JOURNAL OF AIRCRAFT
Vol. 38, No. 5, September–October 2001

Development and Evaluation of a Cockpit Decision-Aid for Emergency Trajectory Generation

Ted L. Chen* and Amy R. Pritchett†
Georgia Institute of Technology, Atlanta, Georgia 30332-0205

The application of intelligent cockpit systems is examined to aid air transport pilots at the task of planning and then following a safe four-dimensional trajectory to the runway threshold during emergencies. The design of a proof-of-concept system is described, including the use of embedded fast-time simulation to predict the trajectory defined by a series of discrete actions, the models of aircraft and pilot dynamics required by the system, and the pilot interface. Then results of a flight simulator evaluation with airline pilots are detailed. In 6 of 72 simulator runs, pilots were not able to establish a stable flight path on localizer and glideslope, suggesting a need for cockpit aids. However, results also suggest that, to be operationally feasible, such an aid must be capable of suggesting safe trajectories to the pilot; an aid that only verified plans entered by the pilot was found to have significantly detrimental effects on performance and pilot workload. Results also highlight that the trajectories suggested by the aid must capture the context of the emergency; for example, in some emergencies pilots were willing to violate flight envelope limits to reduce time in flight, in other emergencies the opposite was found.

Introduction

RESPONSIBILITY for the safe completion of a flight rests primarily with the pilot in command. During emergencies onboard air transport aircraft, this responsibility can be demanding, due to the large number of tasks to which the pilot must attend, including detecting and resolving failures in aircraft systems; continuing to monitor aircraft system health; coordinating with cabin crew, airline dispatchers, and air traffic control; controlling the aircraft; and deciding on (and then following) a course of action that will result in a safe landing. This inherent difficulty is compounded by a significant number of stressors, including physical danger, an uncomfortable physical environment (heat, smoke, noise, etc.), an overwhelming amount of information to consider, and the need to make decisions in a short period of time. In addition, the aircraft may have degraded performance and handling qualities, limiting the extent to which the pilot's past experience is relevant to the present problem.

The objectives of this research were to investigate how pilots generate and then follow a four-dimensional trajectory to the runway threshold during emergencies and to examine the functions needed in pilot aids for these tasks. This paper first presents relevant research from a number of domains, highlighting the important aspects of these tasks, pilots' needs in cockpit aids, and available technologies. Then, the design of a prototype aid is described. The results of a flight simulator evaluation with airline pilots are detailed. The paper concludes with a discussion of pilot performance at these tasks and design recommendations for future cockpit systems.

Background and Motivation

Once an emergency condition exists, effective generation of a safe trajectory (and then following this trajectory) becomes crucial to a safe landing. If done well, this can prevent a serious failure from evolving into an accident; if done poorly, a comparatively minor problem can lead to aircraft damage and fatalities. This trajectory must address multiple conflicting objectives including minimizing to time-to-land, bounding stress on the aircraft imposed by maneuvering, meeting airspace and regulatory limits and flight envelope limits, and ensuring the plan is robust against uncertain and unpredictable elements of the environment.

In this paper, emergency trajectory generation is defined as the determination of a course of action with specific detail to describe all aircraft dynamic states for the remainder of the flight. This combination of a high level of detail and a long timescale differentiates it from other types of trajectory generation and standard methods of flight planning. For example, strategic planning activities such as flight planning share extended timescales with emergency trajectory generation, but utilize a low-fidelity representation of the aircraft.[1] More specifically, plans generated through strategic planning are often described by waypoints and altitude crossings, not through a detailed trajectory. A common example of a strategic planning aid is the flight management systems (FMS) currently found in modern air transport aircraft; air traffic control instructions and flight plans are also typically at this level of detail. Likewise, whereas time-critical planning requires the same detailed aircraft model as emergency trajectory generation, it does so over a timescale on the order of seconds to minutes.[1] Because of the limited timescale, such time-critical plans usually encompass only a single action or maneuver that meets a singular goal. Cockpit systems that provide this level of planning include the traffic alert and collision avoidance system (TCAS), ground proximity warning system (GPWS), and Rotorcraft Pilots Associate's actions on contact functions.[2]

Emergency trajectory generation instead falls under the definition of tactical planning proposed in Ref. 1. This type of planning requires both a high level of detail and a long timescale to avoid generating a trajectory that is later found to be lacking. For instance, not including the detailed effects of aircraft dynamics may result in a delayed landing due to missed localizer or glideslope intercepts (when assumptions about turn rate, descent rate, etc., cannot be met), or the execution of an overly extended flight path (when maximum performance maneuvering is not used by the flight plan). In an emergency, either of these situations can be a serious detriment to the safety of the flight.

The representation of a plan used in this study was that of a procedure. Specifically, a flight plan and its associated trajectory were defined and communicated as a series of actions, for example, turn to heading 300 or descend to 8000 ft, initiated by discrete triggers and linked by the aircraft's continuously evolving dynamic states. This representation was chosen for several reasons. First, procedures are a common representation of tasks in high-workload, complex environments, including aviation.[3,4] Second, trajectories

Received 9 March 2001; revision received 21 June 2001; accepted for publication 22 June 2001; presented as Paper 2001-4268 at the AIAA Guidance, Navigation, and Control, Montreal, 6 August 2001. Parts published at AIAA Aerospace Sciences Meeting, Reno, NV (2000-0300).

*Graduate Research Assistant, School of Aerospace Engineering.

†Assistant Professor, Schools of Industrial and Systems Engineering and Aerospace Engineering.

are typically represented in civil aviation as procedures, with published charts dictating, for example, the turns, descents, and speed changes demanded by specific arrival routes and approaches; therefore, a cockpit aid using this representation in emergencies would provide a familiar view to pilots and establish a flying task for which pilots are already highly trained. Finally, because this representation is so prevalent in nominal operations, autopilots and FMS have been designed to fly the aircraft by initiating distinct new control behaviors and target states at discrete points.

The time or place each action should be initiated, and its severity, for example, the rate of a turn, descent rate, etc., are dependent on the aircraft trajectory and states. For example, the time to start a turn onto the final approach course and the required rate of turn, are dictated by aircraft speed through its impact on turn radius. As multiple actions are placed in series, a cascading effect ensues, with each action altering the aircraft trajectory and dynamic states at the time of subsequent actions. Continuing the example, a high-rate descent preceding the turn onto the localizer can increase the airspeed, which subsequently increases the turn radius and, therefore, may require changing the inbound course, which will subsequently affect the distance traveled and the descent rate need to reach glideslope intercept altitude, etc. This complex coupling prevents the decomposition of the trajectory into separate independent flight segments. Additionally, the coupling between such properties as descent rate, speed, and turn radius prevents the separation of the plan into lateral and vertical components. This makes it difficult to plan a complete set of actions for the entire arrival and approach.

Generation of a detailed emergency trajectory can, therefore, be viewed as a task that may prevent problems such as taking too long to land (important in smoke and fire situations) or requiring extreme maneuvers to intercept the localizer and glideslope (important in situations with degraded aircraft stability and maneuverability). Whereas several studies have examined replanning in general,[1,5–7] and military tactical planning aids in particular,[2] little experimental data exist on how air transport pilots plan a trajectory in emergencies. Likewise, cockpit voice recorder transcripts and accident reports provide only sparse and anecdotal evidence of how pilots perform this task.

