Assessing Pilot Workload in Single-Pilot Operations with Advanced Autonomy

Alexander J. Stimpson[1], Jason C. Ryan[2], Mary L. Cummings[1]
[1]*Humans and Autonomy Laboratory, Duke University, Durham, NC, USA*
[2]*Aurora Flight Sciences Corporation, Cambridge, MA, USA*

The proposed transition to single-pilot operations (SPO) in commercial and military aircraft has motivated the development of advanced autonomy systems. However, a detailed analysis of the impact of advanced autonomy on pilot workload through various phases of flight and contingency scenarios has not been conducted. To this end, this paper presents the development of the Pilot-Autonomy Workload Simulation (PAWS), a discrete event simulation model that allows the investigation of pilot workload under a variety of advanced autonomy capabilities and scenarios. Initial utilization results from PAWS of nominal and off-nominal point-to-point missions demonstrate that the workload for a single pilot assisted by advanced autonomy varies considerably over different phases of flight and various contingencies. These results suggest that advanced autonomy to offset pilot workload is not needed for low-workload phases, but could be critical during periods of high workload.

INTRODUCTION

The economic advantages of single-pilot operations (SPO) and reduced-crew operations (RCO) and the projected pilot shortage through 2022 are motivating both government and commercial organizations to investigate technologies and approaches for realizing SPO (Comerford, et al., 2013; Croft, 2015). Organizations involved in this effort include U.S. organizations such as NASA and the Defense Advanced Research Projects Agency (DARPA), international researchers, and even commercial entities (Wolter & Gore, 2015; DARPA, 2014; Harris, 2007; McCartney, 2010). Some proposed SPO architectures incorporate an operator at a ground control station to offload some of the functions currently provided by a co-pilot e.g., (Wolter & Gore, 2015). Another proposed architecture includes the development of advanced in-cockpit autonomy to assist the remaining pilot (Schutte, et al., 2007).

Under either architecture, the development of advanced autonomy will be a key aspect of implementation. As a result, an important area of research is the determination of the roles of the pilot and the advanced autonomy (Wolter & Gore, 2015). Such allocations could be either static or dynamic (Parasuraman, Mouloua, & Hilburn, 1999). In static allocations, tasks are assigned to a fixed entity across the entire flight regardless of circumstances. In dynamic allocations, tasks can be shifted between entities based on timing or contingency scenarios. In any SPO human-autonomy architecture design, workload will be a key consideration in ensuring the appropriate tasks are assigned to the human pilot, which could be dynamic under different phases of flight. Simulation is a useful methodology for examining different potential allocation strategies for nascent systems, such as proposed SPO configurations (Laughery, Archer, & Corker, 2001). Additionally, such simulations can be used to identify phases of flight or circumstances under which functions of advanced autonomy may be critical for maintaining appropriate workload for the pilot.

To this end, this paper presents, for the first time, the development and initial results of the Pilot-Autonomy Workload Simulation (PAWS), which takes a utilization-based approach to estimating pilot workload under a variety of conditions. Such approaches have been used as a reasonable approximation for workload (Rouse, 1983; Cummings & Guerlain, 2007). To create this model, a generalized task-flow model of multi-crew flight operations was combined with information from subject matter experts (SMEs) to generate estimates for task completion time distributions for both nominal and off-nominal procedures.

PAWS utilizes a discrete event simulation (DES) representation of the task flow and distributions to create stochastic estimates of pilot workload over the course of a flight from takeoff to landing. Such DES approaches have been shown to be useful in estimating operator workload under a variety of environments, e.g., (Schmidt, 1978; Gao & Cummings, 2012; Jun, Jacobson, & Swisher, 1999; Xiling, Qishan, & Wei, 2002). PAWS allows for workload estimates to be generated based on varying capabilities of advanced automation such as primary flight control, communications functions, or checklist execution. Additionally, by inserting off-nominal or emergency procedures at varying points in a flight, an array of contingency scenarios can be simulated. The present paper discusses the creation of PAWS and preliminary data generated based on an initial set of nominal and off-nominal point-to-point missions.

