5.9 Modeling Pilot State in Next Generation Aircraft Alert Systems

Modeling Pilot State in Next Generation Aircraft Alert Systems

Alan S. Carlin, Amy L. Alexander, Nathan Schurr
Aptima, Inc.
acarlin@cs.umass.edu, aalexander@aptima.com, nschurr@aptima.com

Abstract. The Next Generation Air Transportation System will introduce new, advanced sensor technologies into the cockpit that must convey a large number of potentially complex alerts. Our work focuses on the challenges associated with prioritizing aircraft sensor alerts in a quick and efficient manner, essentially determining when and how to alert the pilot. This "alert decision" becomes very difficult in NextGen due to the following challenges: 1) the increasing number of potential hazards, 2) the uncertainty associated with the state of potential hazards as well as pilot state, and 3) the limited time to make safety-critical decisions. In this paper, we focus on pilot state and present a model for anticipating duration and quality of pilot behavior, for use in a larger system which issues aircraft alerts. We estimate pilot workload, which we model as being dependent on factors including mental effort, task demands, and task performance. We perform a mathematically rigorous analysis of the model and resulting alerting plans. We simulate the model in software and present simulated results with respect to manipulation of the pilot measures.

1.0 INTRODUCTION

The introduction of Next Generation Air Transportation System (NextGen) technologies into the cockpit is expected to dramatically increase the responsibilities of the pilot (JPDO, 2007). In particular, additional aircraft alerting systems will be introduced, and the pilot will need to adapt to the increase in both the number and types of possible hazard alerts. NextGen will also introduce additional automation technologies into the cockpit, capable of addressing alerts with minimal assistance from the human pilot. However, interfacing these technologies with both the human pilot as well as the large number of possible hazard alerts introduces a set of research challenges, including how to prioritize the alerts, how to plan the interaction between human and automation to address the prioritized hazards, and how to adjust the plan according to the state and capabilities of the pilot.

In order to address these challenges, Aptima, Inc., in cooperation with SAIC and under the supervision of NASA, is developing a NextGen aircraft system called ALARMS (ALerting And Reasoning Management System). The system has four parts: Bayesian reasoning to determine type and priority of existing hazards, a Time Dependent Markov Decision Process (TMDP)-based planner to address the hazards in a timely fashion, a human performance estimator to inform the planner as to the state and capabilities of the pilot, and an interface to inform the pilot of alerts in the best possible manner.

In this paper, we concern ourselves with the third item, how to estimate pilot state and capabilities in order to inform a plan for the human and automation to cooperate. Defining a plan to cooperate has been the subject of empirical research (Galster, 2003; Galster & Parasuraman, 2003; Parasuraman & Riley; Parasuraman, Sheridan, & Wickens, 2000). One approach is to describe a level of automation in a continuum between fully automated hazard response, to fully manual hazard response (Wickens, Mavor, Parasuraman, & McGee, 1998; Sheridan & Verplank, 1978). An extended method, proposed by Parasuraman et al. models human information processing in four stages: Sensory Processing, Perception/Working Memory, Decision Making, and Response Selection (Parasuraman, Sheridan, & Wickens, 2000). Differing circumstances may call for differing stages of automation.

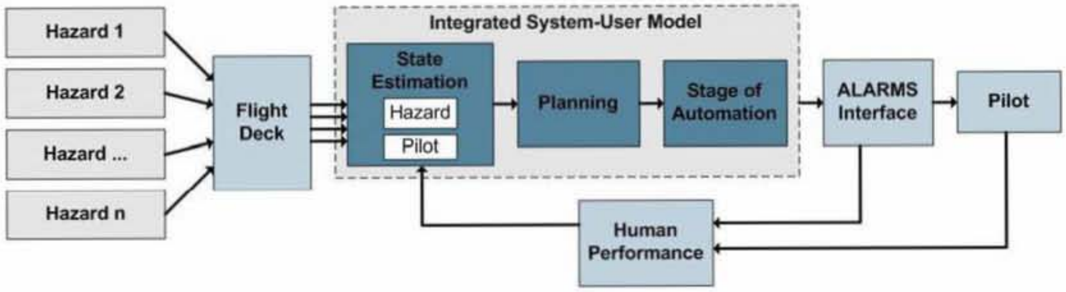


Figure 1: ALARMS System Diagram

But which stage of automation is appropriate may depend on several variables, including the characteristics of the hazards, as well as "pilot state," the ability of the pilot to perform under a given stage of automation. In this paper we introduce a model for estimating pilot state on the aircraft, for the purposes of informing the hazard alerting system. The model leverages existing literature on pilot workload as well as pilot performance. The model in this paper uses three stages of automation, each of which corresponds to a flight deck display. In Stage 1 of automation, increasing workload will greatly decrease quality and increase duration for high workload conditions as compared to low workload conditions. In Stage 2, increasing workload will decrease quality and increase duration. In Stage 3, the effects will be negligible.

The outline for the rest of the paper follows: First, we introduce the ALARMS system architecture and briefly outline its components, including a hazard state estimation module and a planning module for stages of automation. Next, we outline the pilot state estimation module, which estimates pilot workload. We show how the pilot state can be used as input for the ALARMS planning module. Finally, we show modeled results for how changes of pilot state will change temporal plans for stages of automation, and conclude with a summary and a discussion of future steps.

2.0 BODY: ALARMS SYSTEM ARCHITECTURE

The ALARMS system architecture is shown in Figure 1. Proceeding from left to right, multiple hazards exist in the environment and result in alerts on the flight deck. The hazards themselves, and the sensor alerting systems, are external to the ALARMS system. The alerts are issued to the ALARMS Integrated System User Model. The system alerts are treated as *evidence,* and from this evidence the ALARMS state estimation module estimates the actual hazards. A more detailed description of this estimate can be found in a companion paper to this work (9). To summarize, a Dynamic Bayesian Network (DBN) is used. DBN's have been found in similar systems, notably in medical diagnosis (Shwe & Cooper, 1991). Our use of a DBN in the ALARMS system is analogous, if the system alerts are treated as "symptoms" to estimate the "disease" of the actual hazard.

The hazard state is combined with the estimated state of the pilot (which we will return to in a moment), to form a complete state estimate. This pilot and hazard estimate is fed into a Planning module. The Planning module recommends the stage of automation for the hazards, which is fed into the ALARMS interface for display. The result is displayed to the pilot.

In this work, we focus on the Human Performance module in the diagram. This module estimates the status of the pilot,

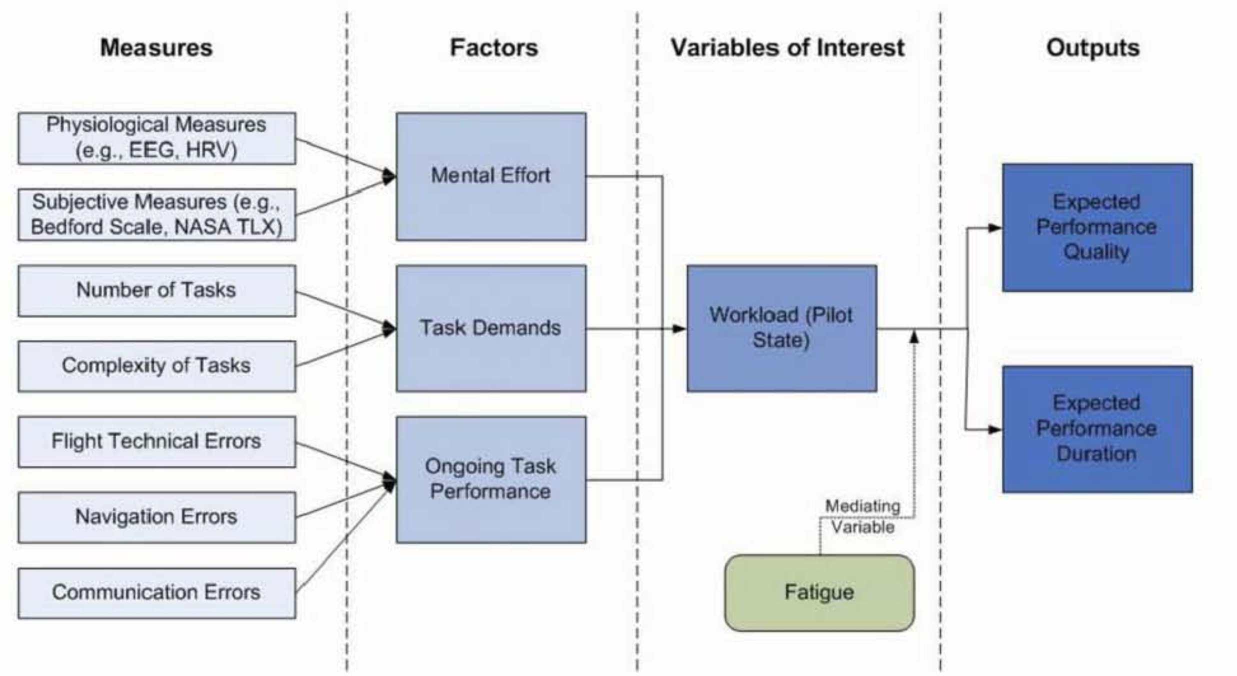


Figure 2: Human Performance Module

which in turn is used to estimate the performance of the pilot at each stage of automation. The planner will then select the appropriate stage of automation that maximizes the effectiveness of the combined pilot and automation.

