they want displayed.

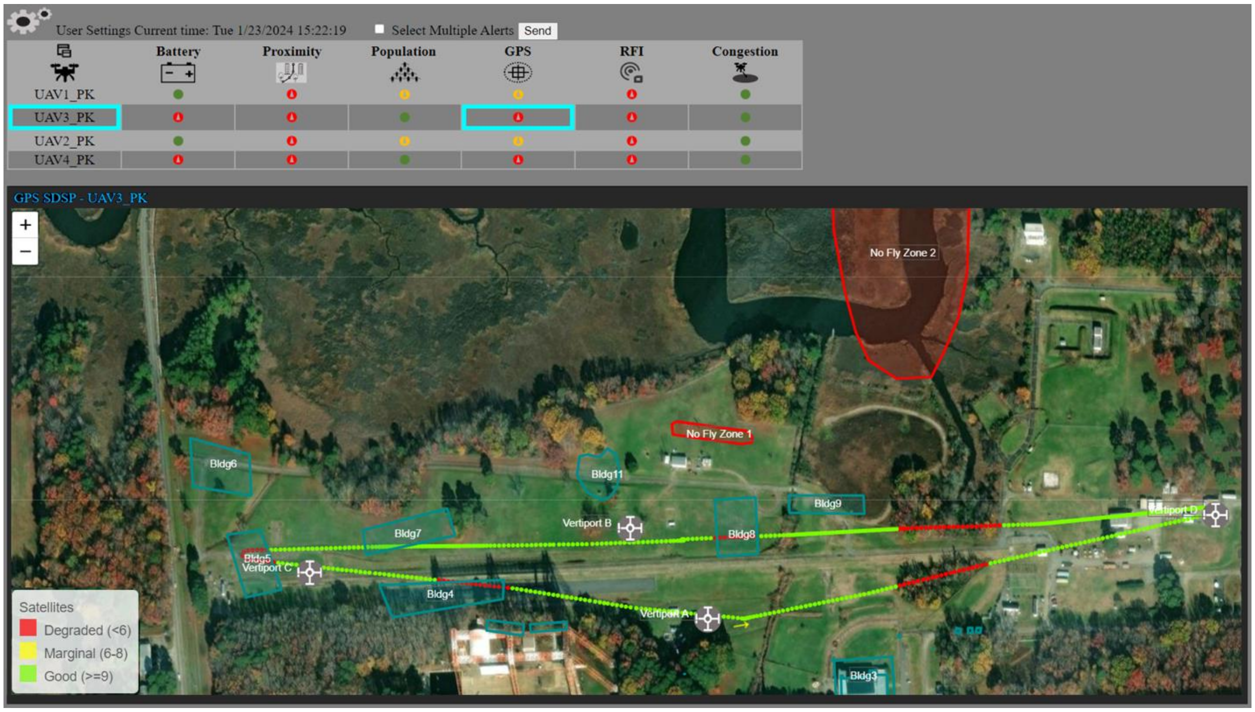


Fig. 1 Image of the current SDSP-CD interface (2024). The dashboard is located at the top left of the display. On the map, users can see the route of a selected vehicle and the locations along that vehicle's route where selected hazard types are located. The present image shows UAV-3's route and where it would encounter a degraded GPS signal. (NASA image)

Fig. 2 The SDSP-CD dashboard. Each row represents one vehicle in the user's fleet. Columns represent the status of each service. Banded rows (i.e., the alternating color pattern for each row) were implemented to aid users in identifying the statuses of each hazard for a particular vehicle. (NASA image)

The interactive map provides users with geographic information, hazard locations, and the ability to visualize their proposed flight paths. When an alert is selected on the dashboard, the location and severity of the hazard are highlighted along the flight's route, which is shown on the map (see Fig. 1). The user can then make actionable decisions based on this information. These decisions include, for example, altering the route to avoid hazards, cancelling the mission if they do not feel it can be safely flown, or keeping the route the same but communicating information about the risk to their remote pilots, among others.

The following provides an overview of the information displayed on the dashboard for the six available services (i.e., GPS signal strength, RFI, battery reserve, proximity to obstacles, population and casualty, and vertiport congestion).

GPS Signal Strength. This service computes GPS signal strength by calculating the number of satellites that the vehicle is able to connect to at each point along a proposed flight path. There are three signal strength designations, good (>8 satellites), marginal (6-8 satellites), and degraded (<6 satellites), which are mapped to the colors green, yellow, and red, respectively. The proposed route is shown as a series of color-coded dots (i.e., green, yellow, and red), indicating the GPS signal strength at each point (see Fig. 3). As shown on the flight path in Fig. 3, the route consists primarily of green dots (indicating that GPS coverage is good at those points on the route) mixed with a few sections of red dots (indicating where the GPS coverage is degraded).

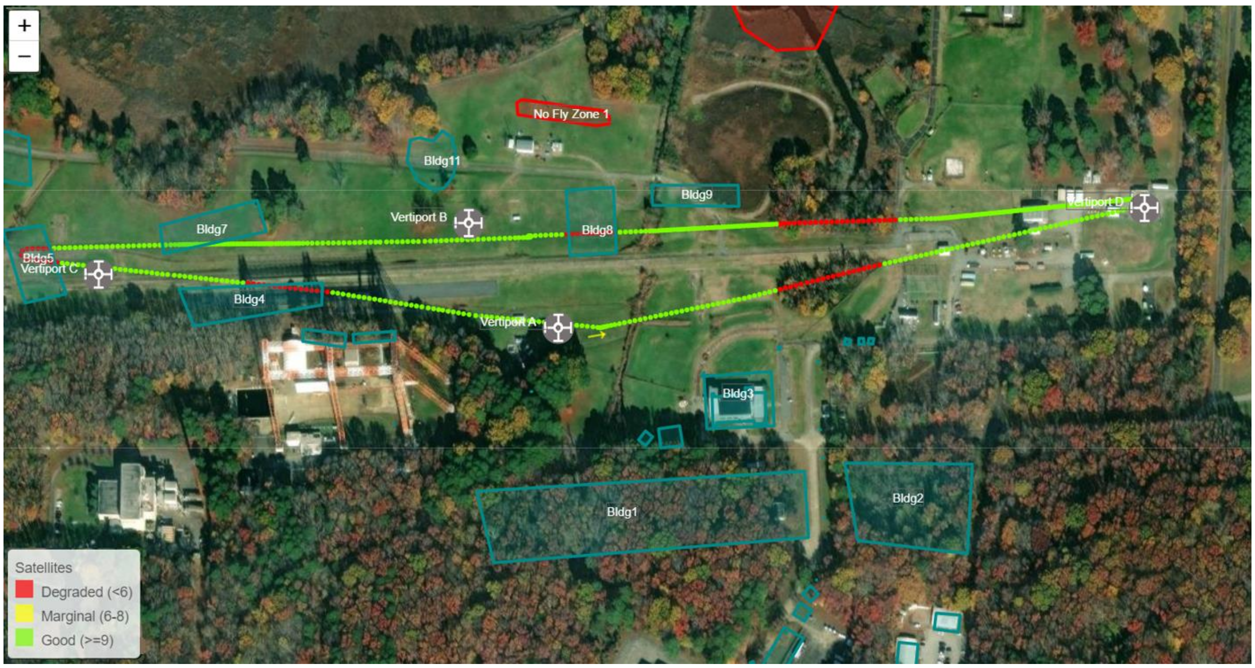


Fig. 3 Image of the SDSP-CD interface showing a proposed flight path with the GPS signal strength service shown on the interactive map. When the GPS service is selected, the route is illustrated as a series of dots, with each dot representing a point along the route. Each dot is color-coded to map to the three GPS signal strength designations (i.e., good, marginal, and degraded). A legend at the bottom left of the map provides users with the meaning of each signal strength designation. (NASA image)

Radio Frequency Interference (RFI). The RFI risk assessment service determines the likelihood of a disrupted signal between the user and their vehicle. Similar to GPS, the RFI service accounts for every point along a proposed flight path. However, on the interactive map there are four interference risk categories, none, low, moderate, and excessive, which are mapped to the colors gray, green, yellow, and red, respectively. Consistent with the GPS service, the route is displayed as a series of color-coded dots, with each dot representing the level of RFI that the vehicle will encounter at that point on the route. However, it should be noted that the dashboard maintains the green, yellow, and red color-coding scheme even when a route has no RFI (i.e., if a route has no RFI and is all gray on the map, the corresponding dashboard alert icon for that vehicle will be green).

Battery Reserve. The battery reserve service differs from the GPS and RFI services in how it communicates information to the user. As opposed to calculating the status of the battery at each point along the route, this service computes two values: (1) the location along the route where the battery charge will reach the user-specified reserve limit, and (2) the location along the route where the battery charge will reach operational depletion (which is typically set to a 30% charge to maintain long-term battery health). On the dashboard, a green alert indicates that the vehicle's battery will complete the route with an adequate charge remaining, a yellow alert indicates that the vehicle's battery has reached its reserve-limit, and red indicates that the vehicle's battery will experience operational depletion prior to completing the route. On the interactive map (see Fig. 4), the route is displayed as a solid line with the locations of where the battery reserve-limit and operational depletion are predicted to occur. On the route, the battery reserve limit is represented as a yellow battery icon and operational depletion (i.e., end of discharge; EOD) is represented as a red battery icon. Users can access additional details via an information table pop-up that provides the estimated time the reserve limit will be reached, the estimated flight time available, and the probability of completing the mission with adequate battery reserve. Finally, it should be noted that the SDSP-CD does not provide users with the current charge of the battery.

