Framework for Evaluation of Human-System Issues with ASDE-X and Related Surface Safety Systems

A *White Paper*

Prepared by

Raja Parasuraman
Catholic University of America
John Hansman
Massachusetts Institute of Technology
Steven Bussolari
Massachusetts Institute of Technology

January 12, 2002

TABLE OF CONTENTS

Summary

Systems that are being developed to enhance surface safety at towered airports need to be designed with attention to human-system issues. This White Paper outlines a framework for the evaluation of human-system issues in the development and use of new technologies such as ASDE-X. The framework is presented in the form of a flow chart of applicable methods that can be applied in an iterative design process. Although the framework is discussed with respect to components of ASDE-X, the general approach is applicable to a wider range of surface safety solutions. The framework describes the best available methodologies associated with the assessment of the following major human-system elements: (1) detection system performance; (2) total system performance; and (3) controller performance, focusing on situation awareness and workload. The methodologies discussed include signal detection theory (SDT) and Receiver Operating Characteristic (ROC) analysis, Bayesian statistics, fuzzy SDT, System Operating Characteristic (SOC) analysis, Monte Carlo simulation, fast-time simulation with probability density function representation of human and system response times, computational cognitive modeling, and modeling of human mental workload and situation awareness (SA). To illustrate the application of the framework in an iterative design process, two examples are provided: (1) the controller alerting function to ASDE-X, and (2) the addition of a runway status lights function to ASDE-3/AMASS. By providing quantitative methodologies for examining human-system issues, the framework allows for an evaluation of the efficiency and safety of surface safety systems. This approach will provide the FAA the means to determine the incremental value to safety of the implementation of different components of proposed systems, as well as to the continuing evaluation of safety once systems are fielded.

Acknowledgements

The authors wish to thank the following individuals for helpful discussion and comments on previous versions of this White Paper: Glen Hewitt, Mark Rodgers, and Mike Curry, Federal Aviation Administration; Tom Reynolds, Massachusetts Institute of Technology; and Jacqueline Duley, Booz Allen Hamilton Associates.

This White Paper reflects the opinions of the authors and not necessarily those of the Catholic University of America, the Massachusetts Institute of Technology, or the Federal Aviation Administration.

List of Figures

1. Introduction

The safety and efficiency of airport surface operations represent important areas of concern in an era of rapid growth in air travel and demand for air services. The September 11th attack and subsequent events have forced renewed attention[1] to aviation security issues and led to a reduction in the volume of air travel. The downturn is expected to be temporary, however, and runway incursions are likely to remain a critical area for concern, as they were in the US Congress prior to the recent terrorist attacks[2]. A recent report has documented the trends in runway incursion severity at US towered airports[3]. During the period 1997-2000, of approximately 266 million airport operations, there were 1,369 runway incursions, of which 3 resulted in accidents. A recent analysis estimates about 15 fatal runway accidents over the period 2003-2022, assuming no new counter-measures are implemented[4]. The recent runway accident in Milan, Italy between an MD-87 and a Cessna Citation II, which resulted in 118 fatalities, adds urgency to the issue.

Several technological and procedural solutions to the runway incursion problem have been proposed over the past few years. These are outlined in the FAA's *National Blueprint for Runway Safety*. The relative efficacy of these proposed initiatives is currently the focus of investigation. At the same time, the effectiveness of other technologies that are being considered to enhance surface safety at towered airports needs further evaluation. One of the systems under development is the Airport Surface Detection Equipment—Model X (ASDE-X). In addition to questions related to the efficiency and reliability of this and related surface safety technologies, the evaluation process needs to consider human-system issues in using new technologies and procedures[5] [6].

2. Framework for Human-System Analysis of Surface Safety Technologies

The purpose of this White Paper is to propose a framework for the analysis and resolution of human-system issues related to surface safety. Application of the framework should allow for identification of the improvement in surface safety that may be expected with the fielding of relevant systems such as ASDE-X. Although this framework is discussed with respect to components of particular technologies, e.g., ASDE-X, the general approach is applicable to a wider range of surface safety systems as well as other human-in-the-loop systems.

Our proposed framework has three elements which are described independently, but with the recognition that they also interact with and have implications for each other. The framework describes the best available methodologies associated with each element and shows how they can be applied. The three elements are:

1. Detection System Performance
2. Total System Performance
3. Controller Performance, Including Situation Awareness and Workload

The first element, detection system performance, is concerned with the application of analytical methods for determining the effectiveness with which surface objects ("targets") and conflicts between targets ("unsafe" states between targets) are detected using new surveillance technologies. Several mathematical/computational procedures for assessing detection sensitivity are described.

The second element addresses total system performance, i.e. the performance not only of the detection component but all other components, including human-centered aspects related to the controller and other humans (e.g., pilots) in the system. This element includes analysis of what kinds of resolution procedures should be adopted once detection has been achieved, and the relative impact of these procedures on issues such as time to alert and controller-pilot communications.

The third element considers the impact of these elements on the performance of tower controllers in using the overall system. While all aspects of controller performance are important, including usability of the interface and acceptance, this element focuses on situation awareness (SA) and workload. Any deficiencies in these aspects of controller performance that are linked to specific design features can severely limit the effectiveness of systems such as ASDE-X in enhancing runway safety.

The methods and models that can be used in this framework are described in more detail in Section 4 of this White Paper. Figure 1 provides a general flow chart to indicate how the framework can be applied in an iterative design process. We assume that the problem of preventing runway incursions is of sufficient criticality that it requires a "layered" defense, in which different technologies and procedures are used so that if any one solution is not sufficient, others can mitigate the problem.

As Figure 1 shows, the initial step must involve a delineation of the operational objectives, including the identification of a set of critical runway incident scenarios that the system is designed to prevent (e.g., landing on a runway occupied by an aircraft waiting to depart). These scenarios can be derived from the critical situations that have been identified by the FAA Runway Safety Program and by the Airport Movement Area Safety System (AMASS) product team. This step should also involve the identification of "worst case" scenarios. Analysis of worst cases can provide estimates of the improvement in safety to be expected from or identify the limits of new surveillance systems and other technologies and procedures aimed at reducing runway incursions.