2.1 Human Performance Module

We model human performance as shown in Figure 2. The Output from the module is an estimate of expected pilot performance, in terms of duration and quality of pilot handling of hazards. The Variable of Interest, **Workload**, is the key parameter representative of pilot state that changes over time and directly impacts performance. The Mediating Variable, **Fatigue**, influences the relationship. Other variables of interest (e.g., situation awareness) or mediating variables may be considered in future work.

In environments with task demands, workload affects the mental resources that a pilot can access to address the demands (Wickens & Hollands, 2000). Specifically, the effect can be modeled through a performance resource function, or PRF (Norman & Bobrow, 1975). When cognitive resources are unavailable or unused for a task, performance will be diminished. As more resources are dedicated, performance will improve, until the task becomes limited by data and not resources. When multiple tasks must be accomplished, such as is the case when a pilot must supervise multiple systems in the cockpit, resource limitation becomes an issue (Kantowitz & Casper, 1988). The workload of the pilot will define the availability of a pilot's resources to handle alerts.

It is possible to assess workload as an index, and several criteria have been specified to compute the index (Wickens & Hollands, 2000; O'Donnell & Eggemeier, 1986). Among these criteria: a satisfactory workload index is *sensitive* to changes in task demands, *diagnoses* the cause of workload variation, is *selective* in that factors that do not affect workload are not included in the index, is *unobtrusive* in that the computation of the index does not affect workload itself, and is *reliable*.

For ALARMS, we identify three factors that predict workload: mental effort, task demands, and ongoing task performance. We also identify relevant measures of these factors from the literature.

2.1.1 Mental Effort

We follow the literature by specifying mental effort as a contributing factor to workload. High levels of performance can be achieved under conditions of normal mental effort while extremely high mental effort situations tend to result in decreased performance. Measures of mental effort include both subjective and physiological measures (Veltman, 2001).

Subjective information in our model includes potential measures such as the NASA TLX scale (Hart & Staveland, 1988), which allows the operator to specify mental demand, physical demand, temporal demand, performance, effort, and frustration level. The Bedford Workload scale (Roscoe, 1984), on the other hand, is a decision tree, and the leaves of the tree provide a workload score on a single dimension.

Physiological information can also be obtained. Examples of potential measures include electroencephalography (EEG) or heart rate variability (HRV). It has been shown that heart rate can differentiate between phases of flight (which require different levels of mental effort) for pilots and co-pilots (Bonner & Wilson, 2002), even when subjective measurements do not.

2.1.2 Task Demands

In the prior subsection, mental effort is described as being necessary to accomplish tasks. The level of effort demanded will depend on the task. Simple tasks will require smaller amounts of resources, while complex tasks will require a higher degree of mental effort. Measures of Task Demands include both the complexity of tasks and the number of tasks.

Task complexity can affect workload; specifically, complex tasks will result in a higher workload. For example, the landing phase of flight produces higher workload than the enroute phase (Bonner & Wilson, 2002). As a second example, more automated tasks consume fewer resources than less automated ones (Schneider & Fisk, 1982).

Number of tasks affects workload as well, in two ways. First, the presence of additional tasks adds to workload. Second, there is a cost to switching among tasks (Rogers & Monsell, 1995). Thus, the contribution of tasks to workload exceeds the sum of the tasks complexities.

2.1.3 Ongoing Task Performance

Workload contributes to the model insofar as it is predictive of pilot performance. Thus, a well-accepted manner of estimating workload is to examine performance directly. Potential measurements include **Flight Technical Errors**, **Navigation Errors**, and **Communication Errors**. These errors can be measured by the ALARMS system at run-time.

2.2 Interface to ALARMS Planner

The goal of estimating pilot Workload is to predict the quality and duration of pilot actions so that a joint pilot/automation plan can be formed to address hazards. In this section, we describe the details of the interface to the planner. We begin by summarizing the planner itself, as introduced in (Carlin, Marecki, & Schurr, 2010) and adjusted in this paper to account for pilot state.

2.2.1 ALARMS TMDP Planner

We model the ability of the pilot and system to address hazards with a Time-Dependent Markov Decision Process (TMDP) (Boyan & Littman, 2000). A TMDP is a tuple $<S,A,P,D,R>$, where S is a set of states, A represents a set of actions, P is a transition matrix, D is a set of probability density functions, and R is a reward. Assume a finite set S of discrete states and a finite set A of actions. When the state s in S, and action action a in A is executed, the process transitions with probability $P(s,a,s')$ to state s' in S. The transition consumes t units of time with probability $d(s,a,s',t)$, where $d(s,a,s',t)$ is a probability density function over t for a given s,a, and s'. Similarly, the reward $R(s,a,s')$ depends on s, a, and s'. Reward occurs when an action terminates. A deterministic TMDP policy is a mapping $S \times [0,\Delta] \rightarrow A$ where Δ (the deadline) is the earliest point in time after which all rewards are zero.

Let Ψ be a set of alert levels (e.g. "Nominal" (N), "Advisory" (A), "Caution" (C), "Warning" (W), or "Directive" (D)), Φ be an ordered set of hazards, and Ω be a set of autonomy stages. A TMDP problem in ALARMS is defined as follows:

- **States** A state s in S is a mapping from the hazards to their aleret levels. For example, given three hazards, state $s = <C,N,W>$ defines that the first hazard is at Caution level, the second hazard is at Nominal level, and the third hazard is at a Warning level.
- **Actions** The actions of the ALARMS system represent the different ways in which the system displays the information about the hazards on the pilot's GUI. Each component of an action represents a stage of autonomy. For example, given three hazards, action $a = <1,3,2>$ will present three hazards at stages 1,3, and 2 of autonomy. It is possible to represent different hazards at the same stage of autonomy (e.g. $<2,2,2>$).
- **Transitions:** ALARMS assumes that all hazards will eventually be addressed (their alert levels will return to N as a result of human or autonomy actions. An exception is when the action is not to address the hazard at any stage of automation (e.g. $<0>$), in which case the state remains the same for that hazard.
- **Durations:** ALARMS models action duration distributions by assuming that actions at differing stages of automation take different durations. The specific durations of actions will be affected by pilot state, as we will specify in the next section.
- **Reward:** Reward is achieved for addressing the hazard and transitioning back to a nominal state. Each hazard can have a different reward associated with it. Reward will also depend on the pilot state, as we will see in the next section.

2.2.2 Pilot State in ALARMS TMDP

As shown in Figure 2, Workload affects the duration and quality of pilot actions in the ALARMS model. This is accomplished by performing a two step process. First, a workload score is computed from measurements of factors. This is accomplished through a linear weighting of the factors:

$$Workload = \alpha * ME + \beta * TD + \gamma * TP$$

where ME represents Mental Effort, TD represents Task Demands, and TP represents Task Performance. α, β, and γ represent linear weights that allow the prioritization of the factors to be varied.