Fig. 4 Image of the SDSP-CD interactive map showing the battery reserve service for a vehicle's proposed flight path (in this example, the vehicle is flying clockwise on the route). The yellow battery icon indicates where the vehicle's battery will reach the reserve-limit that was set by the user and the red battery icon indicates where the vehicle's battery will drop below that threshold. The red EOD marker indicates that, with the current flight path, the vehicle would land with a battery reserve below the user's set threshold. The information table pop-up is located in the top left and can be moved by the user to any desired place on the interactive map. (NASA image)

Proximity to Obstacles. The proximity hazards service calculates the juxtaposition of static ground obstacles along a proposed route (specifically, the current version of the SDSP-CD assesses a vehicle's proximity to buildings and trees). Similar to the GPS and RFI services, sections of the flight path that violate a user-specified threshold are marked in yellow or red. Yellow markings indicate that the vehicle will pass by an obstacle with a distance approaching, but not exceeding, the user-specified required obstacle-proximity threshold, whereas red markings indicate the proximity will be closer than the specified threshold. When a vehicle with either a yellow or red proximity alert on the dashboard is selected, an additional information table appears on the interactive map. The information in the pop-up table includes the type of hazard (e.g., building or tree), time to hazard, safety margin, shortest distance between the flight path and obstacle, obstacle identification, whether threat is vertical or lateral, and ability to highlight the hazard on the map.

Population and Casualty. The population hazard assessment service determines the density of populations that may intersect with a proposed route and the probability of injury or casualty in the event of an unscheduled landing (e.g., if a vehicle in user's fleet experienced an uncontrollable failure). A unique feature of this service is the heat map (see Fig. 5) which is overlaid onto the interactive map and shows areas with sparse, low, medium, and high population densities (color-coded as white, yellow, orange, and red, respectively) as well as corresponding impact points.

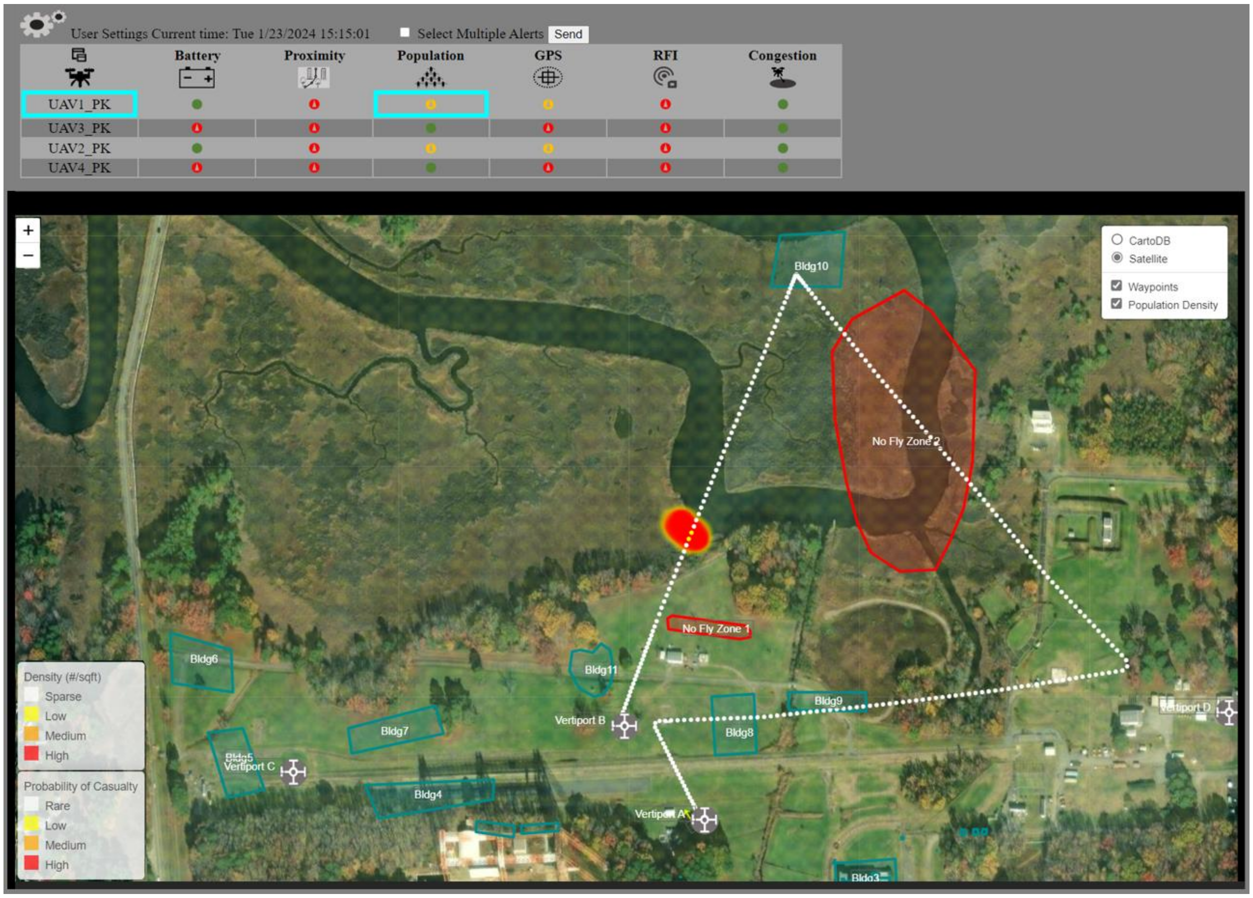
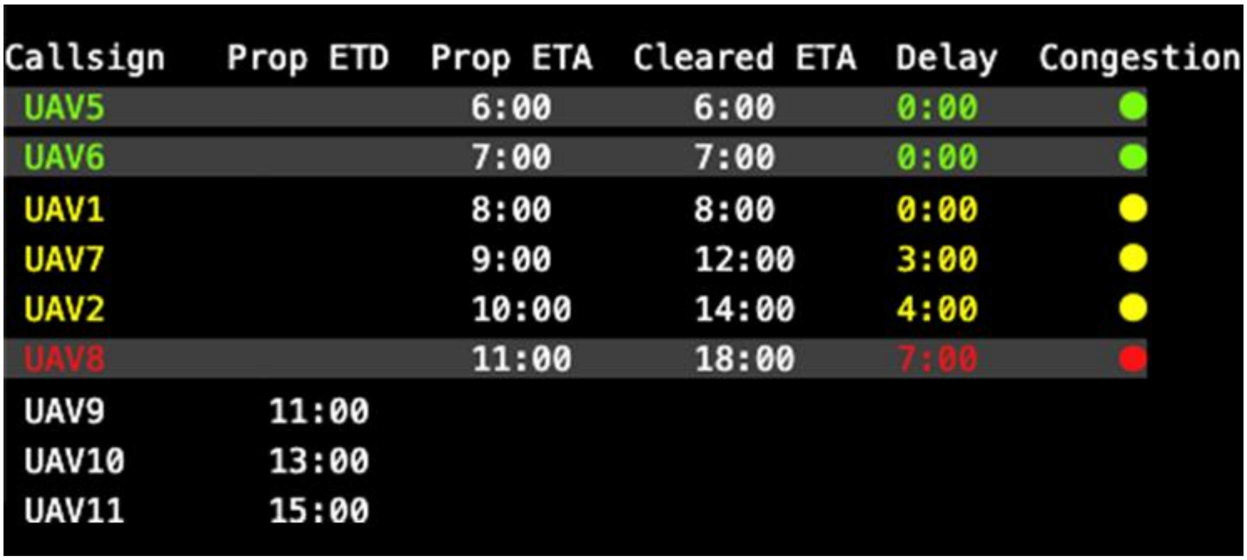


Fig. 5 Image of the SDSP-CD interface showing the population hazard service heat map. Each dot represents a point along the vehicle's proposed route. The white dots indicate that the population density is sparse, meaning that the probability of casualty is low. The yellow dots on the route (which intersect with the high-density population area) indicate that the probability of casualty is low to medium. (NASA image)

Vertiport Congestion. The vertiport congestion service, the newest service added to the SDSP-CD interface, determines whether the proposed arrival and departure flights would result in excessive delays for arrivals due to a lack of parking spots at the destination vertiport. In contrast to the previously discussed services that rely on information from a single flight to formulate a risk mitigation strategy, resolving congestion requires users to comprehensively examine the traffic flows for the congested time period. To support users in solving congestion issues, a flight schedule view was developed that shows the congestion risk of a selected vehicle as well as that of the surrounding vehicles that are landing at the same vertiport in temporal proximity to the selected vehicle.

The flight schedule view (see Fig. 6) provides users with information about predicted traffic congestion at a selected vertiport. For each vehicle departing or arriving a target vertiport, the flight schedule window shows vehicle callsign, either originally proposed estimated time of departure (*Prop ETD*) from the target vertiport or originally proposed estimated time of arrival (*Prop ETA*) into the target vertiport, estimated time of arrival reserved for the vehicle based on vertiport congestion (*Cleared ETA*), expected delay (*Delay*), and the severity of the congestion risk (*Congestion*). The expected delay and the congestion risks are color-coded to indicate the operational severity of the delay at the target vertiport for a given vehicle. For example, a delay of less than two minutes may be considered trivial (negligible or low/green), a delay between two to four minutes may be able to be accommodated (moderate/yellow), and substantial delays that are predicted but not mitigated during preflight may result in an increased risk of an adverse event inflight due to the vehicle's relatively small battery capacity (high/red), which would impact their ability to hold the vehicle in the air to absorb the necessary delays. The congestion information in Fig. 6 is based on study-specific constraints for the vertiport operations which is shown as the number of minutes of delay for a given flight and the category of the average delay (e.g., negligible, low/green, moderate/yellow, high/red) for each 3-minute schedule window.