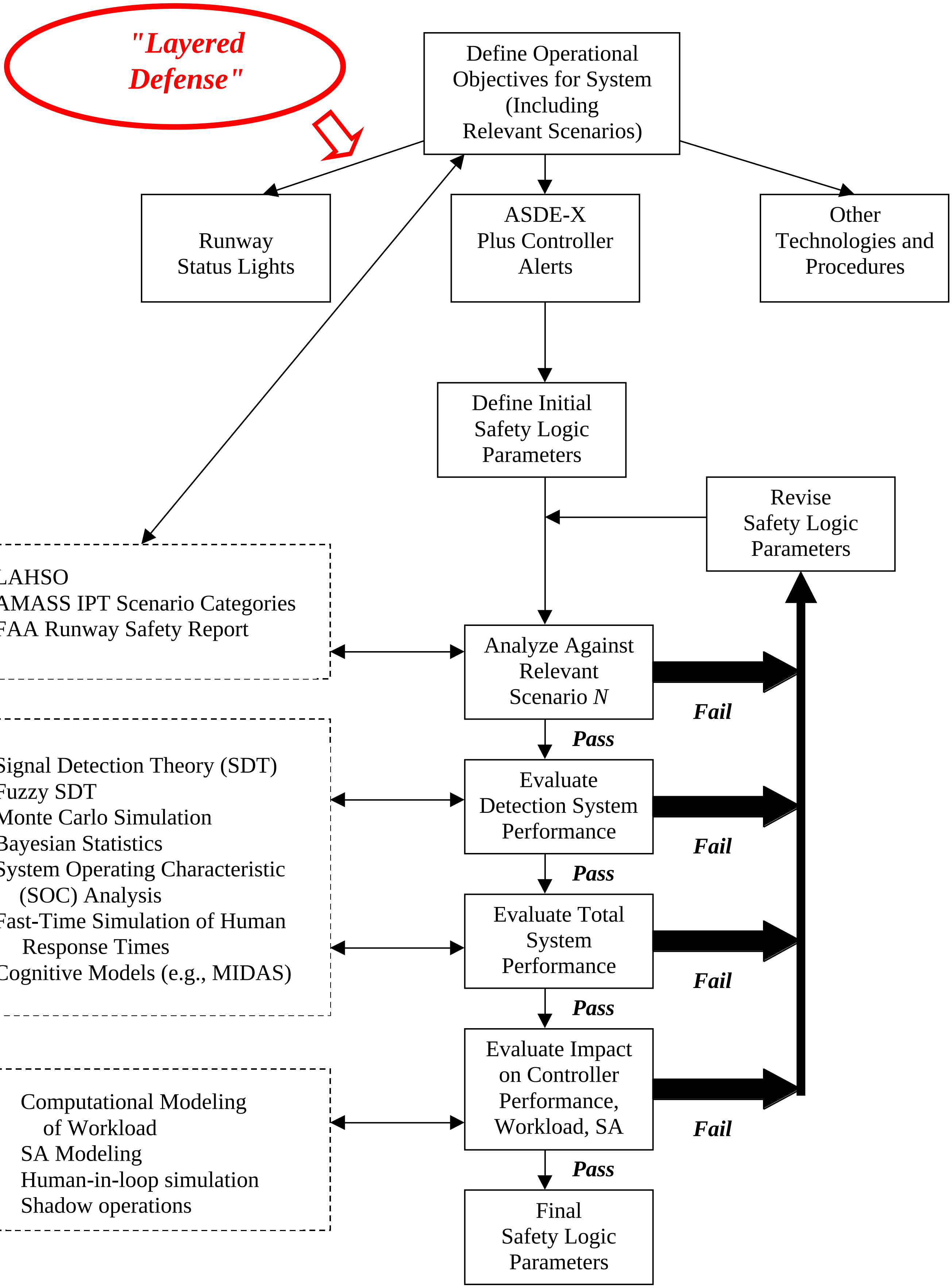

Figure 1. Flow Chart for Application of Human-System Analysis Framework

The next step, shown in Figure 1, is to identify the means (technologies, procedures, or both) by which the surface safety system will aim to prevent incursions and accidents. An example (illustrating the layered defense concept) would be a combination of an ASDE-X surveillance system, tower displays, alerts for controllers, data link traffic for pilots, and runway status lights. For each scenario identified in the previous step, all possible mechanisms should be identified. For example, there would be several defenses in the case of an aircraft landing on a runway occupied by an aircraft waiting to takeoff: the ASDE-X display improves the situational awareness of the controller who identifies the problem and intervenes; alerts that warn the controller of the impending hazard and restores lost situational awareness provide a further layer of defense are; in addition traffic displays on either aircraft warn the flight crews (in this scenario, runway status lights are not contributors to the defense). Note that we treat the surveillance system (sensors and trackers) as functionally different from controller displays and controller alerts, even though a composite program like ASDE-X may combine elements of each. The functional distinction between components is important and part of the rationale behind the framework put forth in this White Paper. The framework provides methods for determining the effectiveness of different functional components as well as the system as a whole. As a result, application of the framework will allow the FAA to determine, for example, whether further investment is necessary for controller alerts in addition to an ASDE-X system comprised of a surveillance system and controller display.

Once each aspect of the layered defense has been identified, the next step is to specify the initial design parameters for each component. The middle blocks of Figure 1 shows the iterative design process for the safety logic of an alerting capability that will be added to the basic ASDE-X surveillance system. (A similar process can be carried out for different ASDE-X parameters).

Following an initial definition of the safety logic parameters, the effectiveness of the system in dealing with the set of relevant scenarios should be evaluated. Successive steps using the models and methods (described later) are then used to evaluate if the system succeeds or fails in meeting the demands of each scenario. For example, the surveillance system detection performance (and the associated safety logic) can be modeled using signal detection techniques. If the detection system performance criteria are not met (e.g., unacceptably high false alarm rate) then the safety logic parameters will need modification. After this, the total system performance needs to be evaluated using such methods as System Operating Characteristic (SOC) analysis and fast-time simulation of human performance (described in more detail below). Finally, methods for evaluating the impact of the system design on controller workload and SA have to be applied. If a particular design parameter results in the workload on the controller being predicted to be too high, then an appropriate change should be made. The methods of evaluation include computational modeling and human-in-the-loop simulation. Because runway incursions are very rare events, however, data collection for performance evaluation may be difficult and alternative methods, such as running a "shadow system," may be needed in evaluating the impact on controller performance, workload and SA.

3. Human Factors in Systems Design

Before describing the human-system issues and the appropriate methods and models to support the iterative design process shown in Figure 1, some discussion is necessary concerning what role Human Factors (HF) should play, if any, in the initial requirements development and design of new surface safety systems. Ideally, a human-centered approach should be an integral part of all phases of system acquisition, from systems requirements definition and design, field testing, operations, and maintenance. It is important to recognize that HF is much more than just consideration of display or Graphical User Interface (GUI) issues. HF also requires user input, but it is not just a cataloging of user preferences (e.g., what controllers like or don’t like about a particular system). It is critical to consider both how the user understands and relates to the system behind the interface and the overall performance of the combined human-automation system. The overall performance is also strongly influenced by the operating procedures which drive how the humans use and respond to the technical systems. There is therefore a need to integrate operating procedures in the system design space and to consider tradeoffs between technical system requirements and modifications to operating procedures as part of the overall system development process. A good example of this type of integration is illustrated in the recent redesign of the Australian airspace coupled with the modernization of the Australian air traffic control (ATC) system[7].