MODEL DEVELOPMENT

The task model used for PAWS was created based upon a generic task flow of multi-crew aircraft constructed from standard flight procedures (including commercial and military aircraft flight manuals), as well as input from SMEs. This input was gathered through guided, open-ended, face-to-face interviews with eleven experienced commercial passenger aircraft pilots (Cummings, Stimpson, & Clamann, 2016), as well as six military pilots. Once the task model was created and verified through discussion with the SMEs, statistical distributions for the time required for each task completion were implemented and validated against minimum, average, and maximum task times provided by the SMEs. Lognormal distributions were utilized for completion times, based upon established literature for representing human task-completion times (Sheridan, 2013). The determination of the lognormal parameters for each task completion time was based upon matching the mean to the SME-indicated average task time, and the 95th percentile of the distribution to the SME-indicated maximum task time.

PAWS implements a DES approach to modeling pilot utilization, coded in Java. In a DES architecture, one or more agents (such as a pilot or advanced automation) act as servers for handling a queue of tasks, each with an associated completion time randomly drawn from the associated distributions. Tasks in PAWS focus on the achievement of subgoals, which generally include multiple-step procedures or actions and are serviced by each agent serially, with only one task serviced at one time by each agent. Detailed models of individual motions, such as Fitts' Law (Fitts, 1954) computations, are not required for PAWS as they are subsumed within the task-completion time distributions. Examples of tasks include setting the landing gear, providing an approach briefing, or providing an altitude callout.

The simulation environment populates all the tasks for the flight, and then assigns them to the agents based upon a predetermined role allocation for each task, assuming prerequisite conditions are met. Prerequisite conditions for each task may include:

- Phase preconditions, which require that the task be performed in the correct phase (e.g., call for gear up during the climb phase).
- Task preconditions, which require that a certain task be completed in a sequence (e.g., flipping a switch before verifying that the switch is in the correct position).
- Temporal preconditions, which require that a certain time is reached prior to initiating a task, and signal the random generation of task times from the associated distribution (e.g., communication with air traffic control (ATC)).

Phase preconditions for the point-to-point flight were based upon the division of a single flight into seven phases (representing takeoff to landing): before take-off check, take off, climb, cruise, descent/approach check, initial approach, and final approach. The simulation advances to the next phase after all tasks for the current phase have been completed, or a specified minimum phase time, whichever is longer.

The PAWS simulation takes inputs of a list of tasks (with associated completion distributions, preconditions, and task assignments), and the number of agents to complete the tasks. Note that an agent for this simulation could represent either a pilot or advanced automation. The task time distributions represent the amount of time that the agent must focus on the completion of the task. For example, for tasks that require simple actions such as checking a gauge, completion times will typically be very short.

The output of PAWS is the utilization of each agent by phase, where utilization is calculated as shown in Equation 1 for each agent *i*, where the numerator represents the time agent *i* spent completing tasks for that phase. For this modeling effort, time spent monitoring the flight gauges was counted as utilization, but any time pilots spend scanning outside of the cockpit for traffic was not included, since this variable can change dramatically under various weather and lighting conditions.

$$Utilization_i = \frac{(Time\ Spent\ Completing\ Tasks)_i}{Total\ Time\ in\ Phase} \qquad (1)$$

Based upon this structure, PAWS can flexibly include tasks of a variety of types and task allocations, under many different agent architectures and role allocations. In its current instantiation, PAWS makes several assumptions about the completion of tasks:

- Allocations of tasks are known a priori and are static, though these allocations may vary across phases. This means that currently there is no dynamic reallocation of tasks during the flight.
- Tasks are treated as equal priority. Agents do not service certain tasks in the queue before others. For example, tasks such as communications with ATC are treated as equal importance to raising or lowering the landing gear.
- Tasks cannot be dropped from agent queues, i.e., there are no balking or reneging queuing behaviors. Specifically, this means tasks are not dropped during periods of high workload or when the nominal phase duration is met.

PAWS takes a task-based approach to modeling workload, and focuses on the aggregate effort required to complete subgoals in the cockpit rather than model a specific individual physical motion or action. Other researchers have modeled individual actions as an approach to evaluate function allocation in the cockpit (Pritchett, Kim, & Feigh, 2014a; Pritchett, Kim, & Feigh, 2014b). However, such approaches, while important to assess the importance of very low-level actions, fail to take variability into account, either from a pilot or environment perspective. In contrast, PAWS directly represents the variability in work environment factors, including numbers and types of tasks, task arrival times, and human variability in performance of tasks.