Callsign	Prop ETD	Prop ETA	Cleared ETA	Delay	Congestion
UAV5		6:00	6:00	0:00	
UAV6		7:00	7:00	0:00	
UAV1		8:00	8:00	0:00	
UAV7		9:00	12:00	3:00	
UAV2		10:00	14:00	4:00	
UAV8		11:00	18:00	7:00	
UAV9	11:00				
UAV10	13:00				
UAV11	15:00				

Fig. 6 Image of the congestion information for a selected vertiport. Similar to the dashboard, each row represents a vehicle in the user's fleet. Columns include the vehicle callsigns, originally proposed ETD, originally proposed ETA, expected cleared ETA based on vertiport congestion, expected delay, and congestion risk severity. (NASA image)

Due to the complex nature of the congestion scenarios and the potential resolution options, SDSP-CD has provided automated resolution options (see Fig. 7) to assist the fleet manager with alleviating congestion problems. Mitigating actions generated by the congestion service include expediting (moving up the departure and/or arrival time), delaying (pushing back the departure and/or arrival time), cancelling, or redirecting a vehicle to an alternate vertiport. Users can also change the landing sequence of any two flights so that they swap positions (e.g., a flight with a high-priority mission that needs to land could take the landing slot of a lower-priority flight). The resolution options in the present study were generated manually prior to the study so that we could examine the desirability of automated resolutions that could be implemented later.

Fig. 7 illustrates an example of the resolution options window providing the user with four possibilities that could be selected to alleviate the congestion for UAV8. The options are ordered so that the option that results in the lowest overall levels of congestion appears at the top of the list. Each resolution option suggests an action (e.g., delay, swap) for a vehicle (e.g., UAV8), and auxiliary information about the action (e.g., number of minutes to expedite/delay). The display also shows the user how well the action will resolve the congestion for the vertiport schedule if updated as suggested, by providing the resulting delay/congestion for each 3-minute schedule window (shown by dashed lines (--) when the delay is completely eliminated, or indicated as low (Lo), moderate (Md), or high (Hi) if delays remain), as well as the residual delay that will be carried into the next 3-minute schedule window (shown by number of minutes of residual delay in parentheses).

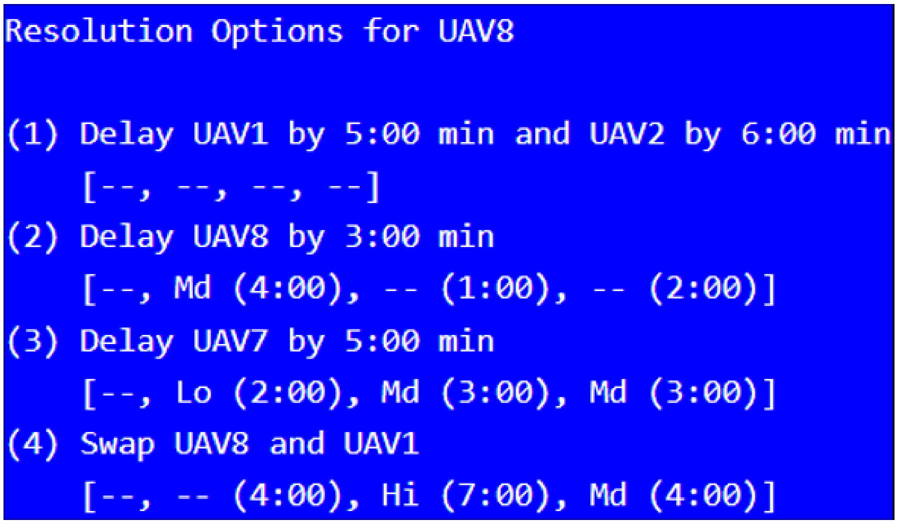


Fig. 7 The resolution options window for a selected vehicle, in this screenshot "UAV8". (NASA image)

III. Methodology

The usability study evaluated the SDSP-CD as a preflight flight planning support tool for sUAS fleet managers. Participants were tasked with evaluating the hazards potentially affecting a set of provided flight plans. They were trained with the interface, asked to respond to a variety of questions about the interface, and asked to make decisions regarding vehicles. Moreover, response time measures and scores on perceived workload and usability were collected. This section provides details of the methodology. This study had three objectives: (1) between-subjects comparison of the updated SDSP-CD and the previous version in the same scenario, (2) evaluation of the updated SDSP-CD on a more complex scenario as well as an additional service, and (3) evaluation of the display of congestion service information.

A. Participants

Sixteen participants completed the usability study. Participants were recruited externally and received monetary compensation for their participation. Participants were naïve to the research purposes and procedures. Participant ages ranged evenly across the different age groups (N = 5 for 18-24 years old; N = 4 for 25-34 years; N = 5 for 35-44 years; N = 2 for 45-54 years). There were 13 male and three female participants. All were right-handed. Two participants wore glasses. Most participants reported spending a low to moderate amount of time on the computer (12 reported less than 20 hours per week). For education level, 12 participants had a college degree or higher while the remaining four had other types of education, such as a high school diploma or aviation certification. Ten participants were educated in a field related to aviation, regardless of the education level.

All participants held an FAA Small UAS Rule (Part 107) remote pilot certificate. Eleven participants had less than 50 hours of sUAS flight time. Five participants flew sUAS only as a hobby. Six participants were also certificated general aviation pilots, including four who were also certified flight instructors. Twelve participants in total, including those who have private pilot license, have had general aviation experience, but none had air traffic control experience.

Since the focused on pre-flight planning by a fleet manager, participants were asked about their past dispatch experience. Although only one participant had dispatch experience, four participants have performed sUAS mission or preflight planning and flight testing, and five participants had experience in multi-aircraft flight control, most of which came from past participation in NASA studies related to multi-flight control. Ten participants have participated in past studies related to sUAS.

Most participants had occupations related to aviation. Three participants worked in the field of sUAS, either as a pilot, flight engineer, or other related fields. Four other participants worked in other aviation related fields (e.g., airline dispatcher) and three participants worked in fields other than aviation.

Finally, participants were asked to list which flight planning tools they use or have used. There were various tools that the participants mentioned, but only a few tools were mentioned by multiple participants. There were multiple mentions of the ForeFlight flight planning tool. Other tools were aviation weather products and apps that can provide airspace and terrain information.

B. Materials

The study was conducted in a small room with one participant and three research team members who were present to conduct the study. The room consisted of one desktop computer and one laptop for the participant to use, and a few other computers for the research team to conduct briefings, training, take notes, etc. Participants interacted with the SDSP-CD interface, which was displayed on one of two LED desktop monitors, and the UI and the user interactions were mirrored on a large wall-mounted TV for the observers. A Microsoft Teams meeting was initiated and recorded to show the SDSP-CD display and record user interactions.

The second monitor for the desktop was used to control the simulation software that ran the SDSP-CD and other supporting systems for the study. To the left of the UI, a laptop computer with a larger LED monitor was set up for the participant to type responses to various questions during the evaluation phase of the study. An Excel worksheet was set up for the participant to type their responses. The briefing and training slides were uploaded to a different computer, which had its display mirrored on a large wall-mounted TV to facilitate the briefings. Finally, each researcher used their own laptop to take notes.

Participants were also provided with paper maps for questions that required them to mark when responding to specific questions. Six screenshots were printed, three from the previous version of the SDSP-CD and three from the current version. There was a screenshot from each version in which the GPS service was selected, a screenshot from each version in which the battery service was selected, and a screenshot from each version where multiple services were selected. Participants saw both printouts side-by-side while an explanation was provided regarding what was different between the two versions. Participants were only told that they were two separate versions (of many) and not made aware that one was a previous iteration.

C. Scenarios

Two simulated scenarios were used. Scenario 1, *Package Delivery*, was identical to the scenario (Fig. 8) that was examined in the previous usability study [5, 6]. Briefly, participants were responsible for a four-vehicle fleet tasked with delivering non-essential items (e.g., lunches and snacks) to people in a fictitious city. This scenario addressed the first objective of the study, that is, analyzing how the updates made to the SDSP-CD interface, when compared to the previous version, affected the proportion of correct responses to objective questions, the type of subjective feedback that was received, the types of decisions that were made, the time it took to complete tasks, and NASA Task Load Index (NASA TLX) and System Usability Scale (SUS) scores. By replicating the Package Scenario from the previous study, direct comparisons were made to the updated version.

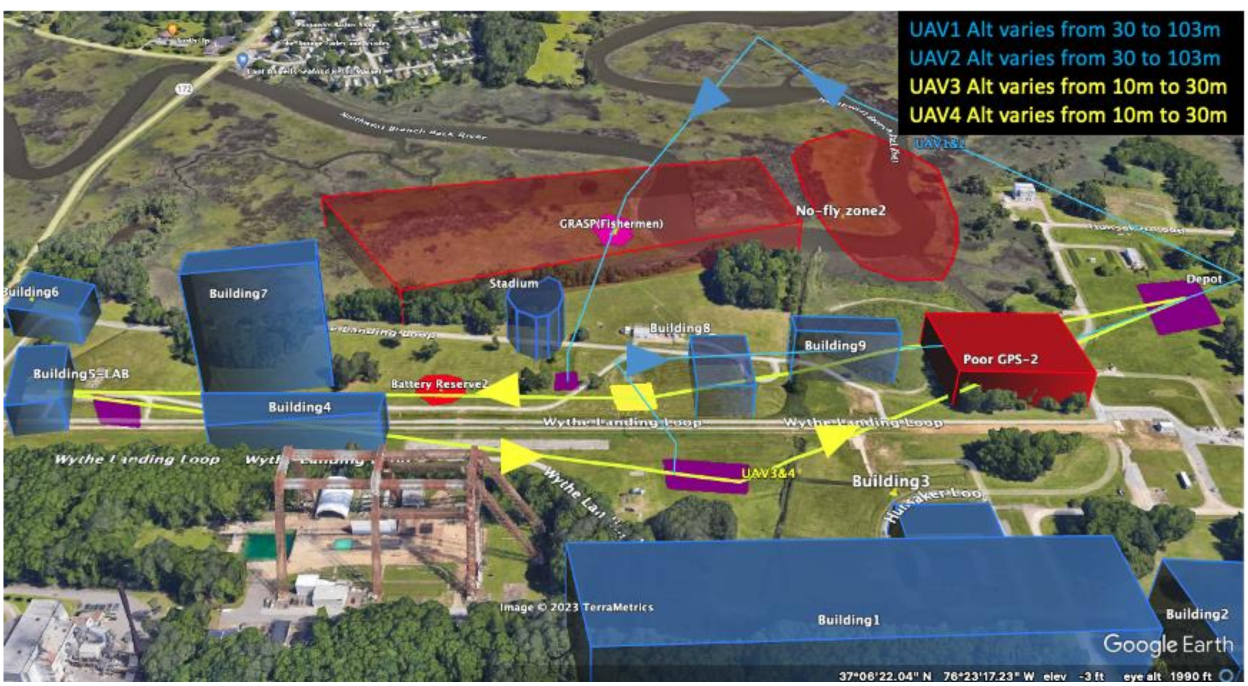


Fig. 8 **Package Delivery Scenario supplementary 3D map. (Maps Data: Google, Image ©2023 TerraMetrics)**