We recognize that inclusion of HF in the system requirements stage is not standard practice at the FAA and that the definition of the operating procedures is often separate from the acquisition process, but the merits of this approach to system design and acquisition must nevertheless be reiterated. It is our view that if HF is to be considered at all, it must be included as part of the system *functional requirements development* process. Otherwise any HF input is likely to be sub-optimal and will not resolve user difficulties that may arise in working with new technologies. The recommendation to apply HF input early rather than late in system design has been made before to the FAA[8]. However, even early HF input may be ineffective when it deals *only* with user interface design or the details of a particular GUI and not with the basic functionality of the system. In such systems, there can be a mismatch between the functionality as specified by the designer, the operating environment (i.e. procedures), and the user's requirements for the system or his or her mental model of system functionality. The result can be inefficient system performance, errors, and possible adverse performance including accidents. User interface design can, at best, provide only a "band aid" when systems are designed this way and cannot compensate for problems in the underlying functionality. In contrast, acquisition approaches that include user task and performance modeling, procedures evaluation, and training assessment in the functional requirements developmental stage result in a better match between the designer's and user's model of the system and can improve system efficiency. Examples of this integrated human-centered approach to system design are Riley's Cockpit Control Language for cockpit flight management systems (FMS)[9] and Vakil and Hansman's operator directed process (ODP) approach to the design of FMS control modes[10].

An example of the integrated human centered approach is shown in Figure 2, which presents both a simplified functional diagram of a surface safety system (that includes ASDE-X as well as other components) and a process flow from the point of view of the tower controller. User interface issues are being considered as part of the ASDE-X and other surface safety systems. This work focuses on the presentation of the integrated surveillance systems output to the tower controller, as represented by the box labeled "Display" between the Sensor/Data Integration system and the controller in Figure 2. Interface design is typically concerned with the *how* of information presentation, i.e., the details of the display to be provided to the controller (icon types, color codes, menu design, control device configuration, etc). These details are important and do play a role in how effectively the controller can use the system. However, fundamentally more important is the core functionality and functional logic of the system, i.e. the boxes labeled "Surveillance/Data sources" and "Sensor/Data Integration and Alerting" in Figure 2 and the associated safety logic or decision algorithms that this hardware uses. This work focuses on the *what* and *why* of information presentation: what functions are part of the system, why certain system parameters are selected, and what the information requirements are for the controller (or pilot).

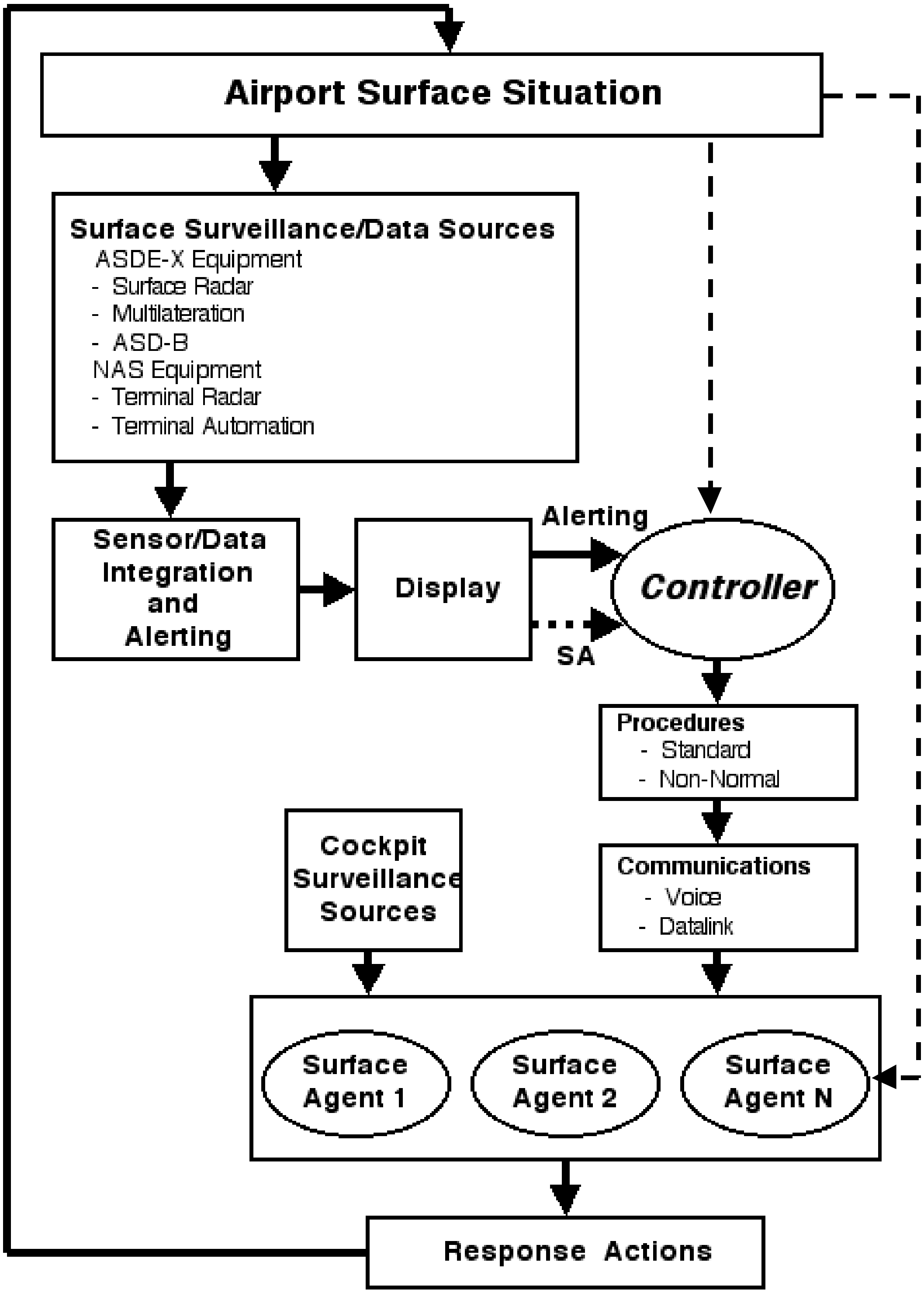

Figure 2. Surface Safety System: Controller-Centered Process Flow

Consideration of the human-system issues discussed below in Section 4 onwards can provide for HF input into system functional specification first, followed by interface design. It is also important to reiterate that the effectiveness of the methodologies discussed in this White Paper for incorporating human-system parameters in the design process can be enhanced if operating procedures are simultaneously considered. Total system performance in the operational environment involves an interaction between operating procedures and technical components. The current paradigm for technical system development appears to assume that the operating procedures are fixed and can only be minimally adjusted (because of the difficulties of implementing changes in procedures). As a consequence the system is developed, often at great expense, to be consistent with the current operating procedures. Although there are working groups within and outside the FAA that are attempting to negotiate changes in operating procedures, these are often only determined once the equipment is in advanced stages of development or even in the field. It is desirable to consider a more holistic development approach where changes in the operating environment and procedures are considered as part of the design space early in the system development process.