The current literature on human decision making suggests that detailed trajectory generation is a very difficult task for pilots, as illustrated by two models of decision making and planning. A rational, analytic model of planning assumes the sequential process of 1) generation of alternatives, 2) imagining the consequences, perhaps through the process of mental simulation, 3) valuing (or evaluating) the consequences of the alternatives, and 4) choosing one alternative as a plan.[8] Models describing observed human behavior in a variety of domains suggest that experienced operators, such as pilots, rely substantially on nonanalytic strategies such as those defined by the recognition-primed decision model.[9,10] Through the use of pattern matching and recognition techniques, these nonanalytic strategies have the advantage of rapidly providing a starting plan that may then be iteratively improved as circumstances allow. For pilots, this method works well in situations covered by their training and experience. However, the effective implementation of this method is reliant on three assumptions: 1) The pilots have sufficient experience, training, and intuition with very similar situations to select a reasonable initial plan of action. 2) The pilot is able to evaluate quickly and correctly the consequences of the plan. 3) The detection of any bad decisions occurs early enough for the pilot to select and evaluate an alternate feasible course of action.

Both types of decision making models note the need for pilots both to identify a reasonable initial plan of action and to evaluate or predict the consequence of that plan of action. However, each emergency situation is highly unique: Each occurs in a different place with a different underlying cause, different goals, and different obstacles to a safe landing. For example, one situation may demand a safe path to a nearby airport with a damaged aircraft; another situation may require the quickest trajectory to a far-away airport.

Some aspects of pilot training may be relevant to these tasks: Specifically, in initial training on single engine aircraft in visual conditions, pilots are required to demonstrate the ability to execute a forced landing in a field in simulated engine-out conditions. However, more advanced training programs typically emphasize nominal operations, in which aircraft trajectory is dictated by published air routes and FMS calculations, rather than determined by the pilot. These programs also emphasize the procedural aspects of emergency responses, such as executing the correct procedures for specific emergencies; however, the common last step of emergency procedures is to land as soon as possible, which does not provide detail as to what the landing trajectory should be. Extensively training pilots on all aspects of trajectory generation would be difficult, given the large number possible situations that would need to be covered.

Therefore, the task of identifying an initial feasible guess for a trajectory cannot be completely trained for, and instead presents pilots with an active and intensive task with only general guidelines as an aid. Likewise, the task of evaluating the performance expected of a planned trajectory is very difficult, given the magnitude of predicting all facets of a highly detailed trajectory all of the way to the runway and the aforementioned limits on decomposing the trajectory into manageable parts.

Unlike the time-critical and strategic planning aids mentioned earlier, no cockpit decision aid exists that directly addresses the needs of emergency trajectory generation. Several cockpit aids intended for other purposes have some applicability. The first are charts and approach plates, which depict published air routes and approach procedures. The trajectories they present are not represented with a high level of detail and are formulated to meet criteria such as traffic flow, which may not be relevant during an emergency; however, they still provide a baseline plan and act as a source of trajectory limits imposed by factors such as terrain. For pilots of transport aircraft equipped with glass cockpits, additional planning aids are available in the form of the trend vector and the altitude range arc, providing accurate turn radius and bottom-of-descent information. However, these are of limited planning use because they are based solely on current aircraft states and, hence, can neither depict the impact of future actions nor indicate whether current actions will ultimately contribute to a safe landing.

At this time, the level of automation most appropriate for this task (i.e., which of the functions the aid should take over, and the ability of the pilot to override the system and/or modify its suggestions) is not known.[11,12] The earlier discussions of decision making highlighted two functions that an aid may perform: identifying a reasonable initial plan of action and evaluating the consequences of those actions. However, other issues must also be considered in assigning the role and function of the aid because of the impact they can have on the pilots' interaction with it. Studies of operator interaction with automated systems have repeatedly identified cases where automated or intelligent systems are not used because they do not bring sufficient benefits to the situation to warrant the time and effort required to use them, a condition commonly called underreliance. Conversely, if the aid is capable of completely taking over a task, operators are prone to either completely rely on the system without verifying its accuracy and appropriateness to the immediate context (a condition commonly called overreliance or misuse), or to be biased by the output of the aid to the point that they can not reason independently (a condition commonly called automation bias).[13–16] For example, in a study of a cooperative flight planning system (examining strategic planning), roughly 40% of pilots were induced to select poor flight plans by the introduction of faulty system information.[7]

This suggests that greater understanding is required of how pilots plan their flights in emergencies and what interventions can be made to aid them and to encourage more-detailed trajectory generation. This study focused on the use of an intelligent cockpit system to examine both these research needs: Interaction with such a system in a flight simulator test provides a preliminary assessment of the qualities and functions pilots require from such a tool and also forces pilot to demonstrate actively and verbalize their approach to planning.

It is envisioned that pilots will use trajectory-generation aids such as the one described in this paper after the decision to land is made. While the aircraft is heading to the destination airport, the pilot not flying will utilize the aid to plan a feasible set of actions for the arrival, approach, and landing. At this point, before committing to any plan, the flight crew can review its consequences on the trajectory. After final acceptance of a plan, the pilots will then fly the plan, either manually using the aid as a reference or through an automatic control system commanded by the aid. Pilots may also opportunistically improve the trajectory, or if the trajectory is found lacking, purposefully revert back to planning.

Beyond the benefits noted earlier in ensuring that near-term actions will lead to a safe landing, this emphasis on first planning and then flying has distinct advantages to the pilots given the cognitive demands they face.[10,17] Planning is a highly cognitive activity demanding the pilots' full attention; as such, it is often limited to preflight and isolated (preferably low-tempo) periods of the flight. By generating the plan, the pilot then makes the subsequent flying task easier by producing a reference trajectory to follow without continuous involvement and replanning.

This cockpit-decision aid complements other recent research efforts. For example, several studies have examined the fault-detection and fault-management processes also associated with emergencies.[18–20] Likewise, several studies are examining the control technologies that can help a pilot fly a reference trajectory (or automatically control the airplane) when the aircraft's handling qualities have degraded.[21,22]

Design and Development of a Prototype Planning Aid

This section outlines the development of a prototype called the emergency flight planner (EFP). This prototype was intended to test the feasibility of providing pilots with a tool that could effectively predict the complex interactions between the actions of a plan. Because no such tool has been documented for this application, this prototype also serves as a means by which to assess the automatic functions and capabilities needed by pilots. A schematic of a complete planner system and the subsystems it requires is shown in Fig. 1. The core functionality of the planner is the ability to predict the aircraft trajectory resulting from a given plan, that is, list of actions. This implies the need for models of the aircraft's dynamics and the pilot's control behavior. A pilot interface is also required.

Because this study sought to assess the utility of the planner to pilots through a controlled flight simulator study, this prototype implemented the subsystems shown by bold blocks in Fig. 1. In the simulator, exact knowledge of aircraft dynamics were used in lieu of aircraft model identification; in an operational flight planner, information regarding the performance degradations of the aircraft would need to be obtained through real-time system identification or directly from the aircraft controller that is compensating for the failure. Likewise, prescripted plans were used because automatic plan generation would require further developments in current methods for hybrid-system analysis and optimization.[23] Specifically, standard methods of optimal trajectory calculation, such as numerical solutions of the Hamilton–Jacobi–Bellman equations, are not well suited to the four-dimensional, hybrid dynamics created by the combination of discrete actions and continuously evolving maneuvers. Likewise, existing solutions to discrete systems cannot accommodate the continuous trajectory segments, and the complex interactions between the discrete and continuous elements prevent their separation into two individual problems.