Scenario 2, *Hurricane Preparation*, explored how users interacted with the SDSP-CD interface when faced with an urgent, high priority scenario. This scenario was motivated by the increased role that UAS operations have taken on in other emergency situations, such as wildfire management [9, 10, 11]. In this scenario, participants were responsible for managing an 11-vehicle fleet tasked with delivering a wide range of items to people that were stranded (i.e., sheltering in place) at a fictitious worksite. In this scenario, a hurricane was projected to make landfall in an unspecified "short" amount of time, and an assistant rushed to equip the vehicles with the requested supplies from those who were stranded at the worksite. The packages that the vehicles were delivering ranged in criticality, from medical supplies to boardgames and lower priority items. Participants were asked to identify risks on the dashboard, locate hazards on the map, and make decisions regarding the operations of each vehicle. They responded to questions in a manner similar to the Package Delivery scenario and provided both objective and subjective responses.

Prior to the two test scenarios, the participants learned how to navigate the SDSP-CD tool in a training scenario. The location and mapped area for the training scenario was based on NASA Langley Research Center's "City Environment for Range Testing of Autonomous Integrated Navigation" (CERTAIN) Area 1. To prevent participants from seeing the same mapped area in both training and evaluation tasks, the location and mapped area for the test scenarios were based on NASA Langley Research Center's CERTAIN Area 2.

D. Procedure

The study consisted of three primary blocks, presented in the same sequential order for all participants: (1) a training session, (2) the Package Delivery scenario, and (3) the Hurricane Preparation scenario. It should be noted that a brief training session was held prior to the Hurricane Preparation scenario that provided participants with instructions on using the congestion service. Various tasks were developed for the study to query how participants used and interacted with the interface and to evaluate the use of the preflight planning services for the role of fleet manager. In each scenario, participants responded to questions, provided feedback, and filled out surveys.

During training, participants learned how to navigate the interface, how to read and make sense of the information being displayed, and how to access various tools and features. Participants received a series of briefings (slide presentations) on the interface and hands-on training in which they used the software in a structured manner with the assistance of the study staff. To successfully complete the training session, participants needed to demonstrate proficiency using the interface and were required to correctly answer questions about the displays. Throughout training, participants were able to ask any questions they had about the SDSP-CD.

Following the training session, participants completed the Package Delivery scenario. Participants were provided with details about the scenario and limitations for each vehicle. Information regarding the mission and the scenario were read to participants to ensure that they understood the scenario. Participants were asked to complete objective questions (e.g., "How many vehicles have a battery alert?"), respond to questions that required them to make decisions about different vehicles and the tasks that they were carrying out, and finally, respond to subjective questions regarding their experience using the interface.

After completing the Package Delivery scenario, participants were asked to compare the current UI ("New") with the prior instantiation used in 2023 study ("Old"). Still images of the two versions of the UIs were provided and participants were asked to rate which UI they preferred. Three different aspects of the UIs were presented to the participants, followed by an explanation of the differences and their ratings.

After the ratings, participants took a lunch break followed by the brief congestion training session. Finally, participants completed the Hurricane Preparation scenario, which was conducted using the same procedures as the Package Delivery scenario.

E. Data Collection

Both objective and subjective data were collected. *Objective data* were those in which a written or verbal response was required with a verifiable correct answer. For example, participants were asked how many battery alerts there were, for which a single answer was correct. The goal of these data was to ensure that participants were able to extract basic pieces of information from the interface.

Subjective data included any type of written or verbal response in which an opinion or judgment about the interface was being made. The goal of these data was to determine if there were trends in the subjective feedback that could help guide future iterations in the development of the SDSP-CD. Additionally, questions regarding direct comparisons to the previous version of the interface were also examined.

Preference ratings were collected to compare the current UI ("New") with the previous iteration that was used in the 2023 study ("Old"). Ratings were collected three times, highlighting different aspects of the UIs for each rating. Other ratings that were collected were the NASA TLX and SUS, which were administered to determine how participants evaluated the workload and usability of the updated SDSP-CD version and compared it to the previous version. These established scales provided quantitative data related to the perceived workload of using the interface and how usable various components of the interface were considered.

IV. Results

A. Performance, Accuracy, and User Experience

A set of questions were asked in this study that queried five aspects of user experience: (1) information retrieval, (2) reasoning based on UI comprehension, (3) usability, (4) perceived workload, and (5) decision making to mitigate risks. Both objective responses and subjective feedback were analyzed. To code the open-ended responses, a list of scoring criteria was developed for each question. After coding the open-ended response data, the raw frequency of each response was counted (unless the participant did not provide a response). Therefore, if participants gave multiple responses to a question (e.g., "what options do you have?"), it is possible that the raw number of responses could exceed the number of participants (i.e., if participants responded with multiple possibilities).

1. Information Retrieval

Information retrieval questions examined if participants were able to use the SDSP-CD to extract relevant pieces of information needed to answer the questions. Objective questions determined if participants could accurately use the UI to locate relevant information. Open-ended questions determined what types of information participants noticed on the displays.

Information Retrieval: Package Delivery Scenario. A series of objective questions, in which there was a simple, correct response, was administered to examine if participants were able to accurately retrieve information from the UI. All participants reported the correct responses to the objective information retrieval questions that were asked in the Package delivery scenario (see Table 1).

Scenario 1: Package Delivery

Questions		Correct (%)
1.	Using the dashboard, what alerts does UAV-2 have?	100%
2.	Using the dashboard, are there any proximity to threat alerts	100%
3.	Are there any RFI alerts? If yes, for which vehicles?	100%
4.	Report: is the battery sufficient for all vehicles?	100%
5.	Is Building 4 a proximity to threat for any UAVs?	100%
6.	Determine if a battery reserve risk will occur for UAV-1.	100%
7.	Which vehicles have a population risk on their flight paths?	100%
8.	From the dashboard and map information, describe the risks to your current flight plan for UAV-2.	100%
9.	Is building 6 presenting a PTT risk to any vehicles? If so, which vehicles?	100%
10.	Using the dashboard, quickly determine the total number of alerts that show values over the thresholds?	100%
11.	Quickly: Which vehicles have a GPS alert?	100%

Table 1 Objective information retrieval questions from the package delivery scenario. These questions required participants to locate and use information displayed on the UI.

To further assess participants' ability to retrieve information from the UI, we administered a series of paper map tasks and asked them to respond by marking the map (we used paper maps because circling directly on the UI display was not available). When asked to use the display to circle the location of a GPS hazard that was being displayed on the UI, 81.3% of participants circled the correct location on the paper map corresponding to the GPS hazard located on the interface. When asked to identify and circle where there was a population hazard, 93.8% of participants correctly identified the risk on the map. Finally, we identified a vehicle and asked participants to determine where it would reach a battery alert and to identify at what distance this alert would occur. All participants responded correctly. Finally, we asked why there was a population alert for UAV-1 and UAV-2, to which 93.8% of participants recorded an accurate response (responses were considered correct if participants mentioned the specific type of population hazard or why the population hazard appeared on the display).

Information Retrieval: Hurricane Preparation Scenario. For objective information retrieval questions asked in the Hurricane Preparation scenario, participants generally performed well, but a performance decline was seen compared to the Package Delivery scenario (see Table 2). Out of six questions, participants responded to two of the questions 100% correctly, three other questions resulted in >85% of participants responding correctly, and one question resulted in 73.3% of participants responding correctly. Questions in the Hurricane Preparation scenario that asked about the alerts for a specific vehicle tended to receive more correct responses compared to questions asking about the number of total alerts (e.g., Question 2).

Scenario 2: Hurricane Preparation

User Interface Questions		Correct (%)
1.	What alerts does UAV-6 have?	100%
2.	How many proximity to threat alerts are there in total? Specify the number of red alerts and the number of yellow alerts.	73.3%
3.	How many alerts does UAV-7 have? Specify how many red alerts and how many yellow alerts.	100%
4.	List all of the alerts that UAV-8 has. Specify the number of red alerts and yellow alerts.	93.8%
5.	Using the dashboard, quickly determine the total number of alerts that show values over the thresholds?	87.5%
6.	Quickly: Which vehicles have a GPS alert?	93.8%

Table 2 Objective information retrieval questions from the Hurricane Preparation scenario.

2. Reasoning based on UI Comprehension

We administered a series of questions that required participants to use reasoning and logic based on the information available in the SDSP-CD to determine the cause of various risks and what actions they would consider taking. The reasoning questions were relatively more complex than the questions that assessed performance with information retrieval.

Reasoning: Package Delivery Scenario. Participants were asked how they could get more information about the population alert for a specific vehicle, of which 87.5% provided a correct response, which consisted of clicking on the alert button on the dashboard (16) and/or finding the information on the map (2) and use the drilldown icon that was used to toggle the population density and the population hazard risk probability information on the map (1).

To assess how participants made novel reasoning decisions based on their UI comprehension, we asked them to determine what they would expect to see in a novel, unplanned situation in which multiple ground vehicles began to park along one of the proposed routes. The most frequent response was that this would result in a population alert

(13), followed by a proximity alert (3), RFI alert (1), and no alerts (1). Although there were no absolute correct answers, predictions about population and proximity alerts seemed to be logical if they reasoned that ground vehicles and/or the people inside the vehicle created new proximity and/or population hazards, respectively. No alerts could also be a reasonable answer if the participant assumed that neither the vehicles nor the number of people that are moving dynamically were substantial enough to trigger the hazard alerts.