DATA COLLECTION

An initial data collection was conducted on the PAWS simulation to demonstrate its capabilities. The simulation was initialized with 225 tasks that represent the full set of generic tasks for a point-to-point mission from takeoff to landing. Figure 1 shows an example portion of the task flow. In this figure, the sequence of tasks for both of the standard roles in the current multi-crew cockpit environment (the pilot flying, or PF, and the pilot monitoring, or PM) can be seen vertically. Arrows can be seen coming to and from these tasks on the right-hand side of the boxes, and these demonstrate the interdependencies in tasks between the multiple crew members in the current cockpit environment.

Table 1 shows a sample of tasks from this model used for the PAWS simulation along with their associated distributions, drawn from the before take-off check. Nominal minimum phase times for a 90-minute mission were assumed as shown in Table 2, and were validated during SME interviews.

Three scenarios were used to demonstrate the capabilities of PAWS in terms of investigating the impact of different role allocations between a pilot and advanced autonomy. The first scenario represents the case where no advanced autonomy has been implemented under normal flight operations, and the single pilot must handle all tasks. This scenario is much like what happened when a Delta pilot was locked out of the cockpit in 2015, requiring the co-pilot to fly the plane alone (Associated Press, 2015). The second scenario represents the same circumstances under off-nominal conditions where an additional emergency task arises, such as an engine fire. The third scenario demonstrates the ability of PAWS to simulate advanced autonomy that could possible handle certain functions of flight operations at the discretion of the pilot.

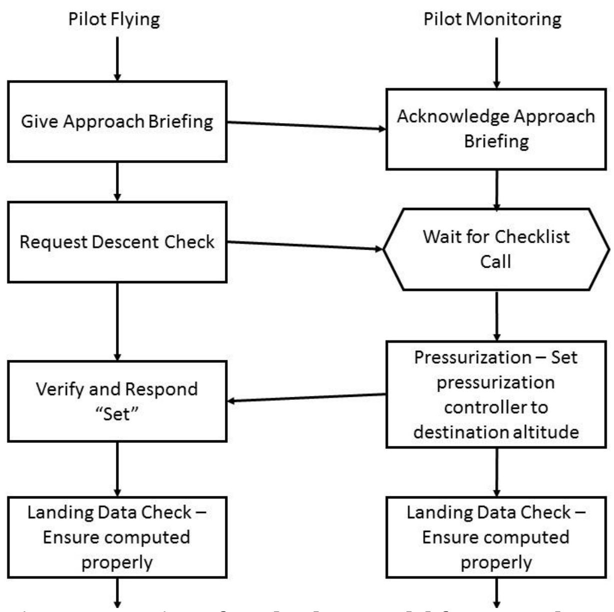

Figure 1. Section of Task-Flow Model for PF and PM Roles

Table 1. Sample Tasks from Generic Task Model

Task Name	**Completion Time Distribution (seconds)**
Request before take-off check	Lognormal(0.6931, 0.1)
Cabin check (communicate w/ crew)	Lognormal(2.9957, 0.5)
Respond to cabin check	Lognormal(0.6931, 0.1)
Request flight controls check	Lognormal(0.6931, 0.1)
Verify flight controls	Lognormal(2.0794, 0.1)
Respond for verification	Lognormal(0.6931, 0.1)
Verify flight controls check completed	Lognormal(2.0794, 0.1)

Table 2. Phase Times for 90-Minute Nominal Mission

Phase ID	**Phase**	**# of Tasks**	**Estimated Phase Time (s)**
1	Before take-off check	25	180
2	Take-off procedure	30	240
3	Climb	18	900
4	Cruise	38	2,580
5	Descent/approach check	24	600
6	Initial approach	3	600
7	Final approach	49	300
	Total	**225**	**5,400**

- Scenario 1 – Nominal mission with no advanced automation. Single agent (human pilot).
- Scenario 2 – Off-nominal mission with no advanced automation. Single agent (human pilot), with an additional emergent task introduced during the initial approach (i.e., investigating a potential electrical problem) with completion time in seconds distributed as shown in Equation 2.

$$\ln N(4.0, 0.5) \tag{2}$$

- Scenario 3 – Off-nominal mission with advanced autonomy. Two agents (one human pilot that supervises and monitors the cockpit and one autonomous agent that handles all primary flight tasks), with an additional emergent task introduced during the initial approach with completion time distributed as shown in Equation 2.