4. Surface Surveillance System: Controller-Centered Process Flow

The controller-centered process flow diagram in Figure 2 includes the basic components of a surface safety system including ASDE-X.. The various human-system issues can be considered in relation to this flow diagram. Data from the Airport Surface Situation is provided by the Surface Surveillance and Data Sources. In the ASDE-X example the sensed data include: Terminal Radar, Surface Radar, Multilateration and Automatic Dependent Surveillance-Broadcast (ADS-B).

Additional data is provided from the Terminal Automation System. Other data sources, as appropriate, could be included in the analysis. The sensor and Terminal Automation System data is integrated and processed in the Sensor/Data Integration and Alerting box which represents both the fusion of the different data sources but also the alert logic which would trigger alerts based on the available data. The output of the Sensor/Data Integration box is presented to the controller through some display mechanism (Display box in Figure 2). The display may be visual (as specified for the ASDE-X system), audio, tactile, or some combination of these sensory modalities. The primary interface between the system and the controller is represented at the interface between the Display box and the tower controller. Note that two types of information are schematically represented at this interface. The dark arrow explicitly represents the output of the alerting system. The shaded arrow represents information from the system which is used to support the tower controller's SA. Note that the SA information provided to the controller is independent of whether the system is in an alert state or not. It should also be noted that the controller also has direct visual and audio surveillance of the surface situation, which is a key factor in building his or her SA. This is also represented in Figure 2.

Given direct surveillance and the output of the system, the controller will assess the situation and determine control actions based on this assessment and the appropriate operating procedures. The control commands are currently communicated by voice, but may be communicated by datalink in future systems. This communication process is represented by the Communications box in Figure 2. In considering the total system performance, it is important to model the processing delays both in the controller's determination of action and in the communications process. It is also important to consider potential errors both in control action and in miscommunications.

The output of the communication process is commands that are transmitted to the various surface agents which the controller is responsible for. In addition to aircraft the controller may be responsible for ground vehicles or even pedestrians on the airport surface. Figure 2 also represents additional input which the agents may have to support their own SA including cockpit surveillance sources for aircraft. Cockpit surveillance sources can be of several types: the simplest is direct visual observation (including enhanced surface markings) of the airport surface area; other sources include any “party line” information which pilots obtain from other pilots; finally, more direct sources can include cockpit displays of traffic information (CDTI) or cockpit moving map displays that support the pilots' awareness of their location on the airport surface.

Based on their own SA and the commands received from the controller the agents in the surface situation will respond with some finite response dynamics. Again, in determining the total system performance, the time to respond as well as the accuracy of response of all the agents must be considered. Methods to do this are outlined in Section 7.

5. Runway Scenario Identification

One of the first steps in applying the framework described in this White Paper is the identification of relevant surface scenarios. Listing and categorization of scenarios is important from the perspective of setting the operational objectives of the surface safety system (see section 4). Scenario identification is also important for testing components of the system using the appropriate human-system analysis methods outlined within our framework (see Figure 1).

There are several approaches to scenario description. One approach is to determine the “worst-case” scenarios which will be most problematic for the system. The challenge with this approach is that it is likely that the runway incursion problem is sufficiently difficult (due to the fast dynamics and the close proximity of safe and unsafe states on the surface) that no system will be able to respond to the “worst cases.” Another approach is to look at the ability of the system to respond to likely events. One method is to base the analysis on the most critical events which have been identified in the analysis of potential Runway Incursions.[3, 11]

6. Detection System Performance

The first element in our framework for human-system analysis is detection performance. This involves the sensitivity of the surveillance system for detecting of critical objects on the surface. The detection problem can be divided into two components. The first concerns evaluation of sensor sensitivity (e.g., radar) for detecting targets. However, sensor detection sensitivity is necessary, but not sufficient. The mere detection of an object (e.g., an aircraft crossing a runway) does not necessarily indicate a problem; rather an unsafe state must be detected (e.g., an aircraft crossing an active runway being used by another aircraft). Hence the second component detection problem concerns the evaluation of the sensitivity of the surveillance and associated decision support system.

Several methods are available for the analysis of the effectiveness of the detection performance of surface targets. These include signal detection theory[12], Bayesian statistics[13],[14] and fuzzy signal detection theory[15]. Signal detection theory (SDT) can be used to determine the effective sensitivity of a given detection system and the appropriate decision threshold that should be used to achieve a given tradeoff between two types of errors of detection—missed targets and false alarms. Bayesian statistics provides a means for assessing the relative likelihood of a hazardous event (e.g., a conflict between an aircraft and a surface vehicle) given that the detection system has detected a target. Finally, fuzzy signal detection theory provides a method to capture the inherent uncertainty and variability in the definition of a surface "conflict" and to provide estimates of detection sensitivity and decision criteria under such conditions.

These methods have proven validity for a variety of applications and can be readily applied to the analysis of detection efficiency of surface targets. An additional advantage is that the methods can be applied to the analysis of machine performance alone (e.g., ASDE-X efficiency in target detection) as well as to *joint* human+machine performance (e.g., controller + ASDE-X). Evaluation of sensitivity in detecting unsafe states, however, may require additional analyses that may not be captured by these detection performance methodologies. In the case where specific hazard scenarios can be identified, it is possible to analyze the time evolution of the scenario and assign a time budget to the various response processes (e.g., surveillance, alert generator, controller response, communication, pilot response, aircraft response, etc.) to determine the overall system performance[16].

6.1 Surface Target Detection Performance

The specifications for ASDE-X (Version 1.1)[17] call for a system capable of "tracking 200 combined real surface and approach (arrival) targets from the sensor plot reports" (Section 3.1.1.1). Additionally, the system must be able to detect targets having a 3 m^2 or larger radar cross-section (ranging from small surface vehicles to wide body jets) with a minimum detection probability P_D of 0.9 and a maximum probability of a false alarm P_{FA} of 10^{-6} (Section 3.2.3). Standard SDT can be used to determine the effective detection sensitivity (d') of the system given that these requirements are met. For example, with these correct and false alarm probabilities, detection sensitivity d' can be computed fairly easily as $d' = z(P_D) - z(P_{FA})$, where z() is the normal deviate of the Gaussian probability density function. With $P_D > 0.9$, and $P_{FA} < 10^{-6}$, $d' > 6$. A system with a detection sensitivity d' of 6 or higher would represent a highly sensitive system. Note however, that even with a very sensitive system, the probability that the output of the detection system (e.g., an "alert") represents a true hazard could be very low if the rate of occurrence of the event to be detected is very low. This is the so-called *base rate* problem and its relevance and effects are described further below.