When asked to determine why certain vehicles had a marginal GPS signal while others had a poor GPS signal, 93.8% of participants successfully reported that the vehicles with the poor GPS signal were flying at a lower altitude and that there was unfavorable terrain (i.e., buildings and trees) along their routes compared to the vehicles with the marginal GPS signal.

Open-ended questions also provided insight into the kinds of reasoning participants were using. We asked participants to report what they noticed about two vehicles (both of which were on the same route). Participants largely noticed that the two vehicles were on the same route (14), followed by noticing that the two vehicles encountered the same hazards (9), followed by noticing that the two vehicles both had the same number of red and yellow alerts (1). Such insights could be helpful in developing a common mitigation plans for multiple vehicles with similar situations.

Reasoning: Hurricane Preparation Scenario. To examine how participants believed the UI would function, we asked them to predict what would appear on the dashboard if UAVs that were not a part of their fleet were to unexpectedly park at some of the local vertiports. The most frequent response was that this would affect the congestion service (14). Participants also predicted that the dashboard would display information about these UAVs (3), a proximity alert (1), a population alert (1), a battery alert (1), and that the dashboard would display updated cleared ETA times (1). Based on the information that was given, the responses that the congestion service would be impacted and the information about the new UAVs, along with updated cleared ETA times seem to be valid outcome of the situation.

We asked participants if moving one of their UAVs (i.e., UAV-1) would relieve the congestion alerts for two other UAVs in their fleet (see Table 3). The word “moving” was purposefully vague so as not to lead participants.

Scenario 2: Hurricane Preparation

Response	Reasoning (of participants who provided a reason)
Yes (10)	Based on the info provided by the congestion resolutions (5)
	Delaying UAV-1 allows UAVs 2 and 7 to take earlier (3)
	UAV 1's cargo will not be compromised by delay (1)
	Moving UAV-1 results in one less vehicle at Vertiport B (1)
No (6)	Moving UAV-1 still results in delays for UAVs 2 and 7 (2)
	UAV-1 does not conflict with the ETAs for UAVs 2 and 7 (2)
	UAV-1 has zero delay, so moving it won't affect UAVs 2 and 7 (1)
	UAV-1 is not in the same time chunk as UAVs 2 and 7 (1)

Table 3 Participant responses to whether moving one of the UAVs in their fleet would relieve congestion for two UAVs with a congestion alert.

Although the question was open-ended and there were no pre-determined “correct” answers, it was designed to elicit the participants’ understanding of the vertiport congestion problem, which consists of traffic demands to the vertiport, exceeding the vertiport capacity, and causing arrival delays before the vehicles could land. Therefore, “removing” one vehicle from a congested timeframe would actually reduce the delays and thereby congestion to the surrounding traffic.

Based on that logic, it seems that the participants who answered “Yes” seem to provide reasonable answers. They either stated that they would follow the resolution options that was provided for them, which would work, or removing UAV-1 from the congested timeframe would help UAV-2 and UAV-7’s congestion problem. In contrast, the participants who answered "No" seem to have missed understanding that removing UAV-1 would help UAV-2 and UAV-7. Two participants who answered that moving UAV-1 still results in UAV-2 and UAV-7's delays seem to be "correct" answers if participants were stating that some delays remain for the two vehicles, even after moving UAV-1. More training might be needed for those participants to understand the nuances of the vertiport congestion problem.

B. User Interface Evaluation

The results from the previous section suggest that the SDSP-CD interface was easy to understand and use when retrieving hazard/risk related information, reflecting favorably towards the general UI layout and design. However, the results also suggested that it could be somewhat challenging to predict what information would be needed for reasoning in hypothetical situations and whether the UI provides sufficient information to reason about novel situations.

In this section, the SDSP-CD interface is evaluated using various ratings scores for direct UI comparisons against a prior UI version, SUS scores that analyzes the UI's usability, and NASA TLX scores that measures participants' perceptions of workload and performance while completing tasks with the SDSP-CD.

1. Evaluation of Interface Updates Based on Previous User Feedback

The SDSP-CD has been iteratively improved over the past few years based on user feedback. Continuing this development approach, the SDSP-CD for this study was updated based on the findings from 2023. To confirm and validate that the UI has indeed improved, participants were shown the previous ("Old") and current ("New") version of three comparison images, one at a time (i.e., one comparison per trial for a total of three trials) and asked to rate their preference. The order of the presentation of the Old vs. New UI were counterbalanced across participants, and the ratings ranged from 1 to 5, where 1 = preference for UI version A and 5 = preference for UI version B with A and B representing different UI versions based on counterbalancing. Ratings scores were adjusted and shifted to a -2 to 2 scale, such that -2 = preference for the Old UI and 2 = preference for the New UI, with 0 = no preference to either UI. Also, the ratings scores were averaged across the three trials per participant.

A one sample t-test was conducted on the UI comparison ratings and the preference for the New UI was significant, $M=1.31$, $t(15) = 8.38$, $p < 0.001$. Fig. 9 shows the results of the UI comparison ratings. Out of 16 participants, individual's average ratings across three trials suggested that all but one participant preferred the New UI over the Old UI, confirming that the UI improvements made since the last study succeeded in improving the user experience.

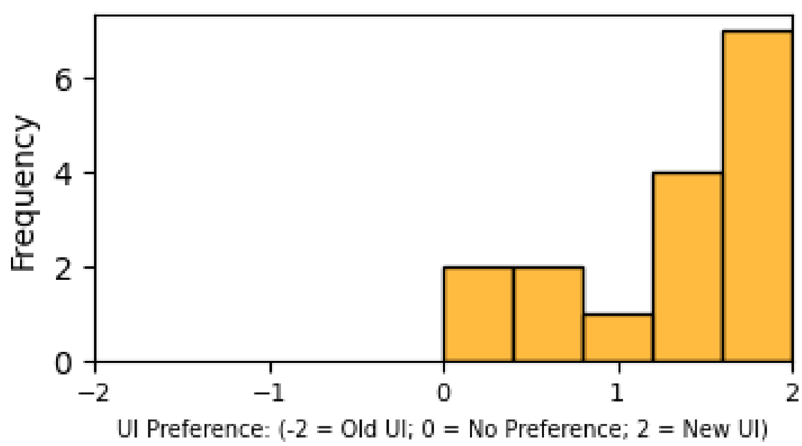


Fig. 9 Histogram of UI Comparison Ratings

In addition to the preference ratings, open-ended feedback on the UI comparison was also solicited. When their preferences were recorded along New vs. Old UI and asked for the reasons, the general feedback was that the New UI in general had better visibility, color choices, layout, and functionality.

Some of the features that improved UI visibility was the use of "banded rows" in the dashboard, each row of the dashboard that highlights all six services for a given UAS to be more visible from the surrounding rows and easier to read. Participants appreciated some of the other details, such as clearer markings of the hazard severity levels that improved their visibility and could be seen more clearly in the New UI.

They commented that the New UI made better color choices. Using green colored routes that had no hazard alerts better matched the risk colors, and the titles on the dashboard were clearer in the New UI. They liked the color coding of the risk assessment on the dashboard that did not change colors when selected, which was the prior implementation in the Old UI.

There was a strong preference for the New UI's multi-selection functionality, clear hazard severity visibility, and grouped table format. Participants liked how hazards were differentiated by severity rather than type, found the red text and banded rows in pop-ups easier to read, and felt that the red color emphasized urgency effectively. The enhanced readability of the information in the pop-ups was also noted as an improvement in the New compared to the Old UI.

They also appreciated the new UI layout and improved interactions with the functionalities. Participants thought that the general selection methods for vehicles and services were easier, as well as the way that the alerts and the UAS were highlighted. They also valued the more detailed information in the header in the pop-up boxes in the New UI.

Finally, participants preferred the overall enhancement of comprehensive battery life prediction information. The New UI's detailed battery information, including icons and warning markers, was also praised for its clarity and accessibility. Participants valued the ability to see battery icons along the route, the visibility of where battery levels exceed thresholds, and the inclusion of multiple warning markers before the EOD.

A small minority of participants preferred specific UI components in the Old version, with 2 out of 16 favoring its icons and selection methods and 3 out of 16 liking its color-coding strategy for hazards, font, and pop-up grouping. Some participants suggested incorporating a blend of both versions' color-coding strategies for improved visibility and alert severity differentiation.

2. *Evaluation of Usability with SUS Surveys*

Three SUS surveys were administered for each participant to evaluate the usability of different aspects of the SDSP-CD. First, under "Package" UI assessment, participants rated the usability of the UI for the five hazard alert services (minus the congestion service) used during the simple 4-vehicle Package Delivery scenario. Second, under "Congestion" UI, participants rated the usability of only the newly added interface for the vertiport congestion service used during the more complex 11-vehicle Hurricane Preparation scenario. And third, under "Hurricane" UI, participants rated the usability of the comprehensive interface with all six hazard services, also used the Hurricane Preparation scenario.

Fig. 10 illustrates the SUS scores for s. The results showed that the usability of the interface for the congestion service was rated the lowest at 55.3 (out of 100). The usability of the comprehensive interface for the high-traffic Hurricane Preparation scenario was rated slightly higher at 65.6. Finally, the usability of the interface for interacting with only the five services (minus the congestion service) used in the low-traffic Package Delivery scenario was rated the highest at 71.7. In general, a SUS score higher than 68 is rated as a "good" UI in terms of usability.

To assess if these three SUS scores were significantly different from each other, a within-subject, repeated measures ANOVA test was conducted. The results were significant, $F(2,30) = 5.38$, $p < 0.02$, suggesting that the SUS scores between the three conditions were significantly different.