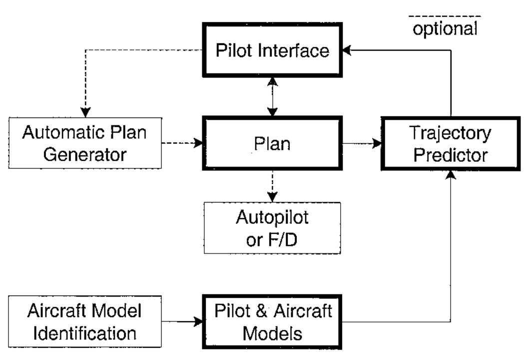

Fig. 1 Schematic of EFP.

Table 1 Arrival and approach actions incorporated in EFP

Attributes	Vertical	Speed	Miscellaneous
Turn to heading	Descend to altitude	Set speed	Set flaps
Fly to a fix	Maintain vertical speed	Set throttle	Set gear
Intercept localizer	Intercept glideslope		

Actions and Trajectory Prediction

The trajectory-defining actions included in the EFP are those relevant to an arrival and approach to an airport, as shown in Table 1. Three types of discrete action triggers were available: elapsed time, aircraft location over a position fix, or elapsed time past a fix.

In predicting the future trajectory with the detail required of tactical plans, the discrete actions must be joined by accurate predictions of the continuously evolving aircraft dynamic state. To meet these needs, the EFP used fast-time simulation to propagate the trajectory forward in time. The differential equations for the pilot–aircraft system are propagated forward, with the triggering of actions changing aircraft dynamics, commanded controls or target states at discrete points in time. For computational efficiency, the EFP utilizes a modified adaptive-timestep Runge–Kutta fourth-order (RK4) algorithm. Standard adaptive-timestep RK4 algorithms maximize the timestep of a continuous system while bounding numerical integration error; however, its timesteps may skip over the triggering of new discrete actions. The modified algorithm, therefore, queries all active actions for an upper bound on the timestep and compares it with that suggested by adaptive-timestep RK4. The EFP extrapolates most 30-min trajectories in less than 2 s on a 450-MHz desktop personal computer.

Representing Pilot and Aircraft Behavior

The trajectory predicted by the fast-time simulation is a product of both the aircraft dynamic behavior and the control behavior expected of the pilot and/or aircraft control system. Research has shown that pilots adapt their control behavior in response to changes in the underlying aircraft dynamics to maintain a consistent closed-loop behavior; many adaptive controllers intended for flight following failures are intended to do the same.[21,22] To replicate these control and dynamic behaviors, elaborate models of the aircraft dynamics and of control behavior may be sought for all failures over all flight conditions. However, these models have obvious cost and complexity penalties; in addition, the behavior of an elaborate control models, if correct, would typically only serve to cancel out changes in the aircraft dynamic model. Therefore, the EFP used a static representation of control behavior and of aircraft dynamics that fits the stable closed-loop behavior achieved with adaptive control under a range of failures.

For the aircraft dynamics, the EFP prototype uses a stable four-degree-of-freedom dynamics model: The longitudinal forces are thrust and drag, pitch and roll moments were governed by the ailerons and elevator, and coordinated flight was always assumed, thereby dictating side force and yaw moment. Failures can be created by reconfiguring aerodynamic coefficients within the model; these effects were selected to represent predicted changes in aircraft performance, as opposed to changes in aircraft stability. Stability and control constraints were modeled as limits imposed on the pitch angle, bank angle, and speed of the aircraft.

The aircraft control is handled by a collection of individual controllers for pitch, roll, and throttle. These are swapped in and out in the same manner as autopilot modes. They control the aircraft toward the target states specified by the active actions and keep the aircraft within the pitch, bank, and speed limits demanded by the aircraft dynamic model.

Pilot Interface

Obviously, many pilot interface designs are possible; at a minimum, they must accept action and trigger information from the pilot

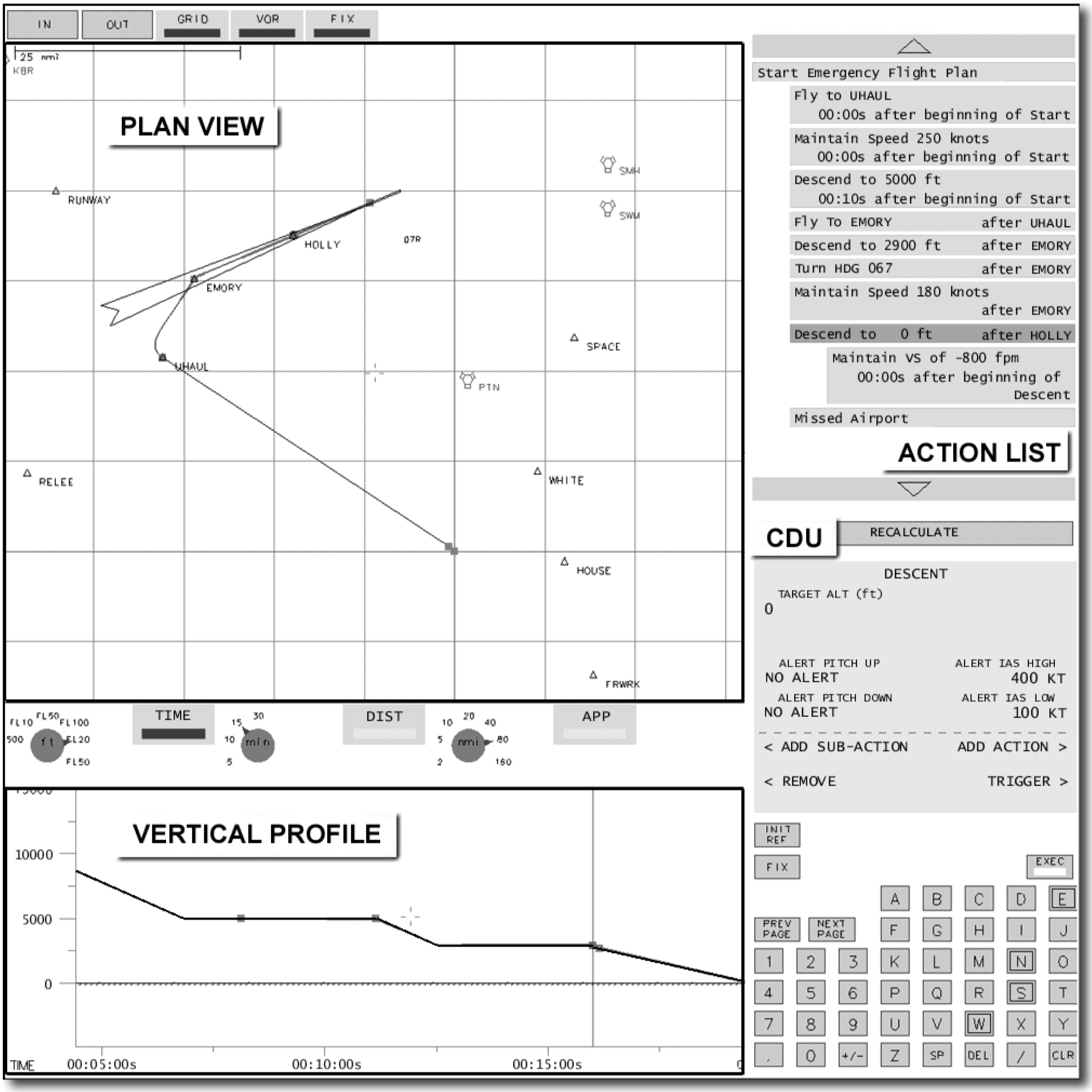

Fig. 2 EFP pilot interface (inverted black and white view for clarity).