RESULTS

Thirty runs of the simulation for each scenario were conducted to observe the impact of random variation of task completion times on pilot utilization. Figure 2 shows a box plot of the utilization results of the simulation runs for Scenario 1 across all flight phases. The difference between phases are statistically significant by ANOVA omnibus test ($F(6,203) = 317.01$, $p < 0.0001$). All pairwise differences between phases were significant by Tukey's comparison test except for Phase 1-2 ($p < 0.05$). Included in **Figure 1** is the notional upper limit of 70 percent and lower limit of 30 percent beyond which decrements in performance may be observed (Cummings & Guerlain, 2007; Schmidt, 1978; Donmez, Nehme, & Cummings, 2010).

Figure 3 shows the results across the three scenarios with varying degrees of autonomy for the initial approach phase (Phase 6), which included an additional task representing an emergency event for Scenarios 2 and 3. The differences between the three scenarios are significant by ANOVA omnibus test ($F(2,87) = 60.32$, $p < 0.0001$). Scenario 3 was significantly different than Scenarios 1 and 2 by Tukey's comparison test ($p < 0.0001$). These results demonstrate the ability of the PAWS model to explore the impact of different task allocation strategies on pilot utilization, in this case for a single phase of flight. It also demonstrates what functionalities must be able to be assigned to an autonomous system in order to achieve the workload relief for the pilot.

DISCUSSION

These results raise several important considerations for SPO. As can be seen in Figure 2, there is significant variation both within and between phases of flight in the model. This variation is primarily a function of the distributions for completion times for each phase. For example, time spent flying the initial approach (Phase 6) has the largest variation due to the differences in typical approach patterns and ATC-dictated modifications to the approach.

Figure 2 also demonstrates that there are several phases of nominal point-to-point flight (before take-off check, initial approach, final approach) for which utilization for human pilots can exceed 70 percent. It is well established within the human factors literature that there are significant performance and safety decrements at high workload levels (Oron-Gilad, Szalma, Stafford, & Hancock, 2008; Cummings & Guerlain, 2007). Thus, it is these areas that need to be targeted in terms of reducing pilot workload.

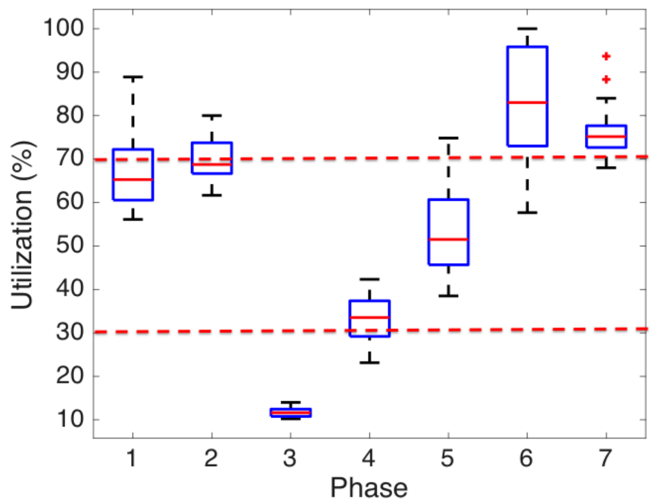

Figure 2. Boxplot of Pilot Utilization for Nominal Point-To-Point Mission (Scenario 1). Dashed lines indicate notional desired upper and lower limits of workload