Based on the significant ANOVA results, a post-hoc, pair-wise t-test was conducted to compare each pair of conditions. Since the test was post-hoc, the p-value was adjusted accordingly using Bonferroni-Holm method, which is a variation of a step-down Bonferroni. These pairwise t-tests showed that the UI for the congestion service was rated significantly worse than both of the UIs for the five other services in the Package Delivery scenario ($T(15)$ - -2.71, $p < 0.05$) and the UI for all six services in the Hurricane Preparation scenario ($T(15) = -2.52$, $p < 0.05$). In contrast, the UIs for the Package and Hurricane scenarios were not significantly different from each other. $T(15) = -1.36$, $p > 0.19$.

3. *Evaluation of Workload with TLX Assessment*

The perceived workload in using the SDSP-CD was evaluated twice in each scenario, for a total of four times per participant. During each scenario, participants were asked a series of questions, but some of those questions were targeted specifically to get participant assessment of the TLX scores. During the question-and-response sessions, participants were asked to rate the task load they experienced answering a set of three targeted questions on the TLX scales, omitting the irrelevant Physical Demand (e.g., physical body movements such as pushing or pulling) dimension. Two types of questions were asked in each scenario. The first set of questions focused on using the UI for simple information retrieval tasks, whereas the second set of questions focused on asking participants to brainstorm risk mitigation actions or strategies based on the risk alerts presented on the UI.

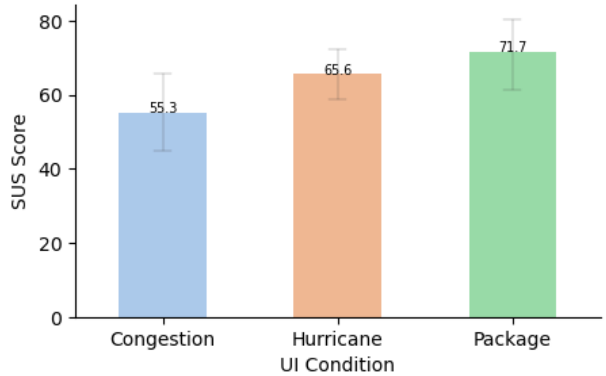


Fig. 10 SUS Scores for UIs for Package, congestion service only in the Hurricane Preparation scenario, and the comprehensive UI in the Hurricane Preparation Scenarios.

The tasks in the first set of questions involved assessing how well the participants understood the UI components and how quickly and effectively they could retrieve the correct information. Some of the questions were timed, creating a temporal pressure in the tasks. Fig. 11 shows the average TLX scores (averaged over the two scenarios because they both followed the same pattern) for this first set of questions across the five dimensions: mental demand, temporal demand, performance, effort, and frustration. The results suggest that task load along mental demand, effort, and frustration was moderately low while the performance was moderately high (M = 34.4 where 0 = high performance and 100 = low performance). Only the temporal demand was above 50%, which is expected since some of the tasks were timed.

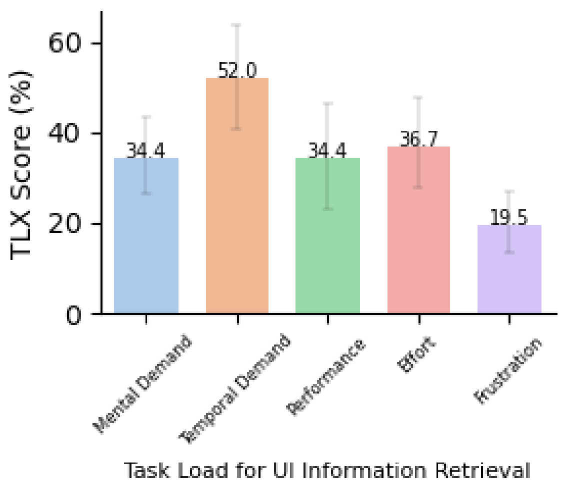


Fig. 11 Average TLX scores for Information Retrieval Focused Tasks (on 0-100 scale; Performance score is inversed: 0 = high performance; 100 = low performance).

The second set of questions (in each of the Package and Hurricane scenario) tasked participants with brainstorming mitigation actions to resolve the hazard risk alerts. The current version of the UI does not provide tool support for suggesting risk mitigation actions, except for the congestion service; therefore, these tasks required a greater mental task load for the participants to answer. Fig. 12 confirms this supposition. The average TLX scores were moderate for mental demand (M = 54.8) and effort (M = 49.2). The performance was worse (M = 40.6; 0 = high performance; 100 = low performance) than the scores for the information retrieval tasks. Meanwhile the temporal demand was moderately low (M = 30.6), as the participants were given ample time to perform their tasks. Even though coming up with risk mitigation actions or strategies were not easy, they still rated frustration as moderately low (M = 35.8).

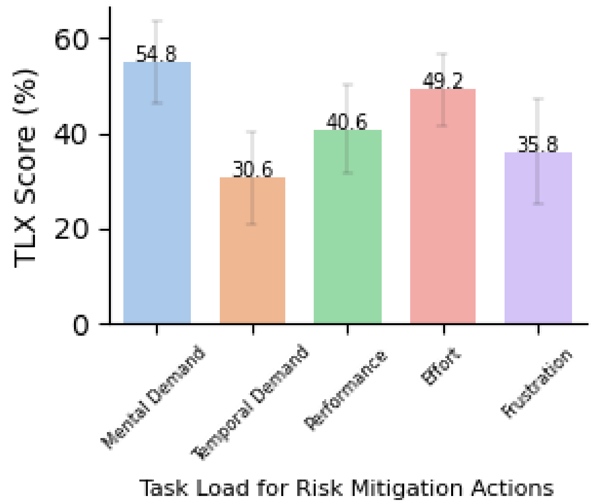


Fig. 12 Average TLX scores for Risk Mitigation Actions Tasks (on 0-100 scale; Performance score is inversed: 0 = high performance; 100 = low performance).

C. User Interface Feedback

After each of the two scenarios, participants were asked subjective questions regarding their experience using the interface. In this section, we summarize their feedback about the overall effectiveness of the UI, whether it provides the necessary information and what additional information would be useful, whether it provides the desired functionality and what additional functionality would increase its effectiveness, and finally, the UI layouts and display features that were helpful and those that could be improved.

1. UI Overall Effectiveness

Participants stated that the SDSP-CD had the capacity to enhance risk management and safety by allowing better planning, reducing the chances of collisions, and providing detailed insights into hazards, battery life, and GPS signal integrity. They suggested that it supported faster decision-making and fewer surprises during flights, ultimately increasing operational efficiency and safety, and facilitating better fleet management.

Participants unanimously agreed that the minimal additional preparation time is well worth the substantial increase in safety. They emphasized that prioritizing safety, even at the cost of added prep time, is crucial, especially in non-stressful, non-time-constrained situations. The ability to quickly identify safety hazards on a broad level, with a color-coded system indicating the severity of threats, allows for a rapid assessment of whether a route is safe or unsafe. Overpreparing, particularly with sUAS, is essential as it ensures that potential hazards are identified before takeoff, when more mitigation options are available. The feedback highlighted that the safety of both sUAS and people along the flight path should always be of utmost importance.

Feedback on the UI reveals a mix of concerns and positive insights. Many users believe the tool enhances safety in preflight planning, although some worry about confusion and errors due to misinterpreting information and forgetting vertical components. Usability issues include difficulty in manipulating routes and identifying specific waypoints. Timeliness and accuracy of updates, particularly regarding No-Fly Zones (NFZ), weather, and other aircraft, are critical, as are potential operational risks like GPS signal issues and unexpected obstacles. Suggestions included comprehensive preflight planning, including checklists, essential to mitigate these risks and emphasizing the importance of training and familiarity with the tool to ensure its effective and safe use.

On the additional workload that might be imposed by the integration of risk assessment process into their pre-flight planning tasks, participants' opinions were divided regarding whether the workload would increase or decrease with the implementation of the UI or software tool. Some believed that the UI would decrease workload by providing all necessary information in one place, making decision-making easier, and streamlining processes. Others expressed concerns that the UI could increase workload due to issues such as a high learning curve, incomplete information, or the need to address alerts and risks identified by the system. Ultimately, the impact on workload may depend on factors such as the user's familiarity with the UI, the comprehensiveness and usability of the tool, and the specific tasks being performed. However, the consensus appears to be that while there may be some initial increases in workload, the long-term benefits in terms of efficiency and safety could outweigh these challenges.

Finally, participants compared this UI with other tools such as ForeFlight and Vigilant Spirit, with mixed reviews. Some users found the UI simple and intuitive, while others felt it lacks necessary details. There was a desire for more customization based on mission-specific needs. Users appreciated the overall concept but saw room for significant improvement in details and functionality. The feedback on the interface highlighted both an appreciation for unique features like color coding and pop-up alerts, as well as suggestions for improvement, including clearer information on battery status.

2. UI Information Requirements

There were mixed opinions on the amount of information provided by the UI. Some participants felt that there could be slightly more information, requesting additional details like "time remaining" and "range remaining" for battery. Others believed that there was enough information for specific scenarios like package delivery or congestion, but suggested enhancements like adding priorities, color coding, and more overlays to the map, including 3-D views and building altitudes. Some participants suggested real time updates for route changes showing the impact of changes on ETA, battery usage and flight safety and display detail metrics like distance to the next waypoint and total route distance. They also recommended customization options to display relevant information and flexible adjustments reflecting immediately on UI. Some suggested moving the User Settings icon to the right-hand side of the UI so that the dropdown menu of the User Settings does not occlude important information on the dashboard. Overall, while the amount of information was generally considered adequate, there were calls for more detailed, real-time data and enhanced UI features to support better decision-making and planning.

The UI provided essential information for safe flight planning but needed enhancements like weather data and UAV performance metrics. Despite minor issues, it effectively organized risks and provided valuable visual feedback. Overall, the UI supported informed decision-making and risk assessment, but additional refinements were needed to handle all alerts and offer comprehensive data.

3. UI Functionalities

In terms of UI functionalities, participants generally did not have trouble finding what they needed within the tool's constraints. The ability to select multiple alerts and see them on the map was praised, and the resolution options provided for the Congestion service were found to be particularly useful for decision-making.

