CONVECTIVE WEATHER DETECTION BY GENERAL AVIATION PILOTS WITH CONVENTIONAL AND DATA-LINKED GRAPHICAL WEATHER INFORMATION SOURCES

James P. Chamberlain & Kara A. Latorella, Ph.D.

NASA Langley Research Center, Hampton, Virginia

Abstract

This study compares how well general aviation (GA) pilots detect convective weather in flight with different weather information sources. A flight test was conducted in which GA pilot test subjects were given different in-flight weather information cues and flown toward convective weather of moderate or greater intensity. The test subjects were not actually flying the aircraft, but were given pilot tasks representative of the workload and position awareness requirements of the en route portion of a cross country GA flight. On each flight, one test subject received weather cues typical of a flight in visual meteorological conditions (VMC), another received cues typical of flight in instrument meteorological conditions (IMC), and a third received cues typical of flight in IMC but augmented with a graphical weather information system (GWIS). The GWIS provided the subject with near real time data-linked weather products, including a weather radar mosaic superimposed on a moving map with a symbol depicting the aircraft's present position and direction of track. At several points during each flight, the test subjects completed short questionnaires which included items addressing their weather situation awareness and flight decisions. In particular, test subjects were asked to identify the location of the nearest convective cells. After the point of nearest approach to convective weather, the test subjects were asked to draw the location of convective weather on an aeronautical chart, along with the aircraft's present position.

This paper reports preliminary results on how accurately test subjects provided with these different weather sources could identify the nearest cell of moderate or greater intensity along their route of flight. Additional flight tests are currently being conducted to complete the data set.

Introduction

Eighty-five percent of the aviation accidents that occurred from 1990-1996, and nearly eighty-five percent of the accident fatalities, involved small general aviation (GA) airplanes. Weather is a factor or cause in nearly a third of these accidents, which equates to approximately eleven weather-related GA accidents per week, with four of the eleven involving fatalities. The Aviation Weather Information (AWIN) program element, which is part of the NASA Aviation Safety program, aims to improve these accident statistics by improving weather information available to aviation users. A particular focus of the AWIN element is to develop technologies and design/use guidelines that provide improved cockpit weather information via graphical displays of data-linked weather products. Goals of this technology and guideline development are to improve pilots' in-flight weather situation awareness and decision quality, ultimately leading to safer flights.

General aviation is particularly affected by convective weather. A survey of GA accidents from 1982 to 1993 revealed that while only 3.5% of these accidents are directly attributed to thunderstorms, a large percentage of these accidents, 66%, resulted in fatalities [1]. Convective weather is challenging because it can be characterized by rapidly changing weather conditions, heavy rain, severe to extreme turbulence, high winds and gusts, hail, icing, lightning, severe downdrafts and microbursts, reduced ceiling and visibility, and instrument meteorological conditions (IMC). Such concomitant weather phenomena were analyzed separately in the aforementioned accident analysis and contribute to many additional accidents. Therefore the incidence of GA accidents attributed to convective activity, and the fatalities resulting

from such weather systems, is likely under-represented by the percentages cited for only thunderstorm effects.

Currently, pilots of small GA aircraft have limited in-flight information on convective weather activity, especially when compared to that available on larger aircraft. Unlike larger aircraft, most small GA aircraft are not equipped with onboard weather detection equipment such as weather radar or lightning detection systems (*e.g.*, Stormscope, Strikefinder) that can indicate convective activity. In addition, the onboard weather radar systems that are available for small GA aircraft are typically expensive, and limited in performance by size and power constraints. When available, these systems can provide improved weather awareness for severe weather hazards, but are limited in range and accuracy [2]. Onboard weather radar systems are workload-intensive to use accurately [3], are subject to attenuation, have a limited range, and provide information that is primarily forward of the aircraft and at the aircraft's altitude [4]. While these systems show severe local weather to avoid, they do not provide the more comprehensive weather picture required to fully support strategic planning or avoidance maneuvers.

Pilots of small GA aircraft today rely principally on aural sources and external, or "out-the-window," weather cues for weather information. Aural sources can include direct queries to Flight Service Station (FSS), En Route Flight Advisory Service (EFAS, or "Flight Watch"), and Air Traffic Control (ATC) personnel, as well as monitoring frequencies to hear other pilots' comments, queries, and the information supplied to them. Pilots can also tune in automated weather information services such as HIWAS, AWOS/ASOS, and ATIS to obtain a broadcast of conditions over a large area or at specific reporting stations. Unfortunately, the information available from these aural sources is limited and, when weather becomes a problem, the frequencies used to obtain this information become saturated, making this information inaccessible at exactly the time it is most needed.

More accessible, complete, and usable weather information would benefit pilots' situation awareness, decision-making, and safety. Graphical weather is a more appropriate representation [5], can more effectively be integrated with other such

information (*e.g.*, terrain) and can be extended using symbology. Accordingly, pilots using prototype graphical weather information systems (GWIS) in static and dynamic desktop simulation experiments were more likely to acquire trend data, have a more comprehensive awareness [6], make better go/ no-go decisions, rate hazard levels higher, have more confidence in weather-related decisions, make fewer calls to ground aviation weather personnel [7][8], and make more correct decisions with graphical, than with verbal or text alerts [9][10].