In the second step, the workload score is used to modify the Duration and Reward function of the ALARMS TMDP. We use the Workload estimate to feed information into the Integrated User Module about the expected capabilities of the pilot, specifically the expected performance quality and the expected duration of pilot actions. The effect of Workload varies according to the stage of automation. In Stage 1 of automation, increasing workload in our model will greatly decrease quality and increase duration for high workload conditions as compared to low workload conditions. In Stage 2, increasing workload will decrease quality and increase duration. In Stage 3, we make the effects negligible.

The specific quantities attached to these terms "greatly decrease", etc, are

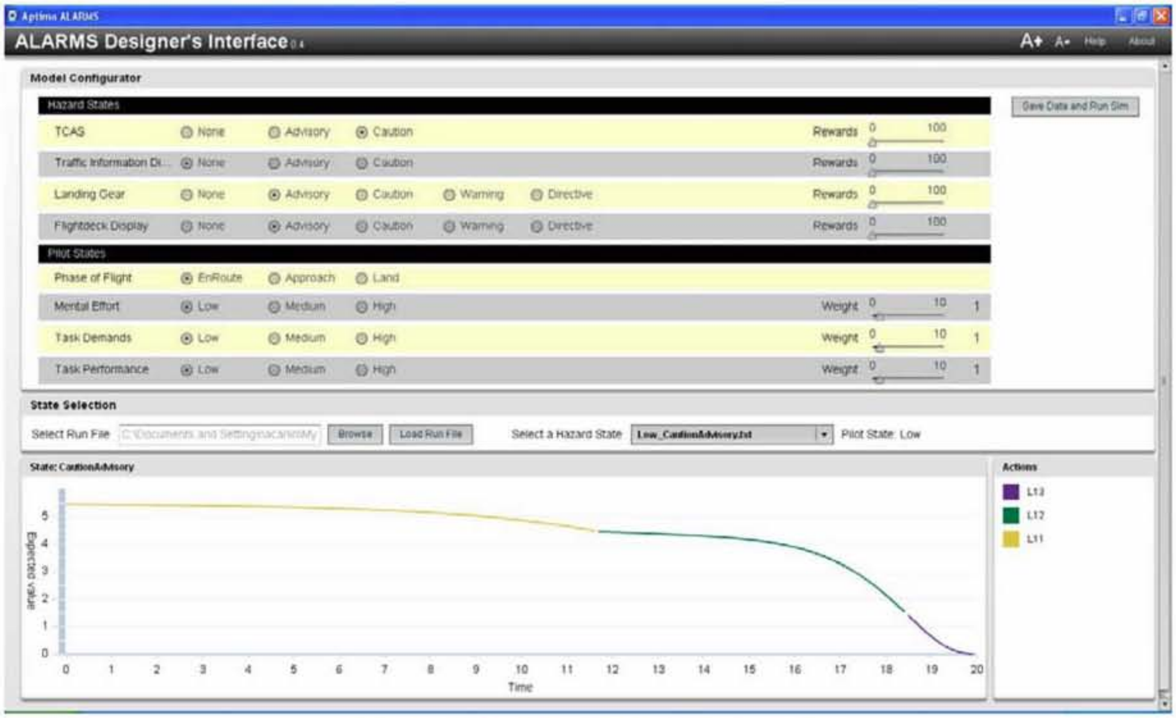


Figure 3: ALARMS System Designer Interface

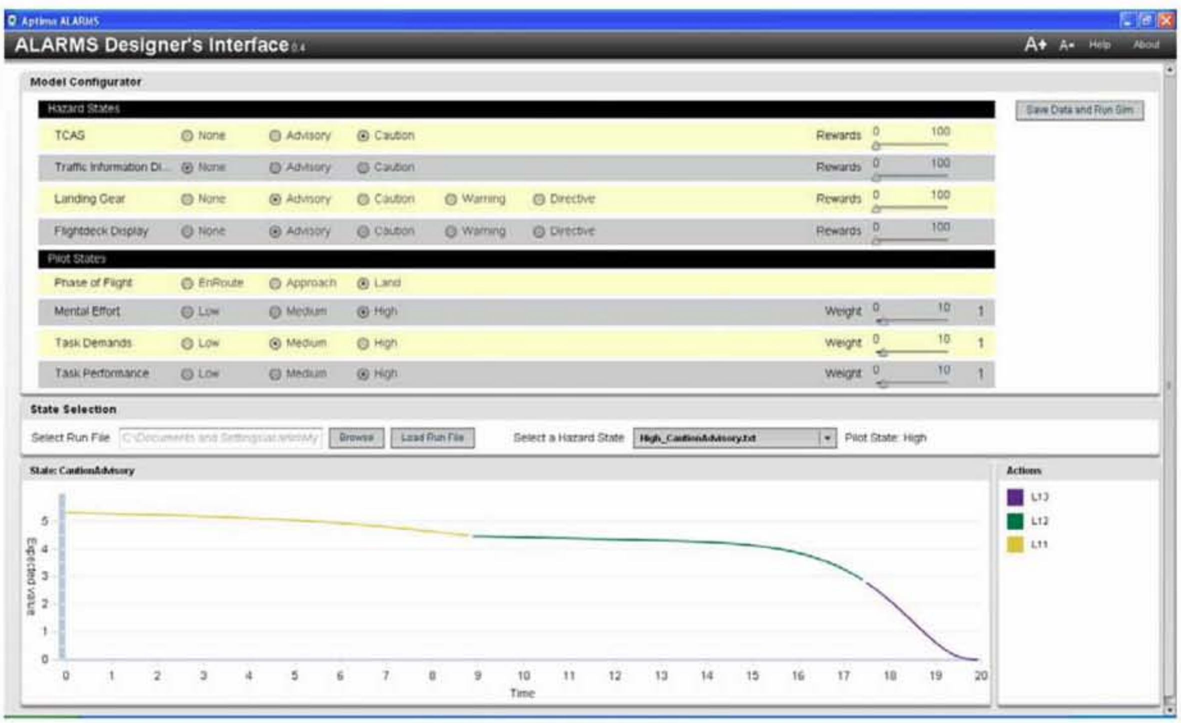


Figure 4: High Workload Plan

parameters in our model. At present, we set quality and duration to halve and double, respectively, in Stage 1, when workload is changed from Low to High. Similarly we set quality and duration to decrease and increase 25% in Stage 2, and to decrease and increase 5% in Stage 3. Medium workload is currently simulated by interpolating between the high and low workload conditions.

3.0 DISCUSSION/EVALUATION

In order to evaluate the effect of the factors on workload and on existing plans, an ALARMS System Designer Interface was developed. The interface is shown in Figure 3. On the top, multiple system alerts can be specified at various levels of alert. Directives are the highest priority of alert, followed by Warnings, Cautions, and Advisories. The entries "TCAS" and "Traffic information display" are indicative of a loss of separation hazard, and the entries "Landing Gear" and "Flight deck display" are indicative of a system failure hazard encountered while landing. Thus, we see in the figure that there is a caution-level alert for loss of separation, and a lower-priority advisory for a system failure hazard. Below the hazards, the factors affecting pilot state are specified, including Mental Effort, Task Demands, and Task Performance. Each factor is given a weight (corresponding to α, β, γ), in this case the weights are all 1.0. The figure shows that factors are all selected as "Low," and thus the pilot is under low workload conditions.

Below, we see a time-dependent plan for addressing the hazards, as computed by ALARMS. The x-axis is in time units. Without loss of generality, it is assumed that the tasks have a "deadline" at the 20 unit mark on the x-axis, and the plan works backwards from that mark. The y-axis shows the utility of the plan (on a relative scale to the ALARMS planning problem). As expected, utility decreases as the deadline approaches. The figure shows that ALARMS produces a 3-part plan for addressing the hazards under these conditions. Each part of the plan consists of the letter "L" followed by a stage of automation for each hazard, thus "L11" denotes that both hazards are handled at a stage of automation of 1, "L12" indicates that the loss of separation hazard is handled at stage 1 and the system failure hazard is handled at stage 2. The figure shows that when there are more than 8 time units remaining, the hazards are both handled at Stage 1 of automation. As the deadline approaches, the recommended stage of automation transitions to "L12," that is, the lower priority hazard is handled at a higher stage of automation. Within 1.5 time units of the deadline, the stage of automation transitions to "L13" the lower priority hazard is handled at Stage 3 of automation.