However, they also suggested several improvements and/or additions to the functionalities. They desired features like the ability to reschedule flights based on time, prioritize flights, and see wind data and building altitudes. Enhancements such as auto-suggested routes, detailed proximity alerts, and detailed information about NFZs were also requested. Issues with sorting aircraft and visualizing GPS coverage and battery life were noted. Users appreciated the comprehensive information provided but wanted more intuitive interfaces, simulation capabilities, and the ability to make route adjustments and see the impacts in real-time.

4. UI Layouts

Participants offered feedback on the UI layouts and how information was displayed. They liked and praised the intuitive nature of the dashboard for managing fleet operations, as well as the various visual representations of the map. Specific interface features that received positive feedback include the following:

- Consolidating all pertinent sUAS information in one place.
- The comprehensive dashboard facilitates making informed decisions, quickly recognizing issues, saves time, improves efficiency, and is likely to reduce human error.
- Graphical/visual representation of missions and routes, which simplifies understanding and enhanced situational awareness by graphically displaying potential issues and hazards, enabling pre-emptive resolution.
- The alert system, route visualization, and resolution options were helpful and effective for managing UAV operations.
- Color-coding (of alerts) and graphical display of information made alerts and hazards easier to quickly see.
- Clear, color-coded battery icons, shown along the route.
- Detailed visual representation of population hazards, display of proximity to threats, and ability to see battery status and the vehicle's remaining distance.
- Ability to get latitude and longitude data.

The UI was generally seen as having great potential but needing more development for full effectiveness. Participants suggested the following improvements (paraphrased and consolidated):

- **Dashboard**
    - Make the title of the service map displays more prominent and visually distinct to quickly identify services without relying on the dashboard. Ensure that the visual differences are significant for easy recognition at a glance.
    - Add a Sorting feature to arrange vehicles alphabetically.
    - Add a Weather / Wind service (expected wind speed) like Terminal Area Forecast (TAF) Wind Service to the dashboard.
    - Ability to Select Multiple Alerts for all services.
    - Move the user settings icon to the right side of the dashboard so that it does not cover the dashboard.
- **Map**
    - Improve map contrast and shading.
    - GPS display to highlight areas with poor coverage, especially near buildings in GPS display.
    - Legend on the map for population display is not able to differentiate probability of causality from the density of population.
    - Have a 3D view of the map and/or have the ability to see the side view on the map to check which vehicle is higher.
    - Add terrain information to the map.
    - Add the ability to view building height by hovering over them instead of clicking on the building.
    - Altitude of the vehicles along the route.
    - Automatically center the map on selected vehicles or alerts.
    - Add another column in the battery alert information box stating the reason why battery is going low.
    - Add a heat map that includes an alert box for GPS and RFI displays that explains why the signal is poor in those areas and a time slider to show signal variation.
    - Improve the visibility of vertiports on the map and provide details about their elevation (rooftop vs. ground), capacity, available battery types (e.g., brands) and charging points.
    - Ability to click and drag the routes (rubber banding) and see robust information directly correlated to the changes made.
    - Show altitude in feet vs. meters.
    - Provide resolution options for all services and rank them in some order.
    - Add an arrow at the end and middle of each line to indicate route direction.
    - Provide more information from the services, such as accessible battery reserve information (both time remaining, and distance, battery state of charge before takeoff and speed of UAV) battery usage projections, and clearer explanations for threshold alerts, as well as more details on proximity alerts.
    - Ability to select buildings, vertiports and waypoints and provide distance between them.
    - Integrate additional and more detailed information, including NFZs, Temporary Flight Restrictions (TFRs), time-sensitive information, UAV performance, real-time population movements, power lines, construction zones and airspace.
    - Integrate mission-related information in the UI, such as landing and takeoff information. The participants also noted inconsistencies between the paper maps they were given to supplement the UI in the study and the information shown on the UI, as well as missing mission related information.

- **Congestions Service**
    - Show only Cleared ETAs without the complexity of aggregate data.
    - The popup box for the congestion service should consistently appear in a designated, consistent location (e.g. bottom right of the map) that is away from other map information to avoid being hidden. The congestion window should have a fixed position to prevent it from covering important information like flight plans and hazards.
    - Selected routes or UAVs should be highlighted for better visibility.
    - Add a prioritization feature for the services, allowing users to set certain UAVs as higher priority. For example, if a lower-priority UAV encounters congestion, the UI would avoid options that negatively impact the higher-priority UAVs.
    - Reduce the number of clicks to reach resolution options.
    - Improve information transparency, by ensuring participants can easily access the necessary information without having to experiment to find what they need.
- **General**
    - Participants suggested enhancing the UI with more robust information for route changes to improve decision making. They want real-time updates showing the impact of changes on ETA, battery usage, and flight safety. Detailed metrics like distance to the next waypoint and total route distance should be displayed. Customization options to display relevant information and flexible adjustments to the UI were also recommended.
    - Participants appreciated the Congestion Service's potential but suggested it could benefit from further refinement, additional visualization options, and enhanced training. While they found the concept of aggregate data initially confusing, they see its potential value and feel that by reducing the mathematical complexity and optimizing the amount of information would improve the clarity of the congestion service, making it more accessible and user friendly.
    - Implement a mix of both the Old and New versions of the interface color-coding strategies for improved visibility and alert severity differentiation. For example, the multi-hazard information in the Alert Box could show both the colors of the hazard types while having some of the fields reflect the hazard alert levels to indicate their severity.
    - Provide additional training: some participants felt constrained by original flight plans but later realized they could modify routes. Additional training was also suggested for the congestion service, as noted above

D. Risk Mitigation Response Strategies

In addition to questions examining participants' ability to accurately retrieve information and their subjective assessments, eight questions regarding risk mitigation were also administered. The following examines risk mitigation data from the Package Delivery and Hurricane Preparation scenarios.

1. Risk Mitigation: Package Delivery Scenario

We asked participants to report which hazards they would solve first and why (see Table 4). The hazard type reported most frequently was battery (9), followed by proximity (5), population (1), and GPS (1).

Participants were asked how they would handle a battery reserve alert. The most frequently reported strategies were to charge the battery more prior to taking off and to get new equipment. Other cited strategies were to land the vehicle along its route and charge the battery and to make an impromptu route change or landing. Participants were also asked what options they would have and what they would do with the battery reserve alert that was present for two vehicles on the same route. Their answers were similar, with the most frequently reported strategy was to land on the route (6), recharge the vehicle's battery along the route (5), and to modify the vehicle's route and/or mission (5).

Participants were asked to report how they would handle the population hazard on the route of two of the vehicles. Most participants reported that they would make lateral modifications to the route to go around the population hazard (14), one participant reported that they would make both lateral modifications and increase the altitude of the vehicles, and one participant reported that their strategy for mitigating this risk would depend on the vehicle's capabilities. Similarly, we also asked participants to describe their strategy for avoiding the poor GPS area for two vehicles on the same route. The most frequently reported strategy was that they would make a lateral modification (6), followed by a directional modification coupled with an increase in altitude (5), altitude increase only (4), and one participant reported that their strategy for mitigating this risk would depend on the vehicle's capabilities.

Scenario 1: Package Delivery

Which alerts would you handle first?	Reasoning
Battery (9)	Battery required for UAV to operate (7) Can manually avoid hazards on the route if necessary (1) There are likely safety features on the UAV for GPS and RFI (1) Needs to know the UAV's capabilities when planning (1)
Proximity (5)	Ensure that the route is safe (1) Location of hazards (i.e., buildings and trees) will remain constant (1) Avoid damaging the UAV and property (1)
Population (1)	Safety of people is a top priority (1)
GPS (1)	Need to maintain connection to UAV (1)

Table 4 Frequency of responses for which alerts participant would solve first and the reason.

Participants were asked to use a paper map to draw new routes so that their risk mitigation strategies could be further examined. Participants were asked to adjust the flights so that they avoided all risks to people. All participants chose to laterally alter the route (16). Twelve participants rerouted the vehicle to the West of the population hazard, thereby extending the route. Of participants who provided a reason for their decision, five participants reported that they chose to fly to the West to avoid the NFZ located to the East and 1 reported that it was less risky overall. Four participants rerouted their vehicle to the East of the population hazard, with one participant reporting that they made this decision to preserve the vehicle's battery.

Two questions were asked regarding situations where a vehicle had insufficient battery life to complete its mission. First, participants were asked what risks there would be if the vehicle had a battery alert on the last leg of its route (it should be noted that the last leg of the route saw a population hazard to the West and a NFZ to the East). As a follow up, they were also asked what actions they could take to mitigate this risk. The most frequent risk that participants reported was the population hazard (13), followed by the NFZ (3), and unfavorable terrain (3) for making an emergency landing. The most frequently cited action that participants reported that they would take is to land at the nearest vertiport (4), followed by landing at the final delivery destination (4), and making an emergency landing (3).

2. Risk Mitigation: Hurricane Preparation Scenario

Participants were asked how they would solve the congestion alerts associated with different vehicles. It is worth noting that, for both questions, most participants selected one of the recommended congestion service's resolution options to remedy the congestion risk, suggesting high value of providing resolution options to the operators in these situations.

Participants were also asked, if they could not resolve every alert, what information they would communicate to their team of remote pilots. The most frequently reported responses were that they would provide their pilots with specific information about the unresolved alerts and risk mitigation strategies for when they encounter the alerts.

Finally, participants were asked how confident they would feel clearing the flights after using the SDSP-CD to assess risks and hazards. Participants were asked, as a fleet manager, how confident they would feel in clearing these flights for departure and, as a remote pilot, how confident they would feel flying the route. Participant responses were coded as either being confident, confident if certain criteria were met, and not confident. Overall, most participants expressed some level of confidence (e.g., they were either confident about clearing the flights or would feel confident if certain criteria were met). Those who expressed not being confident generally reported wanting additional information (e.g., weather and terrain data).