and display the predicted trajectory to the pilot in such a way that the pilot can both assess the performance of the plan and then execute it. The pilot interface used with the EFP is shown in Fig. 2. All action specific information is located on sidebar on the upper right, providing a chronologically sorted list of the actions and their triggers. The primary input device is a control display unit (CDU), a common interface for air transport aircraft equipped with FMS. In the EFP, it provides a detailed textual display of a selected action and is the entry device by which pilots can modify actions and select functions.

The predicted trajectory was displayed to the pilot on two spatial displays (the plan and vertical profile views) using a format analogous to that on pilot charts and approach plates. The trajectory is normally shown in white, except for any segments that violate flight envelope or stability constraints, which are shown in red. The current location of the aircraft is also displayed, allowing the pilot to monitor conformance to the plan. The plan view is a scalable and scrollable north-up representation, with symbology based on the Boeing 747-400 electronic horizontal situation indicator (EHSI). Although this view could be conceivably integrated with smaller existing EHSI displays, issues regarding clutter and resolution would need to be addressed. There is no widely used vertical profile display in air transport cockpits at this time, and no one best display format has been experimentally demonstrated. Therefore, the EFP provides three pilot-selectable formats for the vertical profile display: The time view displays trajectory altitude with respect to the elapsed flight time, the distance view displays altitude along an unwrapped ground track, and the approach view provides a projection along the localizer beam, similar to that found on an approach plate.

Because the trajectory has been simulated using reasonably detailed dynamic models, the EFP can also display to the pilot a complete picture of aircraft state at any point in the future trajectory, including attitude, throttle settings, flight envelope limits, fuel status, airspeed, and aircraft configuration. The query view, shown in Fig. 3, displays this information at any point in the trajectory as selected by the pilot using a presentation similar to a glass cockpit

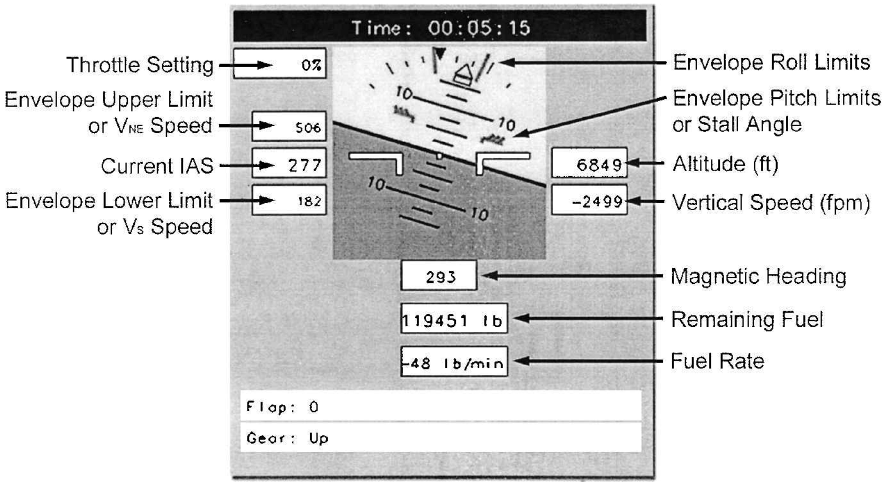

Fig. 3 EFP query view (inverted black and white view for clarity).

primary flight display (PFD). While planning, the pilot can select any point on the trajectory to see the aircraft state predicted there; while flying the aircraft, the query view can be set to automatically display the aircraft state at the point on the EFP trajectory closest to the current aircraft location.

Automated Functions

For the preliminary study described in the remainder of this paper, two variants of the EFP were created. The basic EFP variant provides a mechanism by which pilots can enter a plan, from which the system then predicts the ensuing trajectory. The preloaded EFP variant additionally provides automatic planning functions by presenting the pilots, at the start of planning, with a preloaded plan that they can accept, modify, or delete. Both variants were otherwise identical, with the same interface, method of predicting the trajectory, etc.

Experiment Design

The EFP was tested in a part-task, desktop flight simulator with airline pilots as subjects. Each pilot participated in two consecutive experiments in one session. The goal of the primary experiment was to investigate how pilots approached the planning task with and without the EFP, to determine quantitatively whether either variant of the EFP aided the pilots in landing safely following a major system failure or emergency, and to gather the data needed to improve the design of in-flight planners. The secondary experiment comprised a single deviant scenario in which the EFP had an erroneous model of the aircraft dynamics and, hence, made erroneous predictions of what a plan's associated trajectory would be. This tested the effect that such an error in the planner would have on the ability of pilots to execute a safe flight; given the sizeable evidence suggesting problems with automation bias, the hypothesis for the second experiment was that pilots would follow the erroneous trajectory prediction, with corresponding drops in performance.

Primary Experiment Independent Factors

In the primary experiment, the following two different factors were examined. The first factor was the planning tool, which comprised three conditions:

1) The baseline condition was charts-only, where pilots were provided with traditional paper en-route charts, STAR charts, and approach plates of the region of interest. An E6B-type flight computer (a circular slide ruler) was also made available.

2) The second condition supplied the basic EFP in addition to standard paper charts and E6B. Specifically, on startup, the basic EFP presented the pilot with an empty action list, to which the pilot could enter actions to create a trajectory.

3) The third condition supplied the preloaded EFP in addition to standard paper charts and E6B. Specifically, on startup the preloaded EFP presented the pilot with a feasible trajectory, which the pilot was able to accept, ignore, clear, or modify as desired.

The second factor was scenario type, which comprised two scenarios:

1) The performance altering (PA) scenarios created conditions in which the pilot needed to plan (and then fly) a trajectory in which the aircraft had substantially different performance from nominal. The failures were engine failure, stuck rudder, and inadvertent spoiler deployment.

2) The nonperformance altering (NPA) scenarios created conditions in which aircraft performance was currently nominal, but a compelling need existed for an immediate emergency landing. The failures were smoke in the cabin, cargo fire, and medical emergency.

Secondary Experiment Independent Factor

The secondary experiment had only one independent factor: The same three tool types as used in the primary experiment. The secondary experiment was restricted to a single PA deviant scenario (asymmetric loss of outboard aileron) in which the ability to turn to the left was diminished, but the EFP showed the opposite information, used this erroneous information in predicting the future trajectory and, in the case of the preloaded EFP, suggested an erroneous trajectory.