Apart from determining the sensitivity of the detection system, the tradeoff between missed targets and false alarms must also be made. This is based on the decision criterion or threshold for deciding whether the given "evidence" (i.e. the raw sensor or integrated sensor data) warrants a "target present" response or not. If the decision threshold is set stringently, few false alarms will be made, but at the expense of some missed targets. Conversely, for a more lax decision criterion, there will be fewer misses, but at the cost of more false alarms. Safety-critical detection systems seek to minimize misses, so that they are typically designed with a relatively liberal decision criterion. In SDT, the criterion ß is computed easily as $ß = y(P_D)/y(P_{FA})$, where y() is the ordinate of the Gaussian probability density function. For example, consider the ASDE-X system with a detection accuracy of $d' > 6$, as discussed previously. Suppose it is required that the false alarm rate should be no more than .001. Then the decision threshold for the system should be set such that ß = 1.7 or higher.

The tradeoff between misses and false alarms is represented in SDT by a Receiver Operator Characteristic (ROC), which plots P_D against P_{FA}[12]. As the decision criterion becomes more liberal (i.e., ß decreases), fewer misses are made (i.e., P_D increases), but at the expense of more false alarms. An example of ROC analysis of the Ground Proximity Warning System (GPWS) is shown in Figure 3, taken from Kuchar[18] [19]. This analysis was based on a model of steep terrain. The chosen decision threshold for the B-767 installation of GPWS is shown. This results in a fairly high false alarm rate. A more stringent threshold could be chosen to reduce the false alarm rate (moving downwards and to the left along the ROC), but this would result in an increase in the miss rate (i.e., reduction in P_D).

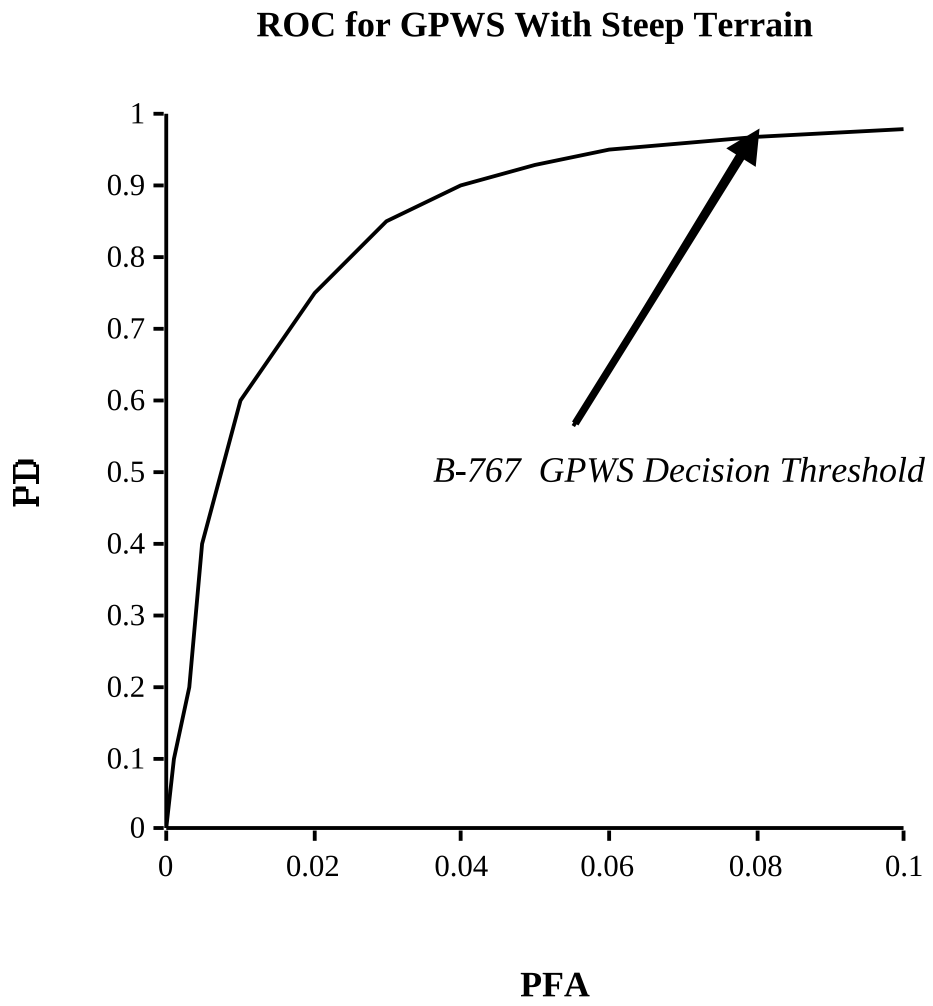

Figure 3. Receiver Operating Characteristic (ROC) for the GPWS System, Assuming Steep Terrain. The Decision Threshold ß Associated with the B-767 GPWS is shown. [From Ref. 18]

Similar types of ROC analysis should be carried out for the ASDE-X system for different surface targets.. Several additional issues must also be addressed other than just the sensitivity of the system in detecting targets and the tradeoff between misses and false alarms. First, the sensitivity of ASDE-X or other system for detecting conflicts between surface objects needs to be examined for different factors such as geometries, airport layout, reduced visibility due to weather, etc.

Second, detection efficiency can generally be improved over time as additional sensor information is obtained. However, in time-critical situations, as many surface operations may involve, there is a premium for early or timely detection of a hazardous condition. Hence, there is a tradeoff between detection sensitivity and time to detection. SDT and ROC analysis can also be applied to the analysis of this tradeoff, just as it can to the tradeoff between missed targets and false alarms. For example, this procedure has been adapted for examining alerting criteria for the Traffic Alert and Collision Avoidance System (TCAS) in aircraft[19].

Third, while SDT and ROC analysis can be used to set the decision criterion *ß* so as to minimize the P_{FA}, adjusting the decision criterion for a low device false alarm rate may be insufficient by itself for ensuring high detection reliability. Alerts that prove not to involve true conflicts will lower controller trust in the system and increase controller workload[20, 21]. This may occur because detection systems are typically tested under conditions with known conflicts that occur at a frequent rate. In the real world, however, conflicts between surface objects will occur much less frequently. Even for a very sensitive detection system, application of Bayesian statistics shows that the low *prior* probability or *base rate* of most hazardous events under real operating conditions may limit the effectiveness of many systems[14]. For example, the rate of runway incursions over the past four years was estimated to be about 5 per every 1 million operations[3], which represents a base rate of .000005. Because of the low base rate, the *posterior* probability of true conflict, or the may be quite low.