V. Discussion

This study aimed to determine whether this iteration of the SDSP-CD was indeed improved from the previous iteration; whether the UI is easy to use and effective, even on a more complex scenario; and how well the vertiport congestion service was integrated. The results section detailed the subjective responses, both qualitative feedback and quantitative ratings.

The direct comparison (section IV-B-1) of the various UI components that were changed clearly showed that the components of the updated UI were heavily favored by the participants. A notable exception was the color coding of the multiple hazards on the Alert Box. The coding was changed from coloring by hazards to coloring by hazard alert levels. The preference on this change was mixed. An alternative design that can code both the hazard alert level and the hazard type may need to be considered.

The results were also favorable for the UI's ease of use. Overall, none of the tasks related to information retrieval were overly difficult and the NASA TLX scores supported the notion that the UI components were easy to comprehend and use, requiring moderately low effort and mental demand. The NASA TLX scores for the information retrieval (section IV-B-3) suggest that it required low effort and high performance. The information retrieval questions (section IV-A-1) supported that finding, with participants answering with very high accuracy. On the simpler 4-vehicle Package Delivery scenario, participants answered all questions with 100% accuracy. On the more difficult 11-vehicle Hurricane Preparation scenario, the average accuracy dropped to approximately 91%, generally due to decreased performance on questions that required counting the alerts displayed for multiple vehicles. Their subjective feedback did not elucidate the increased challenges that would explain the decreased performance, but we hypothesize that the long list of vehicles that extend beyond the visible dashboard and need to scroll to see the entire fleet perhaps had a negative effect. Despite the challenges with that aspect of the dashboard, the participants' open-ended feedback suggested that the risk dashboard was an especially good idea for presenting the risks in a simple, comprehensible way. They also agreed that the general layout of the UI was well structured; nonetheless, they offered many suggestions for how to improve the visual interface and interaction with it. The SUS scores (section IV-B-2) reflected these pros and cons. For the Package Delivery scenario, the UI was considered "good." For the Hurricane Preparation scenario with the congestion service omitted, the UI scored somewhat lower, and even lower when the congestion service was included. The lower rating for the Hurricane Preparation scenario minus the congestion service is likely due to an averaging effect of the score for the original five services (i.e., the UI for Package Delivery scenario) with the score for the new congestion service.

To determine the effectiveness of the UI in providing necessary information for deciding on a risk mitigation strategy, we asked the participants several questions with no single correct answer. Answering these questions required integrating the information displayed on the UI, an accurate mental model of the services, and knowledge about the mission. The TLX scores suggested that planning risk mitigation actions required more effort and the participants felt they achieved lower performance. Overall, participants were able to merge necessary information and a reasonable mental model to decide on an effective mitigation strategy. In some cases, however, participant answers suggested challenges in understanding the meaning of the UI information, deficient mental models, reluctance in referring to paper-based information, and insufficient information about the flight conditions. Implications of these challenges include improved training for the services, UI integration of mission information where appropriate, and inclusion of additional sources such as weather, airspace constraints, and details about the vehicles and the flight plans, among others called out in Section IV.

A key functionality that most participants indicated would improve mitigation decision making was a *what-if* capability. The participants wanted to try alternate flight plans, changing altitude, lateral position, departure/arrival time, or redirecting a flight to an alternate vertiport (e.g., to mitigate congestion), and then rerunning the risk assessment to confirm a good risk posture, iterating until a route with acceptable risk is found. We believe the results of this study can inform the development of such a route planning functionality.

Finally, the third objective of the study was to evaluate the usability of the congestion service. As alluded to earlier, usability as evidenced by the SUS scores was not yet considered "good" for the Hurricane Preparation scenario with tasks exercising the congestion service added. The descriptive feedback from participants suggests that improvement is needed more for the UI interactions than for the fundamental functionality. The integration of the congestion service needs further improvements in the next iteration. The participants found the displayed congestion information difficult to understand. Part of that was due to how it was presented, but also, part of it was due to insufficient training received. The training was limited to a few minutes and explained in a couple of slides. That cursory explanation did not help the participants develop an intuitive feel for how it worked (that is, not transparent); thus, it was challenging for them to come up with mitigation options on their own. They found the automated resolution options helpful in that respect. Further, the service was initially designed for inflight use. The display was not altered for preflight planning, where a simplified view would have been more helpful. Lastly, one of the difficulties arose from a UI constraint in that the map and the congestion service could not be displayed simultaneously. That did not allow participants to see the routes when addressing congestion, making it challenging for them to evaluate a mitigation strategy.

VI. Conclusion

This study continued the user-centered design approach utilized for developing an effective and usable tool to support sUAS This study continued the user-centered design approach utilized for developing an effective and usable tool to support sUAS fleet managers in performing a preflight hazard and risk assessment. Participants were asked task-specific questions evaluating the usability of the UI for information retrieval and making inferences for mitigation decision making. They were also asked to fill out the TLX workload assessment and SUS questionnaire. Finally, they were asked open-ended questions to solicit feedback on what worked well and what needs improvement. The objectives of the study were to determine whether changes from the previous UI version improved the experience, evaluate the effectiveness of the UI for more complex scenarios, and evaluate how well a new risk assessment service for vertiport congestion was integrated into the UI.

The participants overwhelmingly endorsed the Supplemental Data Service Provider – Consolidated Dashboard (SDSP-CD) as a tool that can enhance risk management and preflight assessment of route safety. They indicated that it allowed better planning to reduce the chances of flying into adverse conditions and provided better insights, and the preparation time was well worth the improvement of the safety margin. The study also provided valuable feedback for improvements, especially for the integration of a new service – vertiport congestion – added in this iteration. Although participants appreciated the functionality of the service, refinements are required to better integrate it into the UI and improve its usability. These results were promising in that UI components with multiple iterations were positively rated while the new UI component with no feedback has room to improve, suggesting the importance of iterative UI design process with experts' feedback.

Participants indicated the SDSP-CD effectively supports preflight risk assessment and should also be extended to support comprehensive route planning. Various tasks given to the participants to elicit their risk mitigation strategies can inform the development of a route planning functionality in future work.

Acknowledgments

This work was supported by the System-Wide Safety project, under the Airspace Operations and Safety Program within the NASA Aeronautics Research Mission Directorate. The authors would like to express their appreciation to the pilots for their engaging participation and thoughtful feedback in this study and to give special thanks to Kyle Smalling from NASA Langley Research Center for his help in providing information for scenario development using sUAS to support hurricane emergency responses, and for sharing his resources with the Ames team.

References

[1] ABI Research, "Consumer Drone Shipments to Exceed 90 Million Units and Generate $4.6 Billion in Revenue by 2025, ABI Research, 19 Jan 2016. www.abiresearch.com

[2] Federal Aviation Administration, "Unmanned Aircraft System (UAS) Traffic Management (UTM) Concept of Operations v2.0," 2020. https://www.faa.gov/sites/faa.gov/files/2022-08/UTM_ConOps_v2.pdf

[3] Moore, A. J., Young, S., Altamirano, G., Ancel, E., Darafsheh, K., Dill, E., Foster, J., Quach, C., Mackenzie, A., Matt, J., Nguyen, T., Revolinsky, P., Smalling, K., Vazquez, S., Martin, L., Feldman, J., Gujral, V., Spirkovska, L., Hooey, B., Bradner, K., Iverson, D., Wolfe, S., Oza, N., Condotta, R., Morris, C., Glover, J. S., Rymer, N., Turner, A., Walter, C., Banerjee, P., Corbetta, M., Kulkarni, C., Cleland-Huang, J., Jones, J., and Bonin, T., "Testing of Advanced Capabilities to Enable In-Time Safety Management and Assurance for Future Flight Operations," NASA-TM-20230018665, 2024.

[4] Young, S., Ancel, E., Moore, A., Dill, E., Quach, C., Foster, J., Darafsheh, K., Smalling, K., Vazquez, V., Evans, E., Okolo, W., Corbetta, M., Ossenfort, J., Watkins, J., Kulkarni, C., and Spirkovska, L., "Architecture and information requirement to assess and predict flight safety risks during highly autonomous urban flight operations," NASA-TM-220440, 2020.

[5] Feldman, J. M., Martin, L., Gujral, V., Walter, C. M., Billman, D., Revolinsky, P., and Costedoat, G., "Developing and Testing Two Interfaces for Supplemental Data Service Provider (SDSP) Tools to Support UAS Traffic Management (UTM)," Proceedings of the AIAA AVIATION Forum, 2023-3966, June 2023. https://doi.org/10.2514/6.2023-3966.

[6] Feldman, J. M., Martin, L., Costedoat, G., and Gujral, V., "Usability of pre-flight planning interfaces for supplemental data service provider tools to support uncrewed aircraft traffic management," Application of Emerging Technologies, Vol. 115, 357-366, Proceedings of AHFE International, 2023.

[7] Spirkovska, L., “Vertiport Dynamic Density,” NASA TM-20230012622: Moffett Field, CA, NASA Ames Research Center, 2023.

[8] Code of Federal Regulations, “Subpart U—Dispatching and Flight Release Rules”, 14 CFR Part 121 Subpar-U [online document], URL: https://www.ecfr.gov/current/title-14/chapter-I/subchapter-G/part-121/subpart-U [retrieved May 2024].

[9] Martin, L., Arbab, Y., and Mercer, J., “Initial Exploration of STEReO (Scalable Traffic management for Emergency Response Operations) System User Requirements for Safe Integration of Small UAS,” *IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)*, San Antonio, TX, USA, 2021, pp. 1-8. https://doi.org/10.1109/DASC52595.2021.9594419.

[10] Martin, L., Arbab, Y., Roberts, L., Mercer, J., Walter, C., McCarty, W., and Sheehe III, C., “Developing an unmanned aircraft system pilot kit (UASP-kit) for wildland fire UAS operators,” Proceedings of the AIAA AVIATION Forum, 2022-4003, June 2022. https://doi.org/10.2514/6.2022-4003.vid

[11] Mercer, J., “Advanced Capabilities for Emergency Response Operations (ACERO),” CWSF Western State Fire Managers Committee Annual Meeting, 2023.