Test Matrix

Each pilot completed a total of seven scenarios. The first six runs spanned all six combinations of independent factors (three tool types times two scenario types) in the primary experiment; the final, seventh run, used the secondary experiment's deviant scenario, with pilots equally divided among the three tool types. The orders of the runs were blocked by tool type to mitigate any learning effects due to increased familiarity with any tool.

Experiment Apparatus

The experiment was conducted at Georgia Institute of Technology utilizing the reconfigurable flight simulator software running on two networked desktop workstations, each with a 19-in. monitor.[24] One workstation and monitor set was dedicated to the EFP. The other workstation and its monitor provided the pilot with cockpit instruments, including a PFD, EHSI and engine indicating and crew alerting system (EICAS), all based on B747-400 displays. Additional envelope limits for roll, pitch, and speed were depicted on the PFD using the same format as the query tool, shown in Fig. 3. Control of the aircraft was enabled through a side stick and throttle, whereas the EFP used a cursor controlled by a trackball.

Experiment Procedure and Scenarios

Following a briefing and two training runs, each pilot was asked to fly the seven data-collection runs specified by the test matrix. For each run, the pilots were told that they were captains of a Boeing 747-400, that an emergency had occurred, and that all relevant emergency checklists had already been performed. In all scenarios, the aircraft was in instrument meteorological conditions with no terrain or traffic considerations. Each run was split into two parts. During the first part, pilots were asked to plan their approach to the airport for 15 min using the available tools; this period was described as an interval where the first officer (not actually present at the experiment) was holding the aircraft in a descent toward a hand-off point nearer the airport. The pilot was asked to verbalize the criteria and methods applied in building each plan.

The second part then required the pilot to take control of the aircraft at the hand-off point, steer it onto the localizer and glideslope of the landing runway, and maintain the approach until 500 ft above the runway threshold. The aircraft dynamic model of the simulator was the same as that in the EFP with one exception: In the deviant scenario, the aircraft model underlying the simulator flown by the pilot utilized a different dynamic model from the EFP.

To avoid pilot familiarity with an airport, all scenarios involved fictitious airports. Whereas all scenarios shared a common airspace structure and were intended to be of similar difficulty, slight differences in orientation and starting conditions were created to prevent learning effects. The starting conditions of all scenarios were calibrated such that the preloaded EFP plan utilized similar amounts of aircraft maneuvering and programming effort. Additionally, the preloaded plans were constrained to be within a flight time of 13–14 min and a track distance of 55–65 n mile, while staying within all published attitude and speed limits; these plans had 13 or 14 actions each, including several configuration actions for extending the gear and each stage of flaps.

Subjects

There were 12 airline pilots participating in this study. All had prior experience with FMS and moving map displays. Of the 12 pilots, 8 were captains, and 4 were first officers. Average flight hours were 14,000 and 8,600 h for the two groups, respectively. Total flight hours ranged from 3,800 to 25,000 h. All but one had received military flight training.

Primary Experiment Results

A total of 72 runs were performed in the primary experiment. Unless otherwise specified, the data sets were analyzed for tool and scenario-type effects by fitting to a general linear model. The tool and scenario type were analyzed as fixed effects; pilots were analyzed as a random factor to allow the results to be generalized to the entire population of pilots. In addition, the general linear model also tested for interactions between the factors. Where significant variation was found, more specific tests identified significant differences, including one-way analysis of variance (ANOVA) and the Tukey multiple comparison procedure with 95% confidence intervals (see Ref. 25). To test the residuals of the fit for the normality assumptions of these tests, the Kolmogorov–Smirnov normality test was applied (see Ref. 26). In cases where the assumption of normality for the data did not hold, a nonparametric Kruskal–Wallis test was performed (see Ref. 26).

Pilot Performance in Planning and Flying Trajectories

The number of missed approaches (here defined as a situation where the pilot could not establish a stable flight path on both the localizer and glideslope by 500 ft above ground level) is an important measure of safety and pilot performance. A missed approach entails the aircraft having to circle for another approach, adding significantly more time and requiring additional low-altitude maneuvering. During the 72 runs, 6 instances of missed approaches were recorded. In five of these six instances, the pilot did not use the EFP as the primary reference. The only other instance of a missed approach occurred with a preloaded EFP variant in a PA scenario. In this case, the pilot did attempt to follow the plan given by the EFP.

Although the small number of samples precludes any rigorous statistical analysis, further insight may be gained by observing the underlying cause of the missed approaches. Of the six missed ap-

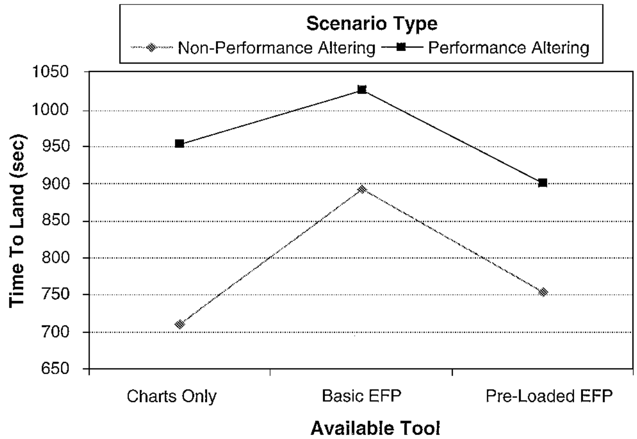

Fig. 4 Average time to land categorized by tool type.

proaches, four occurred during rapid-descent maneuvers in the time-critical NPA scenarios. In none of these runs did the pilot follow the plan in the EFP. One possible explanation for the high number of overly rapid descents is the lack of comprehension of the consequences of high descent rates and close-in instrument landing system (ILS) intercepts. The other two missed approaches were both PA scenarios with no apparent common denominator.

Another important metric of pilot performance is time to land; even in situations where time is not the highest priority, extending the duration of a flight is risky due to the unknown lifetime remaining in damaged aircraft systems. The average of the time to land measure (defined as the length of the pilots' flying time from the hand-off point to when the aircraft reached a height of 500 ft above ground level) is shown in Fig. 4, with one outlier data point removed.

An ANOVA and Tukey test found that the time to land for NPA scenarios is, on average, significantly lower than for the PA scenarios ($F = 18.80$, $p < 0.001$). The difference between NPA and PA scenarios' times can be attributed to the time-critical nature of NPA scenarios such as medical emergencies or fires. Conversely, pilots appear to be more conservative in PA scenarios for the sake of aircraft stability. In addition, analysis of the data found that the availability of the basic EFP variant resulted in a greater time than the other two tool options, as shown by a Kruskal–Wallis test ($H = 6.68$, $p = 0.035$).

From experimenter observations and pilot comments, it was noted that pilots did not always follow the EFP's plans, most likely due to several factors such as difficulty in entering a plan (in the case of the basic EFP) and concern regarding the adequacy of the preloaded plans (in the case of the preloaded EFP). This suggested that a more detailed factor could be used to provide more insight; results for both EFP variants were each broken down into two subcategories, one for whether the pilots at least partially used the EFP and the other for when the pilot did not follow the plan in the EFP at all. EFP usage was defined as situations where the pilot followed its plan for at least a portion of the flight, as judged by comparing track and vertical profile data from both the EFP plans and actual flight data. This created five distinct categories as shown in Fig. 5. The time to land values are referenced to the length of the unmodified preloaded EFP plan around which the scenario was designed.