Figure 4 shows a second plot. Here, the Phase of Flight has been changed to Land, Mental Effort and Task performance are labeled as indicating "High" workload conditions, and Task Demands are "Medium." This is a higher workload condition than the first example. As a result, the ALARMS planner is informed that low stages of automation will be less effective. The resulting plan at the bottom of the figure shows that higher stages of automation are selected at earlier points in time.

4.0 CONCLUSION

In this paper, we introduced a model designed to predict pilot performance in the cockpit, proposed to be implemented as a component of NextGen alerting systems. The larger ALARMS system design consists of Bayesian reasoning to determine type and priority of existing hazards, a Time Dependent Markov Decision Process-based planner to address the hazards in a timely fashion, a human performance estimator to inform the planner as to the state and capabilities of the pilot, and an interface to inform the pilot of alerts in the best possible manner. In this paper we focused on a model to contribute to the human performance estimator.

Key components of the model are that it estimates workload, it predicts the duration and quality of pilot performance, and it can be used to recommend what information will be displayed to the pilot, and what information processing stage will be supported.

Future work consists of several directions. First, we will focus on the real-time nature of the measures, and how such measurements can be integrated into the cockpit in an unobtrusive manner. Second, we will embellish the model further. For example, the literature on task switching as well as issues related to attention (Yerkes & Dodson, 1908) can be added to the model.

5.0 BIBLIOGRAPHY

Bonner, M. A., & Wilson, G. F. (2002). Heart rate measures of flight test and evaluation. *International Journal of Aviation Psychology, 12,* 63-77.

Boyan, J., & Littman, M. (2000). Exact solutions to time dependent MDPs. *Proceedings of Neural Information Processing Systems*, (pp. 1026-1032).

Carlin, A., Marecki, J., & Schurr, N. (2010). ALARMS: Alerting and Reasoning Management System for Next Generation Aircraft Hazards. *The 26th Conference on Uncertainty in Artificial Intelligence.* Catalina Island, CA: UAI.

Galster, S. (2003). *An examination of complex human-machine system under multiple levels and stages of automation.*.(Technical Report No. AFRL-HE-WP-TR-2003-0149). WPAFB, OH.

Galster, S., & Parasuraman, R. (2003). The application of a qualitative model of human-intereaction with automation in a complex and dynamic combat flight task. *Proceedings of the 12th International Symposium on Aviation Psychology.* Dayton, OH: Wright State University.

Hart, S., & Staveland, L. (1988). Development of NASA-TLS (Task Load Index): Results of empirical and theoretical research. In P. H. (Eds.), *Human mental workload.* Amsterdam: North Holland.

JPDO. (2007). *Concept of Operations for the Next Generation Air Transportation System.* Joint Planning and Development Office.

Kantowitz, B., & Casper, P. A. (1988). Human workload in aviation. In E. L. Wiener & D. C. Nagel (Eds.), *Human factors in aviation* (pp. 157-187). San Diego, CA: Academic Press.

Norman, D., & Bobrow, D. (1975). On data-limited and resource limited processing. *Journal of Cognitive Psychology , 7*, 44-60.

O'Donnell, R., & Eggemeier, F. (1986). Workload assessment methodology. In L. K. K. Boff (Ed.), *Handbook of perception and performance (vol 2).* New York: Wiley.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors , 39*, 230-253.

Parasuraman, R., Sheridan, T., & Wickens, C. (2000). A model of types and levels of human interaction with automation. *IEEE Transactions on Systems, Man and Cybernetics , 30*, 286-297.

Pritchett, A. R., Vándor, B., & Edwards, K. (2002). Testing and implementing cockpit alerting systems. *Reliability Engineering and System Safety , 75*, 193-206.

Rogers, R., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General , 124*, 207-231.

Roscoe, A. H. (1984). Assessing pilot workload in flight. *Conference Proceedings No. 373. Flight Test Techniques.* Paris: AGARD.

Schneider, W., & Fisk, A. (1982). Concurrent automatic and controlled visual sarch: Can processing occur without cost? *Journal of Experimental Psychology: Learning, Memory, and Cognition , 8*, 261-278.

Shwe, M., & Cooper, G. (1991). An empirical analysis of likelihood-weighting simulation on a large, multiply connected medical belief network. *Computers and Biomedical Research*, 5, 453-475.

Sheridan, T., & Verplank, W. (1978). *Human and Computer Control of Undersea Teleoperators.* Cambridge, MA: MIT Man-Machine Systems Laboratory.

Veltman, H. (2001). *Mental Workload: Lessons learned from subjective and physiological measures..*

Wickens, C. D., & Hollands, J. (2000). *Engineering Psychology and Human Performance.* Upper Saddle River, NJ: Prentice Hall, Inc.

Wickens, C., Mavor, A., Parasuraman, R., & McGee, J. (1998). *The Future of Air Traffic Control: Human Operators and Automation.* Washington D.C.: National Academy Press.

Yerkes, R., & Dodson, J. (1908). The relation and strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology* , 459-482.

6.0 ACKNOWLEDGEMENTS

We would like to thank Gilbert Mizrahi for the graphical interface, and Andy Chang for helpful visualizations of stages of automation in the cockpit.

This paper is based upon work supported by the National Aeronautics and Space Administration (NASA) under Contract No. NNL08AA20B issued through the Aviation Safety Program and monitored by Kara Latorella, whom the authors wish to thank. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NASA.

Modeling Pilot State in Next Generation Aircraft Alert Systems

Alan S. Carlin, Amy L. Alexander, Nathan Schurr

APTIMA
HUMAN-CENTERED ENGINEERING

NASA SAIC IBM

NextGen Aircraft Alerting Systems

APTIMA HUMAN-CENTERED ENGINEERING

- Future flight deck systems will require sophisticated **information management**
- Need integrated alerting and notification (IAN) function to:
  - Continuously monitor info from various sources to evaluate hazart potential
  - Consider immediate hazards (current) and situations requiring re-planning or coordination (future)
  - Provide caution/alert/warning automated notifications
  - Provide context-relevant decision support to pilot and other aircraft automation functions
- Interdisciplinary effort