One early implementation of an uplinked radar mosaic GWIS, developed at MIT Lincoln Labs with funding from the FAA Datalink Operational Requirements Team (DLORT), had a 15-minute update rate, 6km-square resolution and employed a "lossy" algorithm (resulting in less well-defined precipitation areas) to compensate for lower available bandwidth (250bps) [11][12]. In desktop usability assessments, all subjects found the high level of lossy compression unacceptable, and some found that the medium level lacked the functional equivalence of the uncompressed image [7][13]. When used in a GA flight test, accompanied by terminal forecasts and surface observations, and integrated with a traffic information service, subjects commented enthusiastically on the utility of this GWIS [14]. More than 82% of subjects had positive responses to the utility of precipitation maps, surface observations, and terminal forecasts individually [15]. All subjects had a positive overall impression of the system; 88% indicating that it would be important to make available to GA operations [15].

The FAA Flight Information Services Data Link (FISDL) program will soon make data-linked weather information systems widely available to GA pilots via commercial FISDL vendors. The FISDL vendors will provide, for no service charge, uplink of textual aviation weather products. These products include weather observations (METARS & SPECIs) and forecasts (TAFS) of terminal environments, as well as reports of severe weather conditions (SIGMETS, Convective SIGMETS, AIRMETS, and severe weather forecast alerts) and pilot reports (PIREPS). GA pilots may augment this basic information by purchasing services that will uplink graphical weather information,

including a national weather radar mosaic (NEXRAD mosaic). This textual and graphical weather information will be broadcast via a network of VHF ground stations, and received and displayed via an onboard GWIS.

NavRadio Corporation (now part of the Bendix-King Division of Honeywell International), in a cooperative agreement with NASA AWIN, developed such a prototype GWIS which was subsequently selected for the FISDL program. FAA FISDL and NASA AWIN jointly funded a simulation experiment at Research Triangle Institute (RTI) to evaluate pilot weather flying with and without a version of this GWIS. In this study, the flight simulator subjects were in IMC, had access to an autopilot, and were given a GWIS display that included a NEXRAD mosaic map but lacked an overlaid aircraft present position symbol. Results indicated that while this GWIS increased awareness of the general location of convective weather, it did not improve pilot diversion decision-making (subjects did not understand the location of weather with respect to their position), increased workload for at least half the subjects, and reduced reliance on ground-based weather professionals [16]. This simulation study suggested several features for GWIS's (*e.g.,* aircraft present position symbol) and concluded that further experimentation is required to develop industry standards for appropriately designing GWIS interfaces and procedures for using these systems.

The AWIN Convective Weather Sources (CoWS) experiment, described in this paper, also uses a variant of the NavRadio-developed prototype GWIS but does so in a flight environment. This particular GWIS variant, hereafter referred to in this paper as the "AWIN GWIS," includes a symbol depicting the aircraft's present position and direction of track overlaid on a NEXRAD mosaic map, which is displayed on a handheld, tethered unit.

The CoWS experiment principally investigates how GA pilots' use of various weather information sources – conventional aural, "out-the-window" visual, and GWIS-displayed cues – affects their in-flight weather situation awareness and decision-making related to convective weather systems. In addition, this experiment allowed us to collect usability data for this GWIS implementation.

An earlier publication [17] reported initial CoWS experiment flight test results of pilots' relative confidence, information sufficiency, and workload ratings when using aural, out-window visual, and graphically represented weather information cues in flight near convective weather. This paper reports on the accuracy and consistency of the test subjects' ability to identify convective weather relative to their aircraft location and flight track. Additional flight tests are currently being conducted to complete the data set.

Methods

Apparatus

Apparatus for the CoWS experiment includes the test aircraft, AWIN GWIS, other airborne equipment, and supporting ground infrastructure. These items are described below.

NASA Langley Research Center's Raytheon B-200 Super King Air was selected as the test aircraft. The B-200 is a nine-passenger, pressurized twin-turboprop airplane with a cruise speed of 265 knots and a service ceiling of 35,000 feet. The B-200 was selected for several reasons:

- It is large enough to carry three non-flying test subject pilots and two experimenters to the same weather scenario;
- It can get to the test area of interest quickly, and then slow to speeds comparable to smaller piston-engine GA aircraft such as those typically flown by the test subjects;
- It is adequately powered, pressurized, and radar- and deice-equipped for safety-of-flight concerns when approaching convective weather; and
- It is equipped and approved for operations with the necessary experimental equipment on board.