ANOVA found significant variation between these five conditions ($F = 2.80$, $p = 0.033$). A Tukey test with 95% limits identified significantly higher times in cases where the basic EFP was used compared to the charts-only condition. The same test with weaker 90% confidence limits shows an increase over all of the other conditions (basic EFP not used and preloaded EFP used and not used). Analysis of the duration of the plan created within the EFP provides a possible explanation. A Kruskal–Wallis test showed significant differences in predicted duration between cases using the two different EFP variants ($H = 6.82$, $p = 0.009$); specifically, the plans created in the basic EFP were an average 1.5 min longer than the preloaded plans. Therefore, because the plans that pilots created in

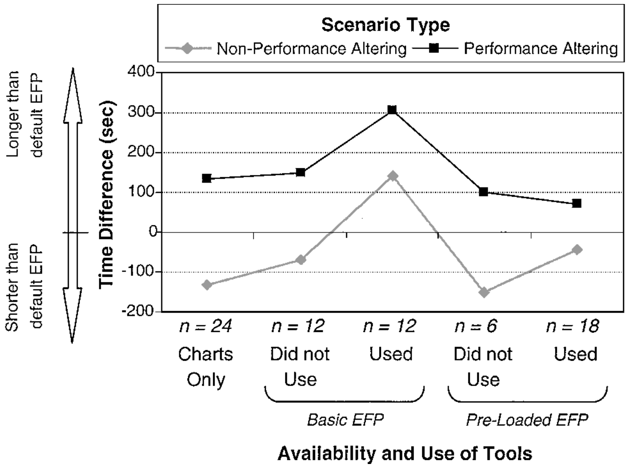

Fig. 5 Average time to land categorized by tool type and usage, referenced to the preloaded plan duration.

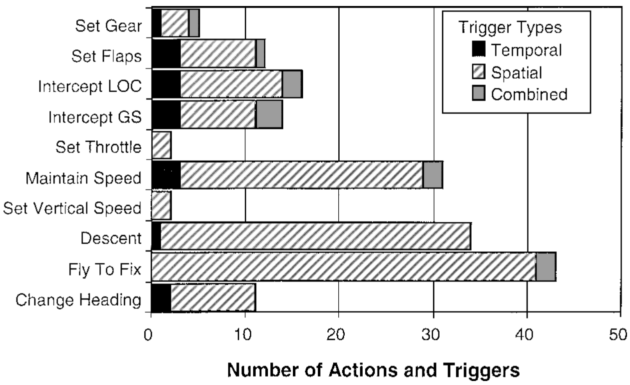

Fig. 6 Actions used in pilot-created plans, subdivided by trigger type.

the basic EFP were longer, adherence to them may have also caused a longer flight than required. No statistically significant differences were found between preloaded EFP and the baseline charts-only tool types.

Planning Constraints and Assumptions

Measures were also made into how pilots planned and flew in the different scenarios. Specifically, significant scenario effects were found in the number of violations of the placard flap and gear speed limits. In the NPA scenarios, where the emergency tended to be time critical, several pilots opinioned that exceeding the flap speed limits was acceptable given the assumption that approximately a 10-kn safety buffer was incorporated into the listed value. The data mirror their opinions, with a significantly higher number of flap violations in the NPA scenarios ($F = 4.47$, $p = 0.038$). However, the data also showed significant results for violations that were more than 10 kn over the listed value. With this revised limit, the NPA scenarios again had higher instances of violations with respect to the PA scenarios ($F = 6.09$, $p = 0.016$). In these cases, several pilots violated their own self-reported limits, apparently to land the aircraft as soon as possible.

These data provide two design insights. First, and most important, pilots' planning objectives change with the context of different emergency situations; correspondingly, flight envelope limits may also need to be relaxed in specific circumstances. Second, even with an undamaged aircraft, pilots may not fully realize the dynamic interactions between trajectory-defining actions and, therefore, may not plan a trajectory that does not exceed aircraft limits.

Actions and Triggers Used by Pilots in Creating Plans

The types of actions and triggers in plans created by the pilots using the basic EFP were recorded. Figure 6 shows the different types

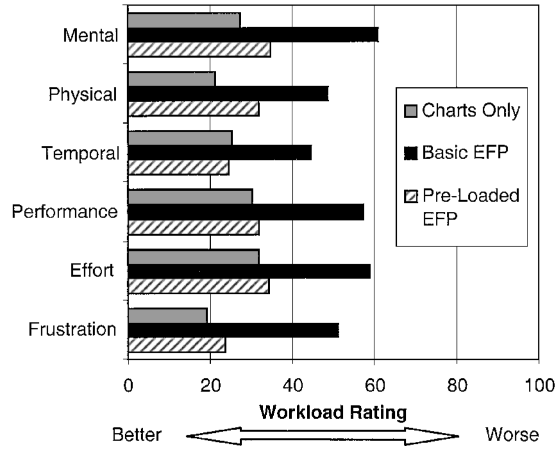

Fig. 7 Average TLX workload ratings for the planning task.

of actions with their cumulative total in all pilot-created plans, including plans that were ultimately not followed by pilots or were infeasible. A substantial number of fly to fix, maintain speed, and descend actions were used. Relative to the default preloaded EFP plans, which contained flap actions for every flap interval on the placard, fewer flap and gear actions were in pilot-created EFP plans. Throttle and vertical speed actions were also lacking from the user-created plans. Whereas these results may indicate pilot-preferred actions, the ability to infer the necessity of the other actions is confounded by both the training provided to the pilots and the EFP interface.

The actions are subdivided by their associated triggering criteria. Most of the actions used a spatial trigger, such as when the aircraft passes over a certain location; pilots often created their own fixes to serve as triggers, rather than querying the tool to identify the corresponding time. The lack of use of the temporal triggers suggests that pilots may prefer spatial representations in conceiving and visualization plans. However, the spatial display of the trajectory itself may have encouraged the use of spatial triggers because the only explicit portrayals of the time of any point in the trajectory were in the query view and in one mode of the vertical profile display.

Pilot Workload

In safety-critical tasks, performance measures are much more compelling than measures of pilot workload. However, workload can be taken as a measure of assistance that the cockpit aid provides to the pilot and as a contribution or detriment to pilot performance. Therefore, at the conclusion of each scenario, the pilots were asked to complete a NASA task load index (TLX) evaluation of workload experienced in both the planning and flying tasks.[27] As indicated by the average ratings shown in Fig. 7, the Basic EFP had higher workload ratings in each of the workload categories than either of the other planning tool types during the planning task; this result was found to be statistically significant by an ANOVA, Tukey test, and Kruskal–Wallis test to at least the 95% confidence level. A similar analysis was performed on the data from the flying stage. However, no differences due to the tool provided were found. The temporal workload measure did have significantly higher values ($H = 4.54$, $p = 0.033$) in the NPA scenarios as opposed to the PA scenarios, as expected.