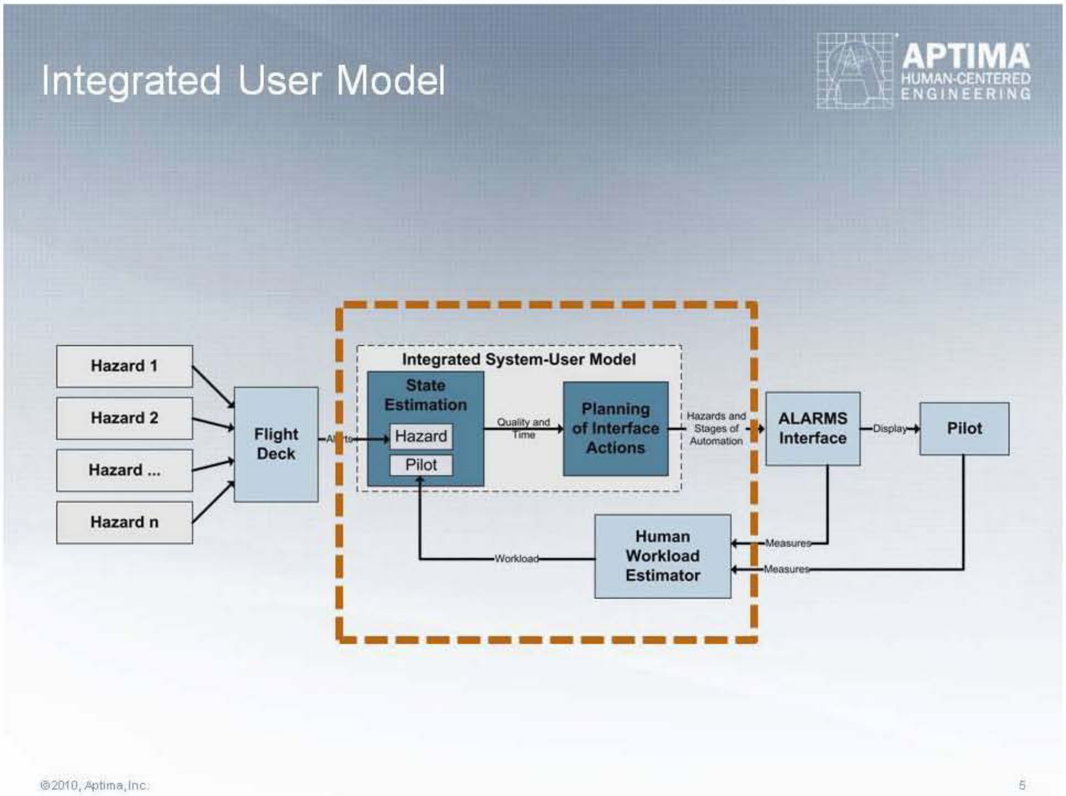
Levels of Automation*

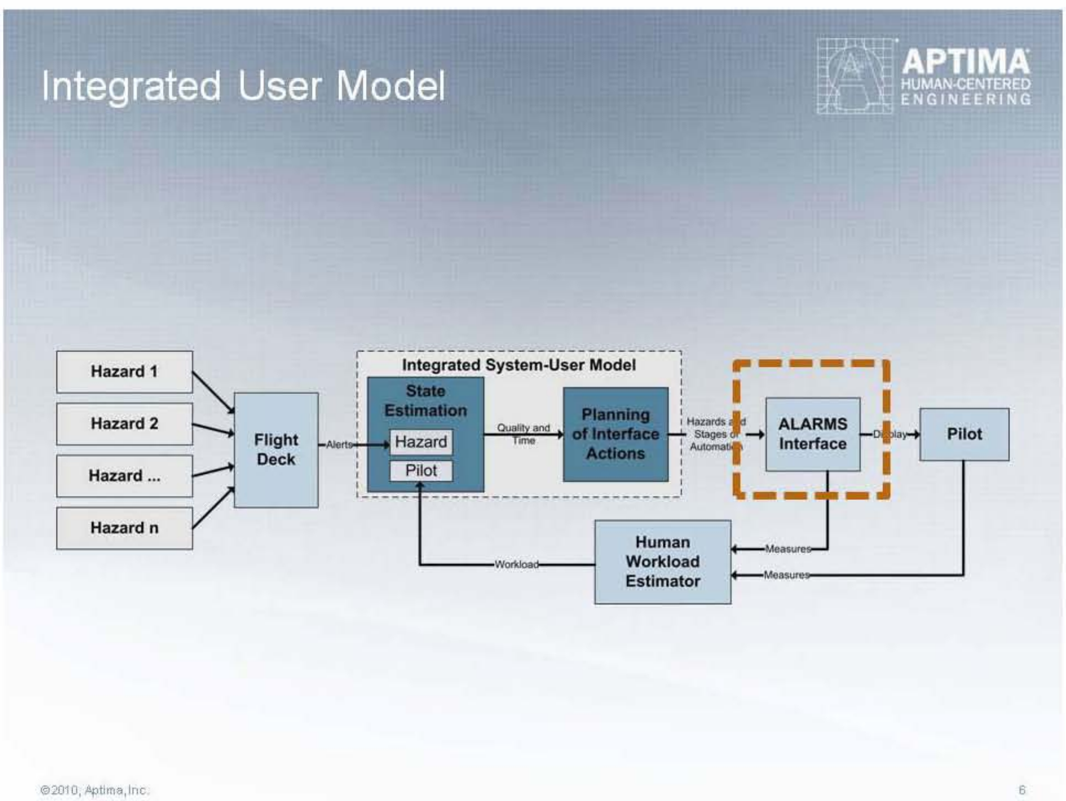
- 10 – Computer Decides everything, acts autonomously
- 9 – Informs human if it decides to
- 8 – Informs human only if asked
- 7 – Executes then informs
- …
- 4 – Suggests one alternative
- 3 – Narrows selection to a few
- 2 – Offers complete set of alternatives
- 1 – No computer assistance

*Wickens 1998, based on Sheridan and Verplank 1978

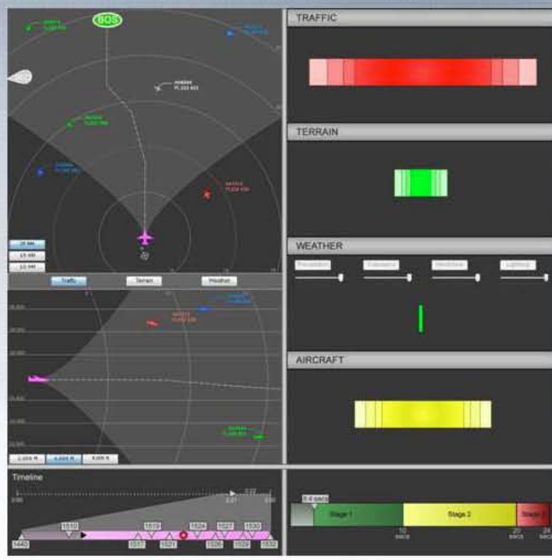
Stages of automation

- Example 4 stage model
  - Sensory Processing
  - Perception/Working Memory
  - Decision Making
  - Response Selection
- Four classes of functions
  - Information acquisition
  - Information analysis
  - Decision and action selection
  - Action implementation


Integrated User Model
APTIMA
HUMAN-CENTERED ENGINEERING
Hazard 1
Hazard 2
Hazard ...
Hazard n
Flight Deck
Integrated System-User Model
State Estimation
Hazard
Pilot
Quality and Time
Planning of Interface Actions
Hazards and Stages of Automation
ALARMS Interface
Display
Pilot
Human Workload Estimator
Workload
Measures
Measures
©2010, Aptima, Inc.
5


Integrated User Model
APTIMA
HUMAN-CENTERED ENGINEERING
Hazard 1
Hazard 2
Hazard ...
Hazard n
Flight Deck
Alerts
Integrated System-User Model
State Estimation
Hazard
Pilot
Quality and Time
Planning of Interface Actions
ALARMS Interface
Pilot
Human Workload Estimator
Workload
Measures
Measures
©2010, Aptima, Inc.
6