The problem can be illustrated by the following hypothetical example. Suppose a detection system with high sensitivity is designed with decision logic such that the system misses only 1 of every 1,000 potential collisions between surface objects, while having a false alarm rate of about 1 in about 1,000. Despite the very high correct detection rate (.999) and low false alarm rate (.001), the posterior odds of a true alert with such a system might be extremely low under actual operational conditions where the base rate of incursions is likely to be low. If the base rate is .000005, as estimated previously, then the system will emit 1,005 alerts for every 1 million operations. However, only 5 of these will be true alerts representing a true incursion situation requiring immediate action, while the remaining 1,000 will be false. Thus, the posterior probability of a true alert will be unacceptably low, .0049 (= 5/1005). Note that even with a 10-fold reduction in the false alarm rate (from .001 to .0001), the system will still give 105 alerts for every 1 million operations, of which again only 5 will represent true alerts. Controller workload in dealing with so many false alerts could represent a problem, as could low trust in the system.

As another illustration, the effect of base rate has been estimated for the User Request Evaluation Tool (URET), which is a system for assisting en route controllers in determining conflict-free trajectories for aircraft[22]. Figure 4 shows the relationship between the probability of a true URET alert (positive predictive value or PPV) and the conflict base rate for three different URET false alarm values, as estimated, measured or specified by ACT-250 of the FAA, the MITRE Corporation, and the URET System Specification Document (SSD). As Figure 4 shows, PPV falls as conflicts become less and less frequent a priori. More importantly, for this range of estimated base rates (1 per 10 to 1 per 1000), PPV is uniformly low: the highest PPV of .18 translates into 1 true URET alert for every 5 alerts.

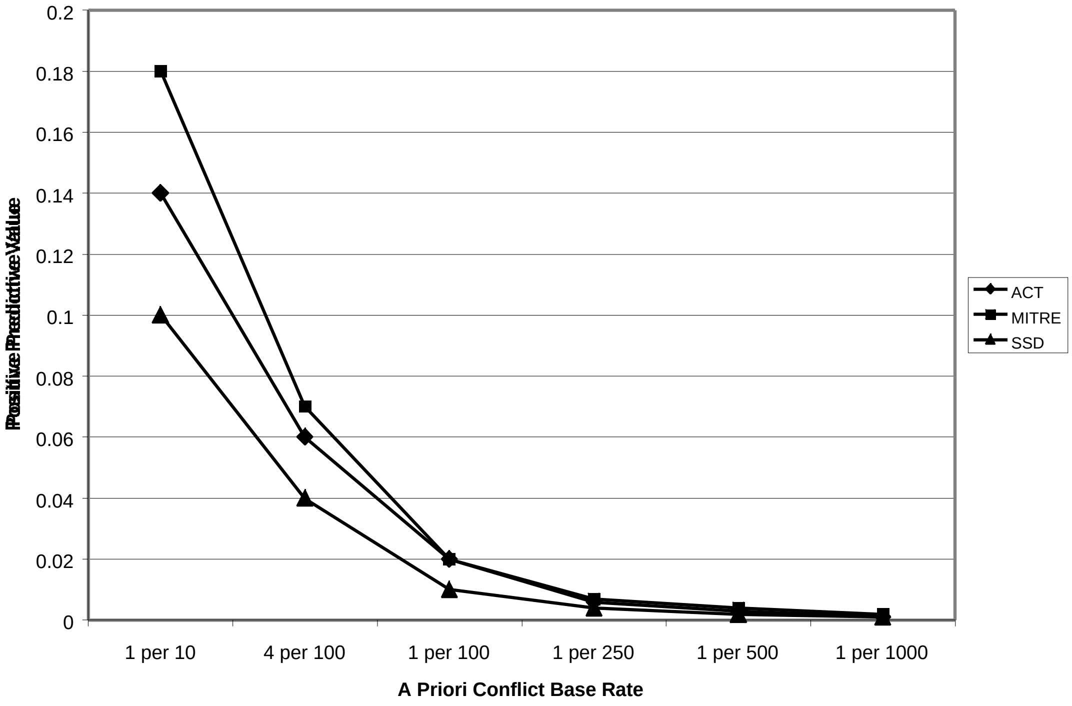

Figure 4. Effect of Base Rate on the Probability of True Alert for the URET System. [From Ref. 22].

This analysis highlights the challenge of designing reliable surface safety systems for detecting low-probability events. To counter the effect of a low base rate, automated detection systems must be designed to have high posterior probabilities of detecting a true hazard. Parasuraman and Hancock[23] outlined a set of equations that could be used by designers to determine the appropriate decision criteria that maximizes the posterior probability of a true alert. For example, the system parameters could be set so that the posterior probability of a true alert is at least .8. This means that of every 10 alerts that occur, 8 point to a true conflict. Applying this solution may work in a number of cases, so long as information on the base rate is available. But in other cases, maximizing the posterior probability of a true alert may result in an unacceptable increase in the missed target rate. The Bayesian approach can be used to supplement the SDT/ROC approach to determine the appropriate tradeoff between missed target rate, the false alarm rate, and the posterior probability of a true alert.

6.2 Unsafe State Detection Performance

Evaluation of target detection efficiency is necessary but not sufficient. In addition, the surface safety system must reliably detect an unsafe state between targets (aircraft/aircraft or aircraft/vehicle). In some cases, unsafe states can develop over time and may be predictable based on the current position and trajectories of the targets, and with some knowledge of pilot intent. This is the approach which is used with airborne conflict alerting systems such as TCAS or various conflict probes. The surface system is, however, a much more difficult case because of the close proximity between safe and unsafe states and the much more transient vehicle dynamics (e.g.. vehicles can stop and start much more quickly on the surface than in the air). For example, it is possible for the surface situation to quickly transition from a safe state into an unsafe state, as when an aircraft which has been holding at a hold line misinterprets a command and suddenly crosses a hold line and blunders onto an occupied runway.

Because of the close spatial proximity between hazardous and non-hazardous states, the fast vehicle dynamics, and the processing delays noted in Figure 2, it is unlikely that any system will be able to detect and prevent *all* potential surface hazards. However, those potential hazard cases which cannot be reliably detected may be addressed through procedural or other means. For example, if an increased buffer zone was provided between the hold short line and the runway incursion impact zone, then more time would be available to respond to a hold short violation. Obviously this may not always be practical due to airport surface constraints. Consequently a total systems analysis which includes both the surveillance system and the operating procedures must be conducted to evaluate the impact of both procedural and systems performance.

7. Total System Performance

7.1 System Operating Characteristic (SOC) Analysis

Given the ability of a surveillance system to detect potential unsafe states as discussed in Section 6, the total system performance must be considered which includes all the components of the system shown in Figure 2. Key components which must be considered are the detection system performance, the delay and correctness of the response of the controller and surface agents, as well as communication delays and other process delays in the system. One approach to total system performance is the System Operating Characteristic (SOC) analysis proposed by Kuchar[19] who used a Monte Carlo analysis to model the different components of the total system. The output of the SOC analysis is a curve (which is much like the ROC of SDT, as described previously) for the individual surveillance components of the system. Like the ROC, the SOC represents the trade-off between the probability of a correct detection and a false alarm. The key difference is that the SOC analysis also models the response actions of the controllers and surface agents, including procedures and response dynamics in determining if a situation results in a missed detection or false alarm.