Secondary Experiment Results

A total of 12 runs were performed in the secondary experiment (one per pilot); each of the three planning tools, therefore, was provided to four pilots for one run. Because of the small sample size, statistical analysis was not appropriate. However, qualitative analysis of the aircraft track data noted interesting trends when comparing EFP usage (which would cause an infeasible trajectory) against EFP nonusage. Results were grouped by whether the pilots had an EFP variant available and followed its trajectory. Four pilots appeared to follow the EFP's plan; of these, three pilots initially overshot the localizer similar to the sample track shown in Fig. 8. Conversely,

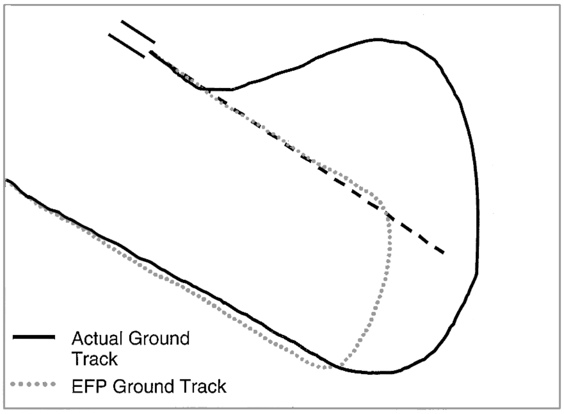

Fig. 8 Representative deviant scenario track.

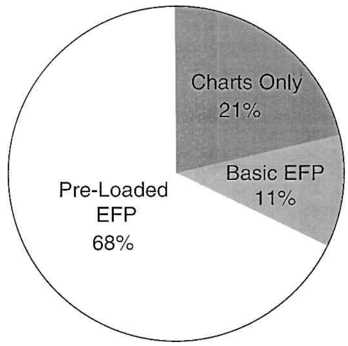

Fig. 9 Pairwise comparison of tool types, analyzed with AHP.

only two of the eight pilots not using the EFP overshot the localizer. Overshoots of the localizer often lead to additional maneuvering and unnecessary time and distance, with a corresponding frequent need for missed approaches. In the cases where the EFP was used, the corresponding localizer overshoot added an average 178 s to the flight time and an average of 12.2 n mile to the track distance when compared with situations where the EFP was not used.

Pilot Ratings of the EFP

At the conclusion of the two experiments, the pilots were asked to provide pairwise comparisons between the three different planning tools. The overall pilot preference shown in Fig. 9 was determined through the analytic hierarchy process (AHP).[28] The relative preference of any two tools can be obtained by taking the ratio of their respective areas. The preloaded EFP has a weak preference over the charts-only condition (68–21%) and a strong preference over the basic EFP (68–11%).

Conclusions

In summary, this research has investigated the tasks of generating and then following a detailed trajectory to the runway threshold in emergencies. Little data currently exist into how air transport pilots perform these tasks, the difficulties they face, and the desired features of a decision aid. This study provided a preliminary investigation of these questions by using a prototype decision aid to examine tool design considerations directly, to gather quantitative evidence about the utility of a prototype aid, and gather data about pilots' planning activities and needs in an intelligent cockpit system for this task.

The results suggest that pilots face problems in creating and comprehensively evaluating a trajectory. In 6 of 72 runs, pilots were unable to establish an approach course. Four of these occurred in aggressive rapid-descent maneuvers without guidance from the EFP. It is reasonable to hypothesize that, had the pilots been able to evaluate fully the adverse consequences of their current actions on their future trajectory, they would have decided to intercept farther away from the airport with a slower descent rate. In addition, the fact that only one of the six incidents occurred when the pilot was using the EFP provides very preliminary evidence that such a tool may be useful in reducing such errors.

Although such tools may be beneficial to pilots, problems found in the proof-of-concept prototypes tested in this study warrant further research and consideration during design. The first of these problems is related to the EFP's pilot interface, which primarily used a keyboard entry mechanism (through a CDU) that pilots described as being cumbersome and occasionally confusing. This suggests that merely attempting to leverage the existing cockpit systems such as the FMS by the addition of predictive routines for emergencies is not enough. A more streamlined interface is required that minimizes the amount of pilot workload required for this concept to be acceptable in an emergency environment.

The second problem associated with the prototype highlights potential issues with the functions the aid needs to perform. Significantly higher times to land were found in cases where the pilot was given the basic EFP. Therefore, simply providing a planning tool that evaluates a pilot-created plan may not be sufficient to guarantee generation of the safest trajectory, although this issue may have been compounded by problems with the interface in this study. The preloaded EFP variant simulated a planner capable of suggesting plans to pilots. Although its plans were not demonstrated to be optimal, it was found that the preloaded EFP still outperformed the basic EFP by every measure, including performance, workload, and pilot ratings.

Giving a cockpit system the ability to generate automatically and suggest plans to pilots raises several interesting research questions. In the deviant scenario, where the EFP provided the pilot with erroneous information, overreliance on the displayed trajectory was common. Conversely, that not all pilots followed the preloaded EFP's plans suggests that the potential also exists for underreliance. Commensurate with studies of other automated systems, pilots in this study reported not relying on plans suggested by the aid due to concern about their validity and the mechanism by which they were created. This suggests that not only does the suggested plan have to be in a clearly understandable form, but its underlying structure and objectives must also match those of pilots if over- and underreliance are to be avoided.

Therefore, the underlying goals and criteria used in automatic trajectory generation must conform to those used by the pilots. However, this study found that these factors change with the context of the emergency. For example, in NPA scenarios the pilots tended to violate overspeed limits in an effort to minimize flight time; in PA scenarios, on the other hand, pilots were generally not as willing to overspeed or overstress the aircraft. Capturing these context sensitivities faces several challenges: accurately eliciting these criteria from pilots, capturing them into a machine-readable representation, giving the system an awareness of the current context, and establishing mechanisms for pilots and the cockpit system to communicate about their criteria and perceived context.

Likewise, methods of representing and displaying the plan need to be examined further. In this study, plans were represented as procedures listing a series of trajectory-defining actions. Pilot comments appear to support this representation; for example, pilot-suggested changes to the display included building in cues to the pilot of newly triggered actions while flying the trajectory. However, in using this representation, many unanswered questions remain: What actions should be used to define the trajectory? What triggers should initiate them? This study considered only a small list of actions and triggers, some of which pilots used heavily and others which were used infrequently. Many other actions and triggers are possible, but to prevent overwhelming pilots with too many options, it will be important to identify those most relevant to the task at hand.

Other research questions address difficulties in automatically generating a plan. Common methods of optimizing trajectories typically require a clearly established objective function from which an absolute best trajectory can be identified. However, in emergency flight planning, a clearly specified objective function may not always be obtainable. Instead, the plan best meeting each several independent objectives and constraints must be found. Likewise, the objective function for these plans may include probabilistic concerns, such as finding a plan that is the most likely to meet all hard constraints in the face of future eventualities. Finally, the representation of a trajectory as being governed by discrete actions requires methods for rapidly optimizing complex hybrid systems.

A final research question examines this study's separation of the overall task into separate planning and flying stages. This delineation may be necessary for a pilot who is creating and flying a trajectory without automatic assistance. However, with the availability of intelligent aids, this distinction may no longer be necessary because the system may be capable of continuously improving the trajectory. In implementing such a system, not only would the appropriate generation routines need to be determined and incorporated, but also its impact on the pilot would need to be studied for the possibility of decreased situation awareness (if the plan is constantly changing without their awareness) and of increased cognitive load (if the pilot is frequently asked to consider new potential plans, diverting attention away from other tasks).