The AWIN GWIS components are installed as an equipment pallet that is strapped into one of the passenger seats. The pallet components include a power supply, AWIN VHF Data Link (VDL) receiver, Global Positioning System (GPS) receiver with recording capability, and two laptop PC's and scan converters with tether cables to two small

handheld AWIN display units. Electrical power and VDL and GPS antenna connections supply the pallet via cables from an overhead panel. The GPS and VDL receivers supply position and weather data, respectively, to both laptop PC's. The PC's provide the image to scan converters and ultimately the portable display units through tether cables. Software applications on the PC's implement the AWIN display user interface. This interface shows textual and graphical weather and map data in response to user commands via each display unit's bezel buttons and joystick.
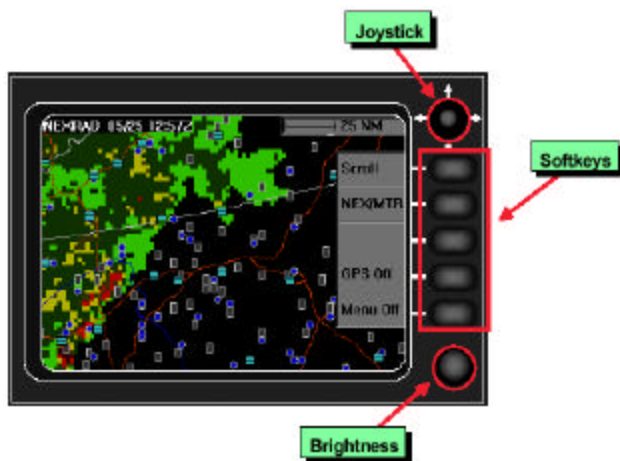

Figure 1. AWIN GWIS Display Unit

Each AWIN display unit's screen is approximately 4 inches tall by 5 inches wide. Five bezel buttons on the right side of the unit actuate soft menu fields, and a rate-controlled joystick controls pan, zoom, and crosshairs for symbol selection (see Figure 1). The unit presents lossless, nationwide radar mosaic imagery at 4-square-km resolution with a 6 minute nominal update rate assuming adequate broadcast reception, and surface weather observations (METAR) in text and symbolic form for reporting stations in the mid-Atlantic region. The display also presents contextual features (rivers, interstates, and state boundaries), airport identifiers, present position and track symbol, creation time stamp for the radar product, a scale legend, and indicates missing data. The features and usability issues of this display will be more fully addressed in a separate report.

Other airborne equipment includes a video camera and intercom/recording system operated by the forward experimenter. A flexible hood is affixed to the aircraft's onboard weather radar display to prevent it from being viewed by the test subjects, while still allowing it to be viewed by the pilot in command (PIC) and videotaped.

The intercom/recording system includes headsets for all aircraft occupants, and has multiple channels to allow the PIC and experimenters to coordinate flight and experimental protocol details offline from communication with the test subjects. All audio communications between the PIC, experimenters, and test subjects are recorded, as are ATC instructions and weather information. The communication system allows subjects to hear real-time conventional aural weather information acquired during the experiment.

The primary supporting ground infrastructure includes four prototype AWIN ground stations equipped with satellite weather receivers, PC-based processors, and VDL Mode 2 broadcast transmitters. The AWIN ground stations are located at four sites in Virginia, and provide a broadcast link of packaged weather data files to the B-200 along several flight routes (see Figure 2). The ground stations all use a single time-shared frequency of 136.275 Mhz to alternately broadcast their respective data files.
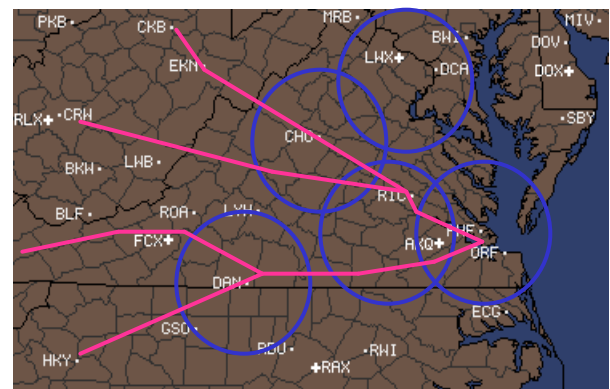
Figure 2. AWIN Flight Routes

Scenarios

The CoWS experiment scenarios are basically a series of flight tests in the B-200 aircraft, in which GA pilot test subjects are provided with different in-flight weather information sources and flown toward convective weather of moderate or greater

intensity. Subjects do not perform flying duties during these flights; a NASA test pilot serves as PIC. The flights are conducted under Instrument Flight Rules (IFR) but in visual meteorological conditions (VMC).

A typical scenario includes a departure from NASA Langley/ Langley Air Force Base (LFI) on a flight path that will obliquely intercept a frontal convective system of at least moderate intensity at approximately 120 nautical miles (nm) from top-of-climb, and at an altitude above the haze layer (typically 14,000 feet). Scenario definition is constrained by the location of the weather and the GWIS's ground-based infrastructure. To accommodate this constraint and minimize training and materials, four potential IFR flight plans were developed from LFI to Hickory, NC (HKY); Charleston, WV (CRW); Abingdon, VA (VJI); or Clarksburg, WV (CKB). One of these four flight plans is chosen on the morning of each flight based on prevailing weather conditions.