Total system performance is represented by the aggregate of the sub-components in Figure 2. For any hazardous situation, the performance of each component must be considered to determine if the system is capable of determining the potential hazard state early enough that it can be displayed to the controller, who can determine appropriate commands to be communicated to and executed by the various surface agents in time to avoid the potential hazard. As in all alerting systems there will be a balance between false alarms and missed detections. Note that these can be considered both at the surveillance level but most importantly at the total system level. A correct detection from the surveillance system which occurs too late for the controller to communicate to the surface agents would be considered a missed detection from the perspective of the total system performance.

7.2. Fuzzy Signal Detection Theory (Fuzzy SDT)

Evaluation of the detection of unsafe states can also be examined using fuzzy SDT.[15] While SDT is well established as a quantitative method for the analysis of detection performance, it assumes a strict division of "states of the world" into one of two non-overlapping categories, *signal* or *noise*. This assumption may be violated in real settings because the "signal" (e.g., unsafe state on the runway) varies with context and over time. A signal can therefore be defined as having a value that falls in a range in between unequivocal presence (a "1") and absence (a "0"). The response of the detection system can be non-binary as well, and could be indicated with different degrees of confidence or strength. Thus, in fuzzy SDT an event belongs to the set “signal” with some degree s between 0 and 1 and the detection response to the set “response” with a degree r between 0 and 1. A *mapping function* transforms the variables that describe the state of the world to the signal value s. Consider a simple mapping function for detection of unsafe separation between two airborne aircraft. As Figure 5 shows, this function maps the horizontal separation distance a between two aircraft to s such that as a decreases, the event becomes more "signal"-like (unsafe state), and conversely, becomes less signal-like as a increases. The legal FAA definition of an aircraft-to-aircraft "conflict" is horizontal separation of 5-nmi or less. The monotonic decreasing function allows an increasingly sharp drop-off of s as a increases beyond the 5-nm cutoff, and yields relatively similar (low) values of s at all high values of a (for additional details, see Parasuraman et al.[15]). Also shown in Figure 5 is the mapping function that would be used in standard SDT, in which a violation of the 5-nmi cutoff defines a signal with value $s = 1$. For $a>5$ nmi, the standard (or "crisp") function assigns no value to signal ($s = 0$).

A mapping function for the detection response r can be similarly defined. Alternatively, r can be strictly binary (0 or 1). Once s and r have been determined, fuzzy SDT assigns category membership to the four detection outcomes, namely correct detection, false alarm, miss, and correct rejection. In standard SDT, the four outcome categories are specified based on logical implication functions (e.g., “If s=1 and r=1, then correct detection”; If s=0 and r=1, then false alarm; etc.). Parasuraman et al.[15] proposed the following functions for fuzzy SDT outcome category membership:

$$\text{Correct detection} = \mathbf{min}\ (s, r)$$

$$\text{Miss} = \mathbf{max}\ (s-r, 0)$$

$$\text{False alarm} = \mathbf{max}\ (r-s, 0)$$

$$\text{Correct rejection} = \mathbf{min}\ (1-s, 1-r)$$

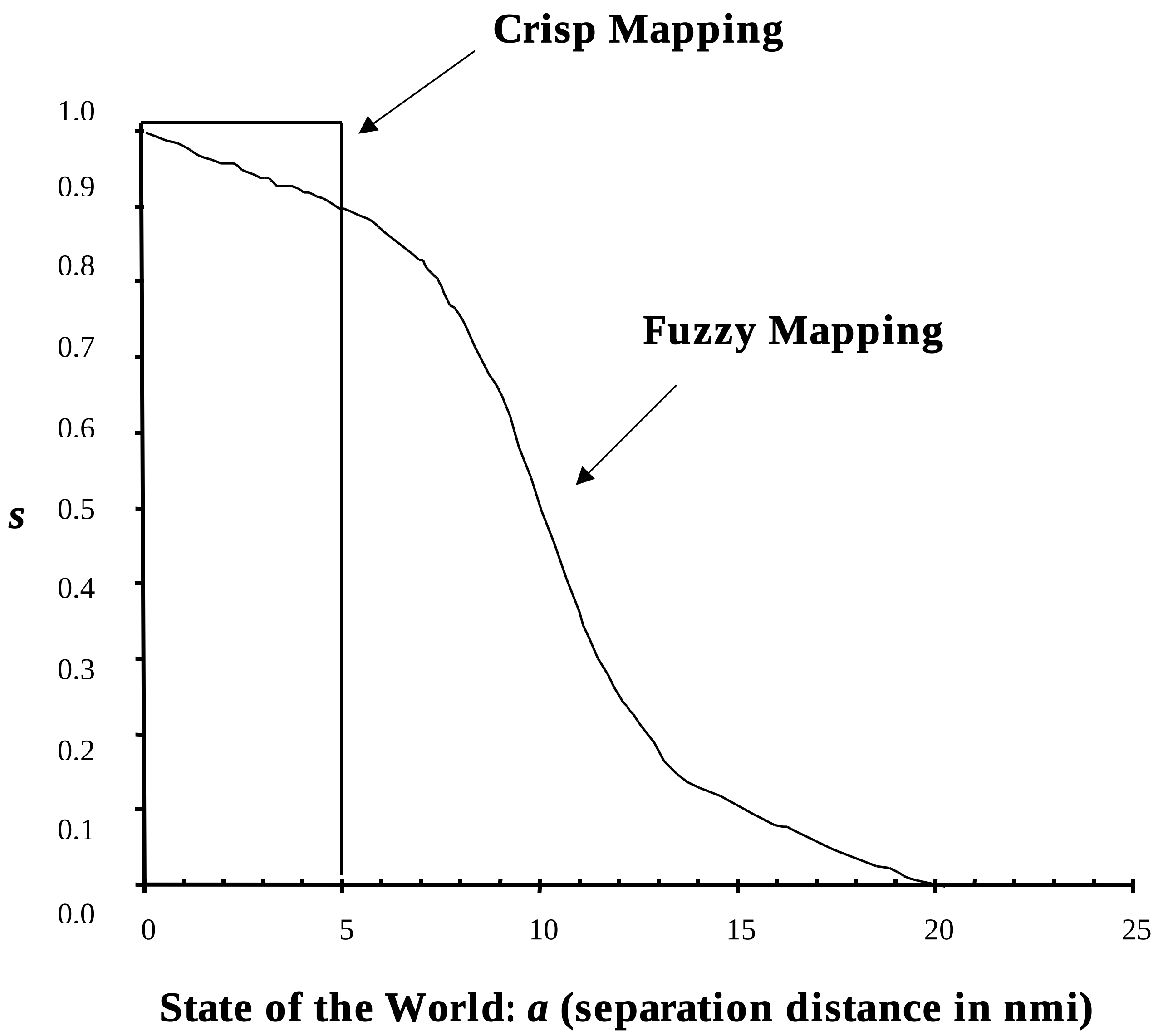

Figure 5. Fuzzy SDT Mapping Function of Horizontal

Separation istance (*a*) to Signal Value (*s*). [From Ref. 15]

As an example of the results of this function set, suppose that *s*=.8 and *r*=.9. That is, the state of the world strongly but not absolutely points to a signal, and the detection system strongly responds that a signal is present. Applying the equations results in the following category memberships: correct detection=.8, miss=0, false alarm=.1, and correct rejection=.1. Hence the outcome strongly points to a correct detection, but unlike standard SDT, there is also some membership in the false alarm and correct rejection categories.