Acknowledgments

This work was funded by NASA Langley Research Center under Grant NAG 1-2014, with Anna Trujillo acting as Technical Monitor. The authors also wish to recognize the airline pilots who volunteered their time to participate in the flight simulator evaluation.

References

[1]Fan, T., Hyams, D. S., and Kuchar, J. K., "Study of In-Flight Replanning Decision Aids," *Proceedings of the Guidance, Navigation, and Control Conference*, AIAA, Washington, DC, 1998, pp. 980–988.

[2]Baker, B., "Rotorcraft Pilot's Associate (RPA) Attack Mission Planner," *AHS International Annual Forum*, American Helicopter Society, Alexandria, VA, 1999, pp. 844–853.

[3]Ockerman, J. J., and Pritchett, A. R., "A Review and Reappraisal of Task Guidance: Aiding Workers in Procedure Following," *International Journal of Cognitive Ergonomics*, Vol. 4, No. 3, 2000, pp. 191–212.

[4]Degani, A., and Wiener, E. L., "Procedures In Complex Systems: The Airline Cockpit," *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, Vol. 27, No. 3, 1997, pp. 302–312.

[5]Patrick, N. J. M., "Decision-Aiding and Optimization for Vertical Navigation of Long-Haul Aircraft," Ph.D. Dissertation, Dept. of Mechanical Engineering, Massachusetts Inst. of Technology, Cambridge, MA, Aug. 1996.

[6]Dershowitz, A., "The Effect of Options on Pilot Decision Making in the Presence of Risk," Ph.D. Dissertation, Dept. of Aeronautics and Astronautics, Massachusetts Inst. of Technology, Cambridge, MA, Oct. 1997.

[7]Layton, C., Smith, P. J., and McCoy, C. E., "Design of a Cooperative Problem-Solving System for En-Route Flight Planning: An Empirical Evaluation," *Human Factors*, Vol. 36, No. 1, 1994, pp. 94–119.

[8]Johannsen, G., and Rouse, W., "Mathematical Concepts For Modeling Human Behavior In Complex Man–Machine Systems," *Human Factors*, Vol. 21, No. 6, 1979, pp. 733–747.

[9]Klein, G. A., "A Recognition-Primed Decision (RPD) Model of Rapid Decision Making," *Decision Making In Action: Models and Methods*, edited by G. A. Klein and J. Orasanu, Ablex, Stamford, CT, 1993, pp. 138–147.

[10]Bainbridge, L., "The Change in Concepts Needed to Account for Human Behavior in Complex Dynamic Tasks," *IEEE Transactions on Systems, Man and Cybernetics—Part A: Systems and Humans*, Vol. 27, No. 3, 1997, pp. 351–359.

[11]Guerlain, S. A., and Smith, P. J., "Interactive Critiquing as a Form of Decision Support: An Empirical Evaluation," *Human Factors*, Vol. 41, No. 1, 1999, pp. 72–89.

[12]Miller, C. A., Guerlain, S., and Hannen, M. D., "The Rotorcraft Pilots Associate Cockpit Information Manager: Acceptable Behavior from a New Crew Member," *AHS International Annual Forum*, American Helicopter Society, Alexandria, VA, 1999, pp. 1321–1332.

[13]Wiener, E. L., and Curry, R. E., "Flight-Deck Automation: Promises and Problems," *Ergonomics*, Vol. 23, No. 10, 1980, pp. 995–1011.

[14]Parasurman, R., and Riley, V., "Humans and Automation: Use, Misuse, Disuse, Abuse," *Human Factors*, Vol. 39, No. 2, 1997, pp. 230–253.

[15]Pritchett, A. R., "Reviewing the Role of Cockpit Alerting Systems," *Human Factors and Aerospace Safety*, Vol. 1, No. 1, 2001, pp. 5–38.

[16]Schroeder, B., and Sarter, N., "Supporting Decision-Making and Action Selection Under Time Pressure and Uncertainty: The Case of Inflight Icing," *Proceedings of the International Symposium on Aviation Psychology*, Ohio State Univ., Columbus, OH, 2001.

[17]Xiao, Y., Milgram, P., and Doyle, D. J., "Planning Behavior and Its Functional Role in Interactions with Complex Systems," *IEEE Transactions on Systems, Man and Cybernetics—Part A: Systems and Humans*, Vol. 27, No. 3, 1997, pp. 313–323.

[18]Menke, T. E., and Maybeck, P. S., "Multiple Model Adaptive Estimation Applied to the VISTA F-16 with Actuator and Sensor Failures," *NAECON '92: Proceedings of the IEEE 1992 National Aerospace and Electronics Conference*, Vol. 2, Inst. of Electrical and Electronics Engineers, New York, 1992, pp. 441–448.

[19]Rogers, W. H., Schutte, P. C., and Latorella, K. A., "Fault Management in Aviation Systems," *Automation and Human Performance: Theory and Applications*, edited by R. Parasuraman and M. Mouloua, Lawrence Erlbaum, Mahwah, NJ, 1996, pp. 281–317.

[20]Deaton, J. E., and Glenn, F. A., "The Development of Specifications for an Automated Monitoring System Interface Associated with Aircraft Condition," *International Journal of Aviation Psychology*, Vol. 9, No. 2, 1999, pp. 175–187.

[21]Rysdyk, R. T., Leonhardt, B. K., et al., "Development of an Intelligent Flight Propulsion and Control System: Nonlinear Adaptive Control," *Proceedings of AIAA Guidance, Navigation, and Control Conference*, Aug. 2000.

[22]Bull, J., Mah, R., Hardy, G., Sullivan, B., Jones, J., Williams, D., Soukup, P., and Winters, J., "Piloted Simulation Tests of Propulsion Control as Backup to Loss of Primary Flight Controls for a B747-400 Jet Transport," NASA TM-112191, April 1997.

[23]De Schutter, B., and De Moor, B., "Generalized Linear Complementarity Problems and the Analysis of Continuously Variable Systems and Discrete Event Systems," *Hybrid and Real-Time Systems*, edited by O. Maler, Springer-Verlag, Berlin, 1997, pp. 409–414.

[24]Ippolito, C. A., and Pritchett, A. R., "Software Architecture for a Reconfigurable Flight Simulator," *Proceedings of AIAA Modeling and Simulation Technologies Conference*, Aug. 2000.

[25]Hogg, R. V., and Ledolter, J., *Applied Statistics for Engineers and Physical Scientists*, 2nd ed., Macmillan, New York, 1992, pp 264, 265, 271, 272.

[26]Siegel, S., and Castellan, N. J., *Nonparametric Statistics for the Behavioral Sciences*, 2nd ed., McGraw–Hill, Boston, 1988, pp. 51–55, 206–216.

[27]Bortolussi, M. R., Kantowitz, B. H., and Hart, S. G., "Measuring Pilot Workload in a Motion Base Trainer: A Comparison of Four Techniques," *Applied Ergonomics, Special Issue: Ergonomics in Aviation*, Vol. 17, No. 4, 1986, pp. 278–283.

[28]Yang, L. C., and Hansman, J. R., "Application of the Analytic Hierarchy Process for Making Subjective Comparisons Between Multiple Automation/Display Options," *6th IFAC/IFIP/IFORS/IEA Conference on Analysis Design and Evaluation of Man Machine Systems*, Vol. 2, Elsevier, New York, 1995, pp. 487–491.