Weather Cue Conditions

Three sets of weather sources, or cue conditions, are provided to subjects. The first cue condition represents the weather information typically available to a GA pilot in IMC. This "Traditional IMC" condition (aural) consists only of aural weather information. The "Traditional VMC" condition (window+aural) augments the aural cues with the visual cues provided by an "out-the-window" view. Finally the "GWIS-IMC" condition (display+aural) augments the aural cues with access to the AWIN GWIS formerly described. Opaque window coverings restrict the views of subjects in aural and display+aural conditions.

Protocol

A subject team arrives at NASA Langley in the morning, and each subject is provided with an introductory briefing, consent form, schedule, and Preliminary Questionnaire. These subjects then receive a mission motivation and briefing; a local terrain, navaid and airport identifier review; a route briefing for the flight to be taken; and practice on forms and procedures to be used during the in-flight phase. Following a short break, subjects have 10 minutes to review a textual DUATS preflight

(standard weather) briefing and associated weather graphics; 10 minutes to listen twice to an audiotaped recording of a FSS preflight briefing; and an additional 10 minutes to review this preflight material. The preflight material is obtained on the morning of the flight, for a departure time two hours prior to the actual departure of the flight. Subjects then complete the Preflight Weather Situation Awareness questionnaire.

After lunch, the subject chosen to receive the AWIN GWIS receives a standardized training presentation, test, and compensatory instruction on this system. This subject is also allowed to practice with the GWIS display upon reaching the aircraft. In parallel to the formal display instruction, another subject receives a weather knowledge survey, and the remaining team subject receives a risk tolerance test and a general personality inventory test. The team then boards the aircraft for the in-flight portion of the experiment.

The in-flight portion of the experiment starts after the aircraft has climbed to cruising altitude and when the aircraft is approximately 120 nm from the first convective weather area of moderate or greater intensity. The outbound leg of the in-flight portion is ended when approximately 20 nm from this area, or at approximately 100 nm from the initial experiment starting point, whichever occurs first.

Throughout the outbound leg of the in-flight portion, subjects are given either a Position Update task, a Weather Situation Awareness questionnaire, or provided aural weather information, on a defined schedule, with one of these events occurring nominally every 4 minutes. The Weather Situation Awareness questionnaires are given every 8 minutes (approximately every 25 nm). The Position Update tasks and aural weather information are alternately provided between the questionnaires, so that each is provided approximately every 16 minutes. Each of these events is described below.

Subjects are given a Position Update task, nominally every 16 minutes (50 nm) during the outbound leg, to compensate for the loss of positional awareness and workload induced by not piloting. For this task, the subjects are required to copy scheduled PIC reports (containing airspeed, altitude, heading, position, next waypoint, and current time) onto a prepared form, plot position on

an IFR low altitude en route chart, and calculate elapsed time and ground speed.

All subjects receive scheduled aural weather information at 3 intervals during the outbound leg, 16 minutes (50 nm) apart. The first cue is obtained from a local automated Hazardous InFlight Weather Advisory Service (HIWAS) broadcast outlet, the second from querying EFAS personnel, and the third from querying ATC.

Subjects are provided with a Weather Situation Awareness questionnaire nominally six times during the outbound leg, at roughly 8 minute (25 nm) intervals. These short questionnaires include items addressing the subjects' weather situation awareness and flight decisions. In particular, subjects are asked to identify the location of the nearest convective cells.

During the outbound leg of the in-flight portion, one of the AWIN GWIS display units is provided to the test subject receiving the display+aural weather cue. The experimenter seated opposite the equipment pallet monitors and controls the AWIN software applications as necessary directly from each PC's keyboard and mouse, and maintains the experiment schedule using the GPS clock display.   The forward experimenter coordinates experiment and flight path details with the PIC, and operates the intercom/recording system and video camera. The video camera is used each time a Weather Situation Awareness questionnaire is administered, to record the NEXRAD radar mosaic product on the extra AWIN GWIS display, the aircraft's onboard weather radar display, the primary flight and navigation instrument indications, and the forward visual scene from the flight deck windows. The visual scene is panned through approximately 190 degrees of azimuth, with vertical panning as necessary to record significant cloud formations and build-ups.

At the conclusion of the outbound leg, subjects are asked to plot the aircraft's position on their en route IFR chart, draw weather within 50 nm of the flight path on the chart, and complete the Inbound Questionnaire. This instrument contains NASA-TLX [18] -derived scales for workload assessment, asks subjects to indicate other weather sources that would have been helpful, and about their flight decisions. After completing the Inbound Questionnaire, the subject using the display is asked to complete a Usability Questionnaire and provide any additional comments. Following the flight, subjects are provided with a short debriefing questionnaire for that flight. At the conclusion of the third flight for a team, when all subjects have been exposed to the display, subjects and experimenters more fully discuss issues of experimental validity and display usability.

Test Subjects

Test subjects were recruited from local regional airports through advertisements. Applicants reported their general and weather flying experience on a Background Questionnaire. Subject selection criteria included: an instrument rating, 10-50 flight hours within the last 90 days, and 50-1000 cross-country or 100-2000 total flight hours. In addition, participants were not selected who had worked for a scheduled air-carrier in the prior year or who had participated in the aforementioned RTI/AWIN experiment.