Stage 1 Automation



Information Support

Horizontal Display

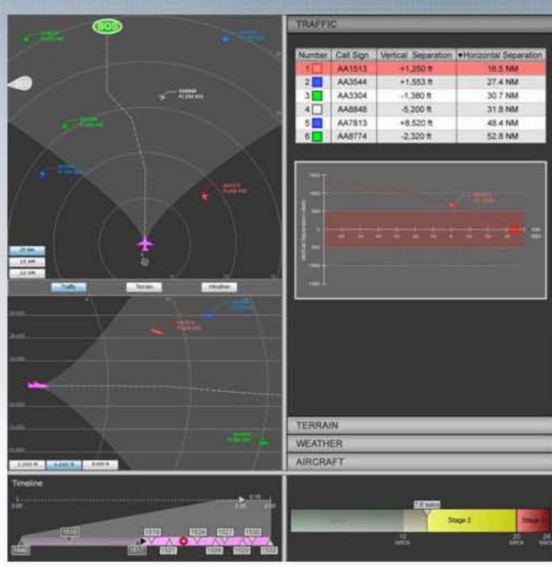
Vertical Display

Timeline

Color-coded urgency

Mock-ups designed by Andy Chang M.S. and Dr. Amy Alexander

Stage 2 Automation

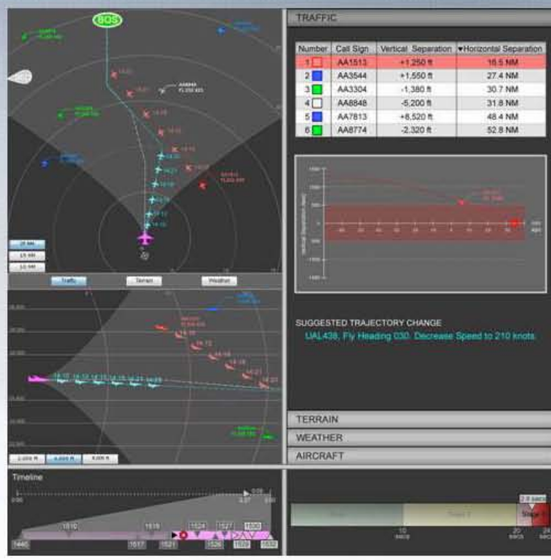


Supports analysis

Information automatically displayed on right

Mock-ups designed by Andy Chang M.S. and Dr. Amy Alexander

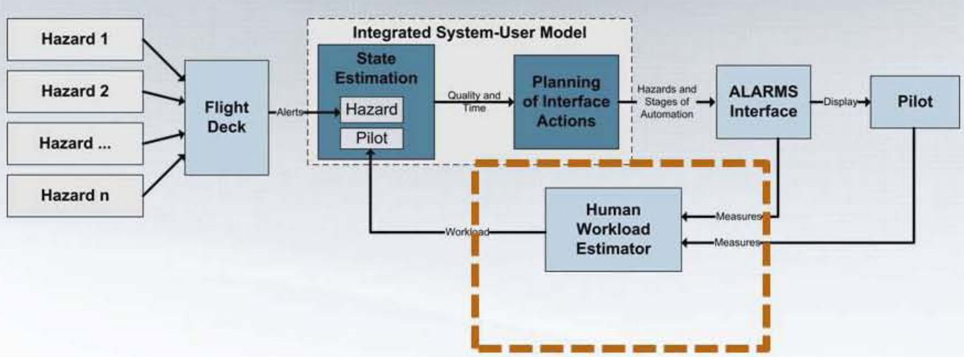
Stage 3 Automation



Decision Support

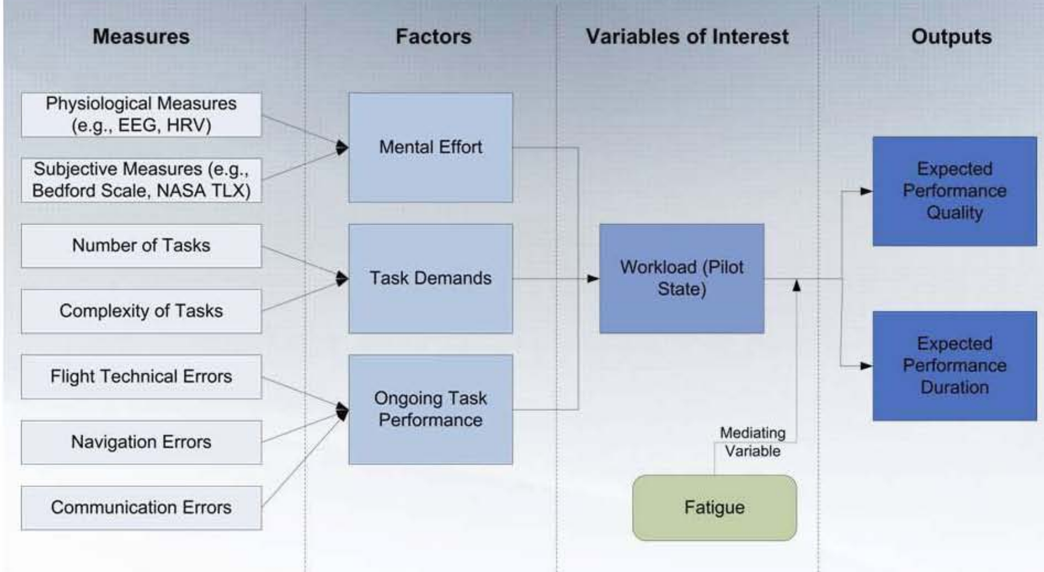
Mock-ups designed by Andy Chang M.S. and Dr. Amy Alexander

Integrated User Model



Pilot State Refinement

APTIMA HUMAN-CENTERED ENGINEERING



Workload

APTIMA HUMAN-CENTERED ENGINEERING

- Workload
  - Performance Response Function (PRF)
    - Finite amount of resource able to be allocated to problem
    - Time Sharing
    - Automaticity
  - Satisfactory workload Index has these properties:
    - Diagnoses cause of variation
    - Selective of factors
    - Unobtrusive
    - Reliable

©2010, Aptima, Inc.

Factors

APTIMA HUMAN-CENTERED ENGINEERING

- **Mental Effort**
  - Subjective Information
    - NASA TLX scale
    - Bedford Workload Scale
  - Physiological Information (EEG, HRV)
- **Task Demands**
  - Task Complexity
  - Number of Tasks
- **Ongoing Task Performance**

 13

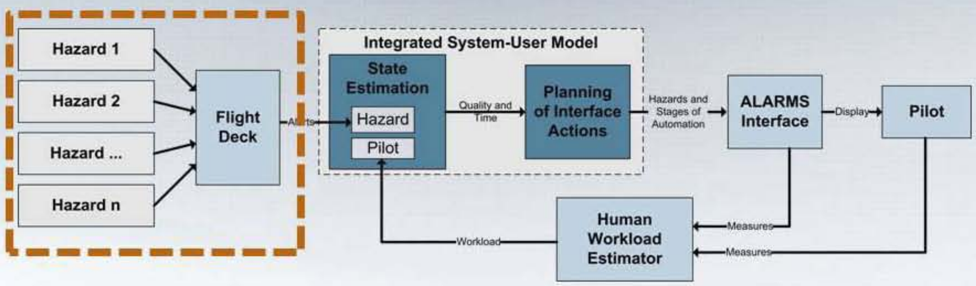
Pilot State Refinement

APTIMA HUMAN-CENTERED ENGINEERING

- **Mental Effort, Task Demands, and Ongoing Task Performance are used to compute Workload**
  - We model these as Low/Medium/High, take linear combination
  - Workload = $\alpha * ME + \beta * TD + \gamma * TP$
  - Workload is combined with information about hazard state and stage of automation to determine the quality and duration of action.
- **Increasing Pilot Workload leads to:**
  - Greatly decreasing quality, increasing duration in Stage 1
  - Decreasing quality, increasing duration in Stage 2
  - Very small effects in Stage 3
- *Thus, increasing Pilot Workload will tend to increase the stage of automation.*

 14

Integrated User Model



Identify CWA (Caution/Warning/Alerting) Applications For NextGen and Develop a System Architecture

- Design Approach –
  - Define Alerting Levels

Alert	Timeframe
Directive	<10 seconds
Warning	10--15 seconds
Caution	< 40 seconds
Advisory	non-critical

23 Existing Systems Reviewed

13 NextGen Planned Systems Reviewed

  - Define Hazard List
    - System Failure
    - System Performance Compromised
    - Loss of Separation
    - Adverse Weather Encounter
    - Altitude Deviation
    - Navigation Deviation
    - Controlled Flight Into Terrain
    - Crew Incapacitation
    - Flight Performance Compromised
    - Structural Failure
    - Life Support
    - Protected Airspace Incursion
    - Loss of Communication
    - Runway Incursion

Hazard Matrix

APTIMA HUMAN-CENTERED ENGINEERING

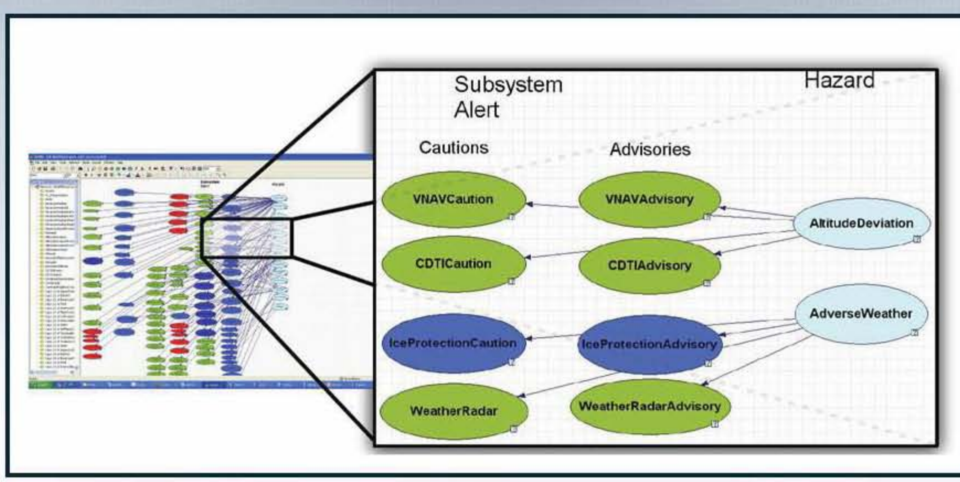
Tools/Technology/Systems	Sub-Systems	System Failure	System Performance Compromised	Loss of Separation	Adverse Weather Encounter	Altitude Deviatio
Electrical (3.1.11)		(C)	(W)			
Hydraulic (3.1.15)		(C)	(C)			
Fuel (3.1.14)		(C)	(C)			
Landing Gear (3.1.17)		(W)	(C)			
Air Conditioning/Pressurization (3.1.21/22)		(W)	(C)			
Ice Protection (3.1.16)		(W)	(W)		(C)	
Fire Protection (3.1.12)		(C)	(C)			
Enhanced Ground Proximity Warning (EGPWS) (3.1.20)		(A)	(A)		(W)	
Navigation Radio (Nav Radio)						
	Enroute (3.1.1)	(A)	(A)			
	Approach (3.1.1)	(A)	(A)			(A)
Flight Management System (FMS)						
	RNAV (3.1.18)	(A)	(A)			