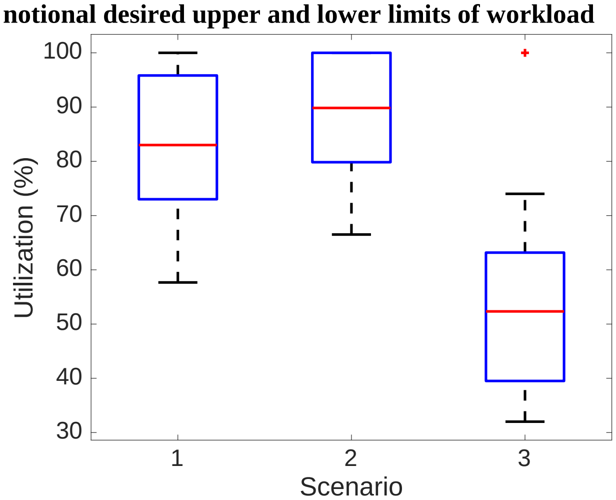

Figure 3. Comparison of Pilot Utilization for Initial Approach for 3 Scenarios

The results in Figure 2 also indicate the potential for unacceptably low workload during some phases. It has previously been reported that pilots spend only a few minutes of each flight actually performing flight tasks in modern commercial airplanes (Cummings, Stimpson, & Clamann, 2016), and low workload has also been reported to result in performance decrements (Cummings & Gao, 2016). The Climb phase is of particular note in Figure 2, with a consistent utilization less than 20 percent. This result indicates that while advanced autonomy may be required for certain portions of a flight, allocating advanced autonomy should depend upon the phase or other circumstances so as to avoid reducing the workload of an already underworked pilot even further.

Understanding that allocating primary flight control to advanced autonomy will be both phase- and environment-dependent, Figures 2 and 3 demonstrate that the initial approach to land constitutes the highest workload and the highest variability in workload. This phase of flight would likely benefit the most from increased autonomy. During the nominal initial approach (Scenario 1), utilization is above the 70 percent threshold, significantly increasing the chances for reduced pilot performance. When an additional emergency is introduced (Scenario 2), pilot utilization further increases, exacerbating the potential for overloading in terms of mental workload. As can be seen in Figure 3, the introduction of advanced automation to perform the primary flight tasks dramatically reduces the utilization of the human pilot to a more acceptable level (Scenario 3). This illustrates not only the importance of advanced autonomy in achieving safe SPO, particularly under contingency situations, but also the potential utility of the PAWS framework.

PAWS was designed to allow for the exploration of potential role allocations under dynamic flight conditions, so that engineers can better understand how increasing autonomy could affect pilot workload. In addition, since PAWS inherently incorporates elements of task analysis, such an approach also indicates to engineers specifically what capabilities an autonomous system would need to have in order to execute those tasks. By understanding which tasks have high variability and high associated workload (as well as low), engineers would have clear specifications for system capabilities (i.e., an autonomous system must be able to detect and handle an emergency such as an electrical problem), as well as the temporal constraints of such interactions. In addition, any added interactions with such autonomous systems would also have to be modeled in order to ensure a human pilot's workload did not actually increase, which can sometimes be an issue when automation is increased (Bainbridge, 1983).

FUTURE WORK

Historically, the introduction of new equipment such as advanced navigation systems has reduced the number of actions pilots use to execute procedures. This is why aircraft today only need two pilots as compared to the 1940s-era Boeing Stratocruiser, which needed five people in the cockpit. Seen through this lens, advanced autonomy is not a revolutionary but rather evolutionary technological advance that simply continues a long-standing trend.

PAWS is a simulation designed to help engineers understand the likely impacts of such advances on an SPO system before system design and implementation, in order to avoid costly mistakes or inefficient designs. The principal goal of any transition to SPO is to maintain levels of safety equivalent to or better than current operations. To this end, these initial PAWS results underscore the need for advanced autonomy to assist the pilot in SPO under certain conditions, but also that the implementation and task assignments between the pilot and autonomy must be carefully examined.

In addition, these results demonstrate many of the useful capabilities of the PAWS simulation. Further investigations will include adding autonomous agents with varying levels of allocation, and the inclusion of stochastic occurrences of contingencies. Specifically, the following additional scenarios will be examined:

- More detailed emergency conditions of increasing complexity under various phases of flight
- Pilot incapacitation that requires the autonomy system to identify and land at an alternate landing site

The findings from this work help to define a roadmap for the levels of safety and applicable missions to which advanced autonomy could be applied. Such understanding is critical for identifying critical design elements, as well as informing policy for regulatory agencies such as the Federal Aviation Administration (FAA). Moreover, while PAWS currently focuses on examining the trade space between pilots and onboard autonomy, it can also be used to represent an operator in any transportation system where advanced autonomy may share the operational space with humans, such as rail and driverless cars.

ACKNOWLEDGEMENTS

This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA). The views, opinions and/or findings expressed are those of the author(s) and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

REFERENCES

Associated Press. (2015). *Delta pilot locked out of cockpit mid-flight*. Retrieved from Fox News: http://www.foxnews.com/travel/2015/01/30/delta-pilot-locked-out-cockpit-mid-flight/

Bainbridge, L. (1983). Ironies of automation. *Automatica, 19*(6), 775-779.