Experiment Design

The full CoWS experiment design requires twelve test subjects (divided into four three-member teams) and twelve test flights. The results reported in this paper are based on the first four test flights, which includes all three of the first subject team's flights and one of the second team's flights.

Each subject team flies on three separate test flights, with individual subjects receiving a different weather cue condition on each flight. This allows us to compare the weather cue conditions in a common weather experience. Subject assignments to weather cue conditions are rotated for each of the team's three flights, so that all three subjects receive each weather cue condition, removing individual difference effects associated with using the different weather sources.

Because weather experience has been found to significantly affect weather-related decision-making and information acquisition [19], candidate subjects were clustered into 3 groups of "exposure experience" using cross-country hours. The midpoints of each cluster are 135 (low), 379 (medium), and 738 (high) cross-country hours respectively for these preliminary subjects. Subjects were selected to form four three-member

teams, each team composed of one subject from each of the clusters to balance exposure experience across flight scenarios. Weather cue condition assignments to subject experience levels were counterbalanced to mitigate against cue condition *vs*. experience level interactions.

The independent variables for each flight are the weather cue conditions assigned to each test subject: aural (representing traditional IMC), window+aural (representing traditional VMC), and display+aural (representing AWIN GWIS-augmented IMC). The dependent measures are derived from the subjects' Weather Situation Awareness questionnaire responses and chart drawings of weather, and are described below in the Dependent Measures section.

Dependent Measures

This section describes how dependent measures were developed for the experiment.

One item on the Weather Situation Awareness questionnaire asks the test subject to identify the nearest convective cell of moderate or greater intensity, and to estimate the bearing (or direction) to the cell and its range (or distance) from the aircraft. The subject is given the option of answering "no cell" if he/she believes that no moderate or greater intensity cell is present within 200 nm of the aircraft.

This questionnaire item is used to generate three dependent measures: Cell Detection, Bearing Accuracy, and Range Accuracy. These measures are derived by comparing the subjects' responses to a reference standard for the actual location of the nearest cell (of moderate or greater intensity) to the aircraft at the time of the response. The Cell Detection measure records hits (cell was present and detected), misses (cell present but not detected), correct rejects (cell not present, and not detected), and false alarms (cell not present but was erroneously detected) for each subject. The Bearing and Range Accuracy measures note the difference between the subjects' bearing and range estimates, in degrees relative to aircraft heading and nautical miles (nm), respectively, and those derived from the reference standard.

The subjects' drawings of weather on their charts at the end of the outbound leg are used to

generate a fourth dependent measure called "big picture weather situation awareness," or BPWSA. This measure is a structured evaluation and ranking of the subjects' overall or "big picture" weather awareness, and is done by comparing different portions of the subjects' chart drawings with the reference weather standard. Specifically, subjects' chart drawings are evaluated in 5 regions relative to the aircraft's present position:

- < 50 nm ahead, < 50 nm to either side
- > 50 nm ahead, < 50 nm to either side
- < 50 nm behind, < 50 nm to either side
- > 50 nm behind, < 50 nm to either side
- > 50 nm to either side

Each chart portion was ranked on a ten-point scale, for: accuracy of hazardous area, cell, and line location, orientation and shape (0-6 points); accuracy of cell motion arrows (0-2 points); and accuracy of cell intensity levels (0-2 points). These rankings were then weighted in descending order by chart region as listed above (*i.e.,* "< 50 nm ahead" region was weighted five times more heavily than "> 50 nm to either side" region). All rankings were then summed and normalized to a number between 0 and 1 for comparison purposes.

The various sources of reference standard or "ground truth" data include individual NEXRAD site data, onboard weather radar video images, NASA pilot observations, and various sources of NEXRAD mosaic products. The reference standard used in this paper is based on consecutive hourly WSI NOWRAD (*i.e.*, NEXRAD mosaic) products as archived by the National Climatic Data Center (NCDC). The method for deriving the reference standard from these consecutive-in-time mosaic products follows:

The nearest mosaic products before and after the time the subjects responded to the Weather Situation Awareness items are selected. The known aircraft position, based on GPS output, is then plotted on each Before and After mosaic product. The nearest cell of moderate or greater intensity (*i.e.*, at least yellow – 40 DBZ) to the aircraft position is then determined on the Before and After mosaics. Bearing and range from the aircraft position to this cell location is then determined. The weather patterns on these mosaic products are

then examined. Based on the patterns the final reference bearing and range is determined, either by interpolation between the Before and After mosaics, or by selection of the nearest-in-time mosaic result. In almost all cases herein, the convective activity was well-organized and established into easily-identifiable lines or cells on both Before and After mosaics, and interpolation was chosen as the best determinant of reference bearing and range. There were two cases in which the standard was based on the Before mosaic alone.