A=Advisory, C=Caution, W=Warning

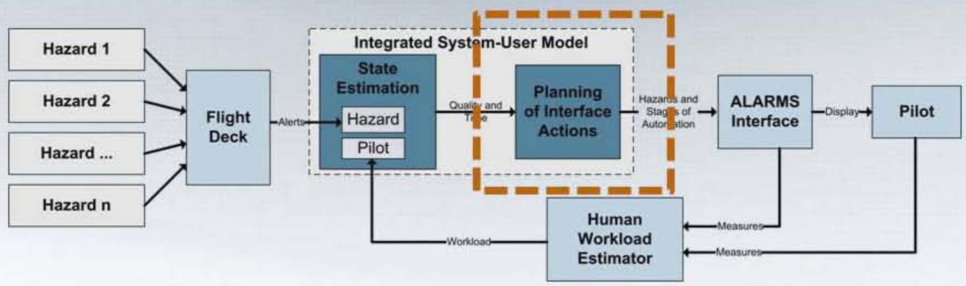
Blue=Aviation, Green=Navigation, Purple=Communication

ALARMS Bayesian Network

APTIMA HUMAN-CENTERED ENGINEERING



Integrated User Model



TMDP model

- TMDP model (Boyan and Littman 2000)
  - <S,A,P,d,R>
  - Set of States: Mapping from hazards to alert levels
    - Example: <N,A,W>
  - Set of Actions: Stage of Automation for each hazard
    - Stage 1 = human-centric, Stage 3 = highly automated
    - Example <1,3,2>
  - Probabilistic Transitions: P(s,a,s')
  - Action Durations: Probability Density Function d(s,a,s',t)
    - Deadline Δ after which there is no reward
    - More automated actions: Quicker
  - Reward for addressing hazard and returning to nominal state
    - More automated actions: Lower reward
  - Policy: Mapping from state and time to action

Pilot State Refinement

APTIMA HUMAN-CENTERED ENGINEERING

- Mental Effort, Task Demands, and Ongoing Task Performance are used to compute Workload
  - We model these as Low/Medium/High, take linear combination
  - Workload = $\alpha * \mathrm{ME} + \beta * \mathrm{TD} + \gamma * \mathrm{TP}$
  - Workload is combined with information about hazard state and stage of automation to determine the quality and duration of action.
- Increasing Pilot Workload leads to:
  - Greatly decreasing quality, increasing duration in Stage 1
  - Decreasing quality, increasing duration in Stage 2
  - Very small effects in Stage 3
- *Thus, increasing Pilot Workload will tend to increase the stage of automation.*

CPH Planner (Marecki et al. 2007)

APTIMA HUMAN-CENTERED ENGINEERING

- Define values V(s)(t)
- Expected Reward (for Bellman Backup)

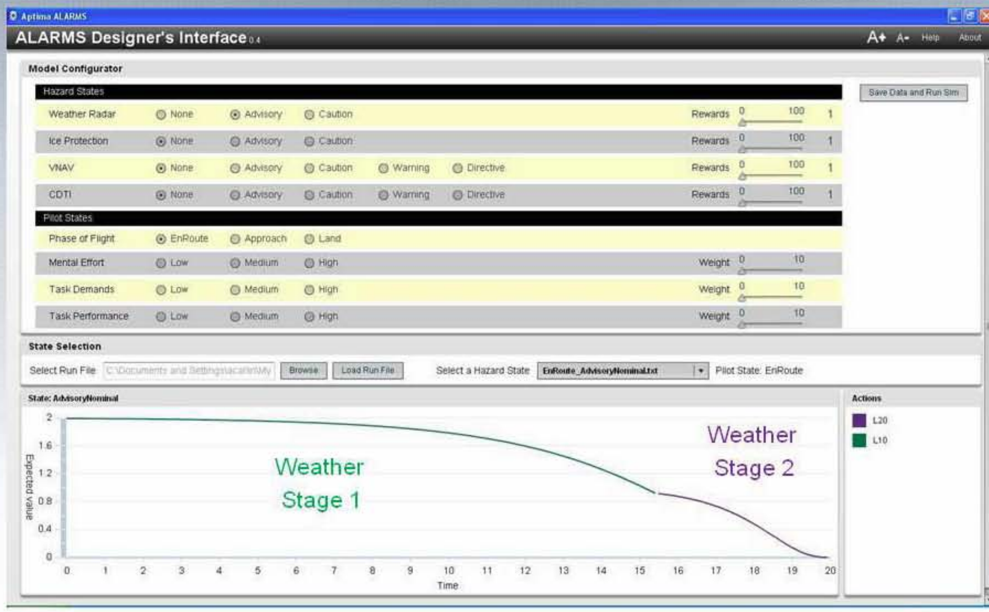
$$\arg\max_{a \in A(s)} \left(\sum_{s' \in S} P(s'|s,a) \int_0^t p_{s,a}(t')(R(s,a,s') + V^*(s')(t-t'))dt' \right)$$

- Since continuous, need approximate value functions
  - Convolution becomes intractable
  - Solution: phase-type distributions
  - Convert to MDP with uniform action durations