Comerford, D., Brandt, S., Lachter, J., Wu, S., Mogford, R., Battiste, V., & Johnson, W. (2013). NASA's Single-pilot operations technical interchange meeting: Proceedings and findings. Moffett Field, CA: NASA Ames Research Center.

Croft, J. (2015, Feb 16). The Coming U.S. Pilot Shortage is Real. *Aviation Week & Space Technology*.

Cummings, M., & Gao, F. (2016). Boredom in the Workplace: A New Look at an Old Problem. *Human Factors, 58*(2), 279-300.

Cummings, M., & Guerlain, S. (2007). Developing Operator Capacity Estimates for Supervisory Control of Autonomous Vehicles. *Human Factors, 49*, 1-15.

Cummings, M., Stimpson, A., & Clamann, M. (2016). Commercial Pilot Views on Futuristic Single Pilot Operations. *Proceedings of the AIAA Aerospace Conference.* San Diego, CA: AIAA.

DARPA (2014). *ALIAS seeks to provide portable, flexible advanced autopilot capabilities.* From http://www.darpa.mil/newsevents/releases/2014/04/18.aspx

Donmez, B. D., Nehme, C., & Cummings, M. L. (2010). Modeling Workload Impact in Multiple Unmanned Vehicle Supervisory Control. *IEEE Systems, Man, and Cybernetics, Part A Systems and Humans, 40*(6), 1180-1190.

Fitts, P. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*, 381-391.

Gao, F., & Cummings, M. (2012). Using Discrete Event Simulation to Model Multi-Robot Multi-Operator Teamwork. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56*, pp. 2093-2097. Boston, MA.

Harris, D. (2007). A human-centered design agenda for the development of single crew operated commercial aircraft. *Journal of Aircraft Engineering and Aerospace Technology, 79*(5), 518-526.

Jun, J. B., Jacobson, S., & Swisher, J. (1999). Application of discrete-event simulation in health care clinics: A survey. *Journal of the operational research society*, 109-123.

Laughery, K. R., Archer, S., & Corker, K. (2001). Modeling Human Performance in Complex Systems. In G. Salvendy (Ed.), *Handbook of Industrial Engineering: Technology and Operations Management, Third Edition.* Hoboken, NJ,: John Wiley & Sons, Inc.

McCartney, S. (2010). *Imagining a day when airliners are flown solo.* From http://online.wsj.com/article/SB10001424052702304011604575564212371734270.html

Oron-Gilad, T., Szalma, J. L., Stafford, S. C., & Hancock, P. A. (2008). The workload and performance relationship in the real world: A study of police officers in a field shooting exercise. *International Journal of Occupational Safety and Ergonomics, 14*(2), 119-131.

Parasuraman, R., Mouloua, M., & Hilburn, B. (1999). Adaptive aiding and adaptive task allocation enhance human-machine interaction. *Automation technology & human performance: Current research and trends*, 119-123.

Pritchett, A. R., Kim, S. Y., & Feigh, K. M. (2014). Measuring Human-Automation Function Allocation. *Journal of Cognitive Engineering and Decision Making, 8*(1), 52-77.

Pritchett, A. R., Kim, S. Y., & Feigh, K. M. (2014). Modeling Human-Automation Function Allocation. *Journal of Cognitive Engineering & Decision Making, 8*(1), 33-51.

Rouse, W. (1983). *Systems Engineering Models of Human-Machine Interaction.* New York: North Holland.

Schmidt, D. (1978). A Queuing Analysis of the Air Traffic Controller's Workload. *IEEE Transactions on Systems, Man, and Cybernetics*, 492-498.

Schutte, P., Goodrich, K., Cox, D., Jackson, B., Palmer, M., Pope, A., . . . Kinney, J. (2007). *The naturalistic flight deck system: An integrated system concept for improved single-pilot operations.* NASA.

Sheridan, T. (2013). Human response is lognormal; plan on waiting if you want reliability. *Ergonomics in design: The quarterly of human factors, 21*(1), 4-6.

Wolter, C., & Gore, B. (2015). *A Validated Task Analysis of the Single Pilot Operations Concept.* Moffett Field, CA: Ames Research Center.

Xiling, L., Qishan, Z., & Wei, L. (2002). A discrete-event system method for solving the single airport ground-holding problem in air traffic control. *2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering. 3*, pp. 1745-1748. IEEE.