Results

Cell Detection

Over all conditions, subjects identified a cell of moderate or greater intensity when one existed approximately 78% of the time, but reported no such cell in the area when one did exist approximately 22% of the time. Examining these results further indicates that while each experimental condition did result in subjects reporting cells when they did exist, these accurate reports occurred most frequently in the display+aural condition. Subjects in both the aural and display+aural conditions were more apt to correctly state that there was no cell of moderate or greater intensity within the specified region.   The display+aural condition resulted in the fewest erroneous reports. When supported by the AWIN GWIS Display, subjects were much less likely to miss significant cells or to falsely report cells when none existed. All errors committed by subjects experiencing the aural condition were failures to detect existing cells and such misses accounted for the majority of window+aural errors as well. When subjects were in the window+aural condition, they falsely reported cells that did not exist. This did not occur in the aural condition, and minimally in the display+aural condition. Table 1 shows the relative counts for hits, false alarms (FA), misses, and correct rejections (CR) for each condition.

We further assessed correct identification situations, hits, to determine the accuracy of bearing and distance estimates to the nearest cell.

Table 1. Detection Category Results.

	Miss	FA	Hit	CR
Aural	7	0	16	4
Window	7	4	16	0
Display	1	1	22	3
Total	15	5	54	7

Bearing Accuracy

Table 2 shows the accuracy of relative bearing assessments in which a cell existed and the subject identified a cell. Table 2 shows the same relative patterns for "very accurate" (within 25 degrees of actual bearing), "accurate" (within 45 degrees), and "relatively accurate" (within 90 degrees) bearing estimates over the experimental conditions. Generally speaking, the window+aural condition supports more accurate bearing estimation than the display+aural condition, which supports more accurate bearing estimation than the aural condition. Subjects provided with the aural condition were much less accurate in their bearing assessments than when provided with the window+aural, or the display+aural condition.

Table 2. Percent Accuracy of Bearing Estimates for Hits.

	$\leq 25^0$	$\leq 45^0$	$\leq 90^0$	$> 90^0$
Aural	31	50	63	38
Window	56	75	94	6
Display	45	68	82	18

Range Accuracy

All three experimental conditions seem to support distance estimates for most reports within approximately 100 nm of the actual distance to the target cell. Subjects were least accurate when only afforded aural weather information. All distance estimates, when in the window+aural condition, were within 50nm of the true cell distance. Over three-quarters of the estimates were within 50nm accuracy when subjects used the display+aural condition, which is more accurate than estimates provided based only on the aural condition.

Approximately half of both the window+aural condition and display+aural condition estimates were accurate to within 25nm and both were more accurate than the distance estimates for the aural condition. The display+aural condition supported marginally more "very accurate" estimates of bearing than did the window+aural condition. When provided with the aural condition and with the display+aural condition, some subjects reported distance estimates that were in excess of 50 nm off from the target distance. In one case, for each of these conditions, these estimates were in excess of 100 nm. For all of the cases in which subjects using aural weather information erred by more than 50 nm, they reported that the closest cell of moderate or greater intensity was further than the target cell we identified. This was true for all but one of the cases in which subjects using the AWIN display erred by more than 50 nm. These data are summarized in Table 3.

Table 3. Percent Accuracy of Distance Estimates for Hits.

	$\leq$25nm	$\leq$50nm	$\leq$100nm	>50nm
Aural	31	56	94	44
Window	50	100	100	0
Display	59	77	95	23

Big Picture Weather SA

The weather that subjects drew on their en route charts at the end of the outbound leg allowed them to indicate their accumulated understanding of how the weather situation had developed over the flight. This understanding was evaluated and ranked with the Big Picture Weather SA (BPWSA) measure.

Results from the Big Picture SA measure are shown in Table 4 for each weather cue condition. As described earlier, the BPWSA measure is a weighted ranking of the subjects' overall weather awareness at the end of the outbound leg. Heavier weighting is given to knowledge of hazardous weather location, orientation, and shape near and ahead of the aircraft. Lesser weight is given to knowledge of areas further from and/or abeam or

behind the aircraft, and for knowledge of cell directions and intensities. The rankings are normalized to a number between 0 and 1, with 1 representing a fully correct and complete knowledge of all weather hazards and their directions and intensities in all areas.

"Correct rejects" (*i.e.,* no weather was present, and no weather was drawn by the subject) were treated using two approaches when calculating the BPWSA measure. In the first approach, correct rejects (CR) are given a perfect score of 10 and included in the weighted rankings; aggregated results from this approach are shown in the middle column of Table 4. In the second approach, correct rejects are eliminated altogether from the weighted rankings; these results are shown in the right-hand column of the table.

Table 4. Big Picture Weather SA Results.

	SA Score1 CR=10	SA Score2 No CRs
Aural	0.615	0.329
Window	0.673	0.327
Display	0.910	0.743
Average	0.733	0.466

The results of the BPWSA measure in Table 4 indicate that subjects with the display+aural condition have significantly higher BPWSA rankings than either aural or window+aural. Further, there is little difference in the scores for subjects with aural and window+aural conditions. This result is true for both treatments of "correct rejects," *i.e.*, with and without including "correct rejects" in the ranking.