$$\int_0^t p(t') V(s')(t-t')\, dt'$$

ALARMS System Designer Interface

APTIMA HUMAN-CENTERED ENGINEERING

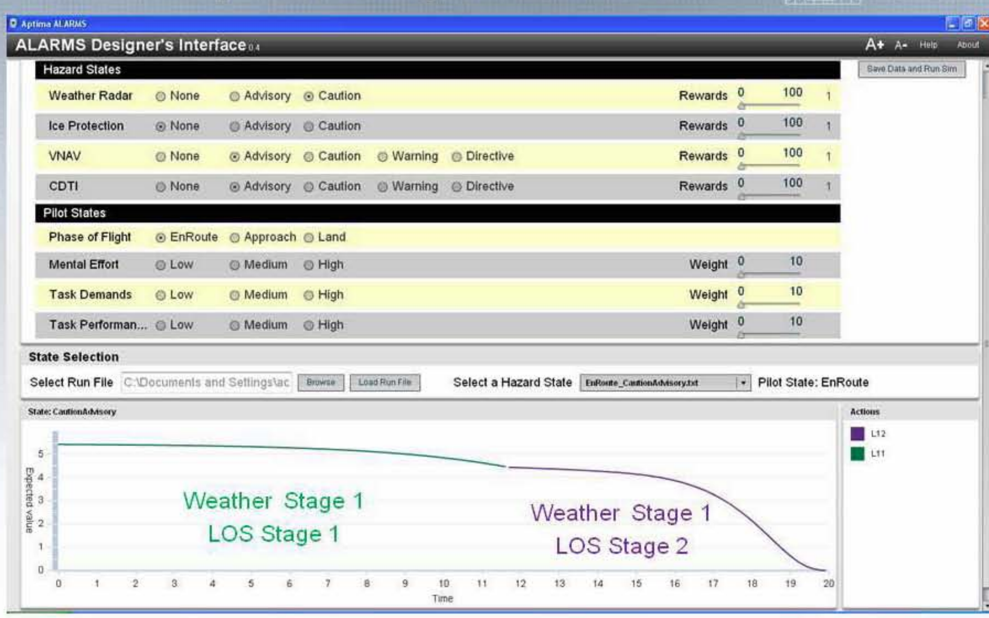


*Interface GUI programmed in FLEX by Gilbert Mizrahi at Aptima

 23

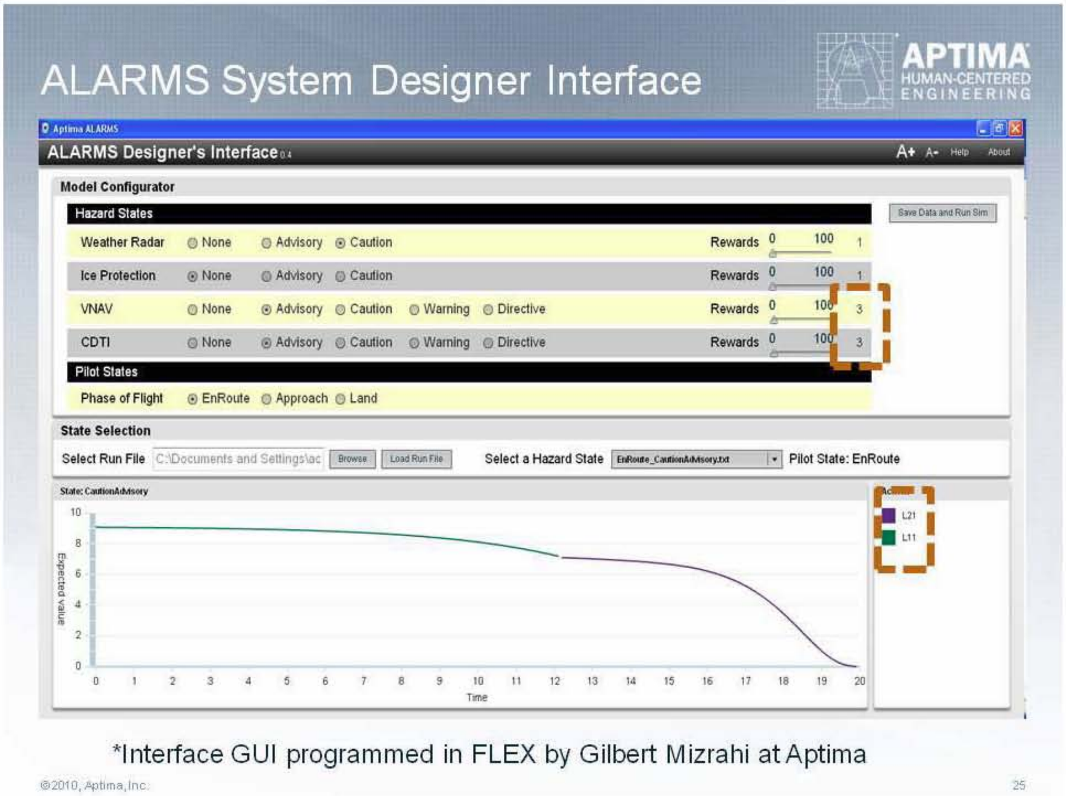
ALARMS System Designer Interface

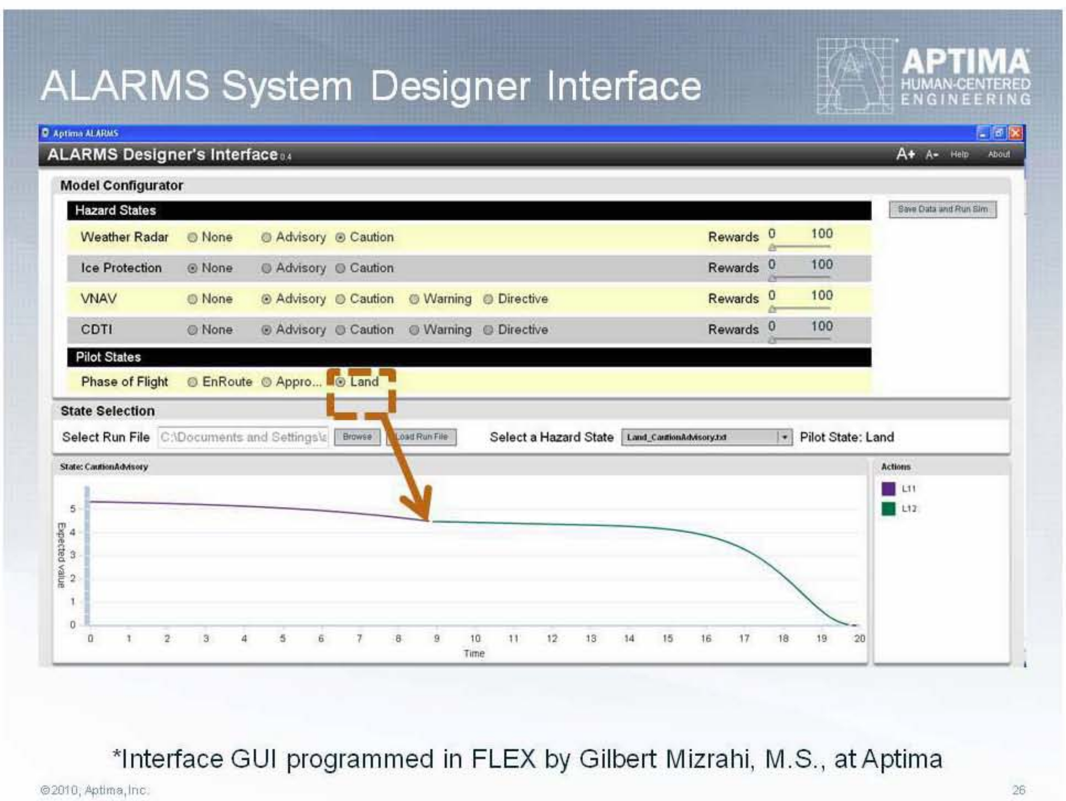
APTIMA HUMAN-CENTERED ENGINEERING



*Interface GUI programmed in FLEX by Gilbert Mizrahi at Aptima

 24


ALARMS System Designer Interface
APTIMA
HUMAN-CENTERED ENGINEERING
ALARMS Designer's Interface
Model Configurator
Hazard States
Weather Radar None Advisory Caution Rewards 0 100 1
Ice Protection None Advisory Caution Rewards 0 100 1
VNAV None Advisory Caution Warning Directive Rewards 0 100 3
CDTI None Advisory Caution Warning Directive Rewards 0 100 3
Save Data and Run Sim
Pilot States
Phase of Flight EnRoute Approach Land
State Selection
Select Run File
Browse
Load Run File
Select a Hazard State
Pilot State: EnRoute
Expected value
Time
L21
L11
*Interface GUI programmed in FLEX by Gilbert Mizrahi at Aptima
©2010, Aptima, Inc.
25


ALARMS System Designer Interface
APTIMA
HUMAN-CENTERED ENGINEERING
ALARMS Designer's Interface
Model Configurator
Hazard States
Weather Radar None Advisory Caution Rewards 0 100
Ice Protection None Advisory Caution Rewards 0 100
VNAV None Advisory Caution Warning Directive Rewards 0 100
CDTI None Advisory Caution Warning Directive Rewards 0 100
Save Data and Run Sim
Pilot States
Phase of Flight EnRoute Appro... Land
State Selection
Select Run File
Browse
Load Run File
Select a Hazard State
Pilot State: Land
Actions
L11
L12
Expected value
Time
*Interface GUI programmed in FLEX by Gilbert Mizrahi, M.S., at Aptima
©2010, Aptima, Inc.
26

Conclusions

APTIMA HUMAN-CENTERED ENGINEERING

- Contributions of ALARMS
  - Introduced system architecture
  - Developed interface mockups
  - Identified current and future hazards
  - Bayesian reasoning over uncertain hazard state and sensor systems
  - TMDP planning of interface stages of automation
    - TMDP planning reasons about time duration uncertainty
    - TMDP model is adaptable to empirical findings
  - Accounts for pilot state
  - Developed system designer interface for understanding system behavior
- Future work
  - System integration and empirical evaluation

Questions?

APTIMA HUMAN-CENTERED ENGINEERING

Questions?