Test subjects with the display+aural condition typically lost points in the BPWSA rankings primarily due to not indicating the direction of motion of the cells, or due to not drawing a weather hazard that existed far abeam or behind the aircraft. While subjects with aural or window+aural also lost points for these omissions, they also frequently lost points for significant errors in weather hazard identification near and/or ahead of the aircraft. In all cases, display+aural subjects correctly identified hazardous weather features near and ahead of the

aircraft, and in almost all cases correctly identified the configuration or shape. Subjects with the window+aural condition would tend to misidentify the location and/or configuration of hazardous areas, *e.g.*, a significant error in the orientation of an extended line of cells. Subjects with only aural information would sometimes miss a hazardous weather area altogether; more frequently, however, they would indicate hazardous weather existing somewhere in a large area, but have no detailed knowledge of its shape or location.

Discussion

Results from all dependent measures – Cell Identification, Bearing Accuracy, Range Accuracy, and BPWSA – show that when subjects were in the window+aural and display+aural conditions they have an improved understanding of the hazardous weather situation over when they had the aural condition. When subjects used the display+aural condition, they had markedly improved detection of cells over their performance when using either the window+aural or aural conditions. Where cells were correctly identified, the distance and bearing estimates were least accurate for the aural condition. The window+aural condition supported more accurate bearing estimates and was most accurate for distance estimates if allowing a relatively lenient buffer of +/-50 nm. The display+aural condition improved the likelihood of more accurate (25 nm) estimates marginally over the window+aural condition. Subjects were most likely to falsely report a cell where none existed when using the window+aural condition. This last point bears further discussion.

In general, the convective weather present during the flight tests reported in this paper could be characterized as well-defined, significant lines and areas of cells, with surrounding towering cumulus buildups and occasional embedded cells. Flight conditions could be generally characterized as unrestricted visibility, on top of any lower cloud layers, and laterally clear of towering cumulus and cumulonimbus cells (we attempted to achieve these weather conditions on the outbound leg of all flights). However, while we strove to minimize such occurrences, there were occasional instances of haze and flight between layers. More frequently, there were instances of towering cumulus in the

vicinity of moderate or greater intensity cells; these towering cumulus clouds were sometimes positioned between the aircraft and the cells. At other times, the cells depicted on the NEXRAD mosaic and/or the onboard weather radar were embedded in larger cumulus cloud masses. For this discussion we will collectively refer to such towering cumulus clouds and masses as "masking clouds."

It is possible that these masking clouds account for many of the false alarm errors experienced by the window+aural subjects. This conjecture is corroborated by our project pilot, who is also an active corporate pilot and has extensive convective weather flying experience in the mid-Atlantic and Southeast regions. Based on this project pilot's experience, first-hand observation of the experiment flights, and correlation of those flight conditions in real time with the onboard weather radar, visual evaluation of cell strength, bearing, range, and configuration can be misleading in the presence of masking clouds. The fact that such visual evaluations can be misleading does not mean that visual weather cues are not valuable – numerous texts on flying technique recommend visual avoidance of cell buildups [20][21], as does our project pilot. In terms of safety, the false alarm errors made by the subjects when using the window+aural condition could be viewed as conservative errors. That is, avoiding a false alarm area is better than not avoiding a missed cell area.

One other discussion point concerns a potentially hazardous use of the GWIS for tactical weather avoidance. As mentioned in our earlier report [17], for several reasons the uplinked NEXRAD mosaic product can sometimes become outdated, and with the present display design pilots often do not notice that the product is old. Even when the uplinked product is up-to-date, the product itself is typically built from 6-to-10 minute old NEXRAD site data, and represents a near-real time, but *not* real-time, weather situation.

Sole use of this time-delayed information by pilots for tactical avoidance of nearby cells is therefore potentially hazardous, but a particularly hazardous action would be to attempt penetration of a line of cells by flying through a gap, or clear space between cells, depicted on the AWIN display. Individual cells in a line frequently move relatively

quickly along a path parallel to the line, and consequently the gaps move along the line quickly as well. On several occasions when returning from our outbound experiment legs, we flew near lines with gaps using the onboard weather radar and visual cues to maintain safe tactical clearance from the cells. On some of these occasions we flew out of range of our AWIN ground stations, and the NEXRAD mosaic became outdated. In these cases, it was not unusual to find that a "gap" in the line, as depicted by the AWIN GWIS display, was now filled with a cell that had moved down the line since the NEXRAD mosaic product was last updated.

Perhaps the best in-flight convective weather situation awareness would occur when pilots use all three weather sources, *i.e.*, both aural and AWIN display sources combined with a VMC view out the window. The weather sources complement each other, in that the AWIN display provides a better strategic or "big picture" view of the weather situation, the window view keeps pilots tactically clear of nearby convective or near-convective activity, and the aural sources provide additional big picture as well as trend data. In particular, some test subjects reported that trend data such as cell direction and speed were most readily obtained from aural sources.

Conclusions

This flight experiment was designed to address how GA pilots use different weather information sources (conventional aural, window+aural, and display+aural) that reflect different operational conditions: VMC, IMC, and IMC augmented with a data-linked graphical weather display. This paper reports on preliminary data that were analyzed to assess the accuracy of pilots' cell identification, bearing and range estimates, and accumulated "Big Picture" weather situation awareness.

These early results emphasize the benefits of graphical weather information systems for improved weather situation awareness, and indicate that the design of such systems must consider how pilots interpret weather depictions and attend to the latency of this information. These results also indicate that the three experimental conditions of weather information sources serve complementary purposes for pilots. GWIS design should strive to incorporate the improved interpretive information

available, potentially, in conventional aural information, and the veridicality and immediacy of an out-the-window view. These results have implications for instructional design for pilot training as well as operational guidance for usage of GWISs to capitalize on the advantages of aural, visual, and displayed cues and to understand where they may be incomplete and/or misleading.

References

[1]AOPA Air Safety Foundation,1996, *Safety Review: General Aviation Weather Accidents - An Analysis & Preventive Strategies.* AOPA Air Safety Foundation: Frederick, MD:

[2]Bussolari, S.R.,1994, Mode S data-link applications for general aviation. *The Lincoln Laboratory Journal, 7*(2), 313-328.

[3]Kelly, W.,2000, EWxR, An Enhanced On-board Weather Information System. *NASA Weather Accident Prevention Project Annual Review Proceedings*, Hampton, VA. May 23-25.

[4]Bernays, D.J., Dershowitz, A., Lind, A.T. & Bussolari, S.R.,1993, The development of a data link-compatible graphical weather service at MIT Lincoln Laboratory. *The Symposium on Worldwide Communications, Navigation, and Surveillance*, Reston, VA.

[5]Wickens, C.D.,1984, *Engineering Psychology and Human Performance.* Scott, Foresman & Co: Boston.

[6]Potter, S.S., Rockwell, T.H., McCoy, C.E.,1989, General aviation pilot error in computer simulated adverse weather scenarios. *The Fifth International Symposium on Aviation Psychology*, 570-575. Ohio State University Press: Columbus, Ohio.

[7]Lind A.T., Dershowitz, A. Chandra, D. & Bussolari, S.R.,1997,. *The Effects of Compression-Induced Distortion of Graphical Weather Images on Pilot Perception, Acceptance, and Performance*. Lincoln Laboratory Report ATC-243, MIT: Lexington, MA.

[8]Lind, A.T., Dershowitz, A. & Bussolari, S.R.,1994, *The Influence of Data Link-Provided*

Graphical Weather on Pilot Decision-Making. Report No. DOT/FAA/RD-94/9, NTIS: Springfield, VA.

[9]Wanke, C. Chandra, D., Hansman, R.J. & Bussolari, S.R.,1990, A comparison of voice and datalink for ATC amendments and hazardous windshear alerts. *The Fourth International Symposium on Aviation Safety*. Toulouse, France.

[10]Wanke, C. & Hansman, R.J.,1992, Hazard evaluation and operational cockpit display of ground-measured windshear data. *Journal of Aircraft, 29*, 319-325.

[11]Chandra, D.C. Bernays, D.J. & Bussolari, S.R.,1995a, Field evaluation of data link services for general aviation. *Transactions of the IEEE*, 258-263.

[12]Chandra, D.C., Banis, K.J. & Bussolari, S.R.,1995b, Design of a human interface for general aviation data link applications. *The Eighth International Symposium on Aviation Psychology*, 142-147.

[13]Lind A.T., Dershowitz, A. Chandra, D. & Bussolari, S.R.,1995,. The effect of data link-provided graphical weather images on pilot decision-making. *The Sixth IFAC Symposium on Man-Machine Systems*, Cambridge, MA.

[14]Chandra, D.C.,1995, Pilot evaluation of datalink services for general aviation. *The Eighth International Symposium on Aviation Psychology*, 679-685. Ohio State University Press: Columbus, Ohio.

[15]Talotta, N.J. *et al.*,1997, *A Field Evaluation of Data Link Flight Information Services for General Aviation Pilots.* Report No. DOT/FAA/CT-97/3, NTIS: Springfield, VA.

[16]Yucknovicz, D., Burgess, M., Heck, M., Novacek, P.,2000, Assessment of the effects of delayed weather information datalinked to the cockpit on pilot navigation decision making. *Presented at IEEE AIAA Digital Avionics Systems Conference.*

[17]Latorella, K.A., & Chamberlain, J.P. ,2001, Decision-making in flight with different convective weather information sources: Preliminary Results. In *The Eleventh International Symposium for Aviation Psychology*. Ohio State University Press: Columbus, Ohio.

[18]Wiggins, M. & O'Hare D.,1995, Expertise in aeronautical weather-related decision making: A cross-sectional analysis of general aviation pilots. *Journal of Experimental Psychology: Applied, 1*(4), 305-320.

[19]Hart, S.G. & Staveland, L.E.,1988, Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.) *Human Mental Workload*, 139-183. Elsevier Science: North-Holland.

[20] Buck, R.N., 1998, *Weather Flying.* McGraw-Hill: New York, NY.

[21] Newton, D., 1983, *Severe Weather Flying.* Aviation Supplies & Academics, Inc.: Renton, Washington.

Acknowledgements