Fuzzy SDT analysis may be particularly well suited to the evaluation of detection performance in real settings where the definition of a signal (e.g., "unsafe state") varies. Parasuraman et al.[15] discussed application of fuzzy SDT to the analysis of the URET tool, which aids en route controllers in determining conflict-free trajectories for aircraft. URET evaluates the paths of all the aircraft that are or shortly will be in the controller's sector and generates either a red or a yellow alert with respect to any pair of aircraft that will approach within certain distances. A red alert is given if the predicted separation distance will be less than the legal criterion of 5 nmi and a yellow alert for separations that are greater than 5 nmi but still considered likely and worthy of attention. In some cases URET gives no alert at all, or may give an alert too late to be of much assistance to the controller. A fuzzy SDT analysis of field data recorded with URET showed more clearly than standard SDT which characteristics of the system produced more effective performance. For example, fuzzy SDT showed that there were a number of instances when the separation distance was less than 5 nmi (below the legal limit) but the URET alert status was yellow, indicating that in these cases, controllers were not receiving the strongest possible alert. Therefore, the fuzzy SDT analysis provided clues as to where room for improvement might lie in the underlying URET algorithms. More generally, fuzzy SDT can provide estimates of sensitivity (*d*') and criterion (ß) that may better capture the temporal and contextual variability inherent in real-world signals that need to be detected by surveillance and alerting systems.

A fuzzy SDT analysis could also be carried out to evaluate the detection of unsafe states in the surface environment. For example, for an aircraft approaching an occupied runway, s could be a function of such variables as speed, initial separation, time, and other factors. Given that these variables can be measured or estimated, the mapping function transforms them into a signal value s which can take on any value between 0 and 1. Category memberships for correct detections and false alarms could then be determined as described previously, followed by sensitivity (d') and ROC analysis. Fuzzy SDT can complement standard SDT and SOC analyses and can provide for a more comprehensive evaluation of the effectiveness of surface surveillance and alerting systems.

7.3 Available Time and Resolution Procedures

As discussed previously, total system performance will be strongly dependent on both the available time and specific resolution procedures. The detection of critical surface targets and of unsafe states must be followed by the selection and application of appropriate resolution procedures. A critical variable is the *available time* for effective action to resolve the problem and return to a safe state. While a certain minimum time is essential for any action to be taken, unlimited time is not necessarily the goal. There will exist a range of times that balance the need for early detection, giving sufficient time for resolution, with the need for avoiding false or nuisance alerts if the system is designed to provide for very long notification times. The FAA's experience with field testing of the AMASS system suggests that the optimal range of times may vary across airports, given differences in approach patterns, procedures, traffic mix, etc. A rough estimate of 20 seconds has been suggested as a reasonable time to give the controller time to take effective action[24].

Fast-time simulation methods are available for determining the probability of detection of unsafe states, given assumptions regarding the detection, communication, and response times of controllers and pilots in certain specified scenarios. Surveys of the runway incursion problem[3] have shown that one of the most frequent incursion scenarios involves a taxiing aircraft crossing a hold short line onto an active runway on which another aircraft is taking off or landing. In a fast-time simulation study of runway incursions, Reynolds and Hansman[25] used probability density functions to represent the time delays introduced by the detection and resolution processes, both at baseline and with an enhanced surveillance system of the AMASS type. They assumed notional values of delay times for the controller, intruder pilot, and evader pilot. This analysis showed the improvement in detection performance and available response times that can be expected with enhanced surveillance technology. This analysis methodology could be extended to evaluate the ASDE-X system if supplemented by empirically-derived (i.e. through human-in-the-loop simulation) or computationally-derived (e.g., with human cognitive models such as MIDAS[26]) values for human response times. In addition, the analysis could be extended to consider other technologies, procedures, and incursion scenario types.

8. Controller Performance: Situation Awareness and Workload

The human performance parameters described in the previous sections can be collected, analyzed and made part of the design process for new surface safety technologies. The result can be a range of possible design options. However, additional consideration of two important human-system issues, controller workload and SA, is needed to ensure effective use of the systems. Hence workload and SA assessment must also form part of the design stage, not merely at the interface evaluation phase.

Controller workload can be defined as the degree to which the controller's tasks tax his or her mental and physical abilities.[27] Workload reflects both the task demands imposed on the controller and the amount of effort expended by the controller in meeting those demands. Controller workload will be impacted by aspects of the functional design of the surface safety technology. For example, while the detection performance analyses outlined in previous sections can be used to design a system with an apparently reasonable false alarm rate, workload may go up if controllers have to deal with several false or nuisance alarms. Hence, the acceptable false alarm rate must be set not only using detection methodologies but also workload analysis. A projected false alarm rate of "1 every 200 hours" was described at the Surface Safety Framework Study meeting.[24] This notional value needs to be validated and analyzed for its workload impact.

Several methods[28] for assessing workload and SA[29] are available. For systems that are not yet built, modeling techniques for estimating user workload are available[30].

The negative impact of false alarms may be counteracted in part through the enhancement of controller SA. As shown in Figure 2, the display system can provide both alerting and SA information to the controller. To the extent that the new technology enhances the controller's SA beyond that obtainable with the baseline system, controllers may be willing to tolerate relatively high false alarm rates. Previous experience with the AMASS and ASDE-3 systems suggests that high false alarm rates can result in an unacceptable system even with the added enhancement of SA; hence the ASDE-X system must be designed with appropriately low levels of false alarms.

Multiple levels of alert (e.g. caution and warning) also promote controller understanding of the system and therefore reduce workload. In any event, the complexity of airport surface operations with their rapid state changes and small separation distances, place significant requirements on the accuracy and reliability of surface surveillance and safety systems. In addition to modeling and human-in-the-loop simulations, "shadow" operations may provide another means of assessing the impact of new technologies on controller workload and SA. This technique works with certain technologies such as runway status lights, which can be evaluated in shadow operations in an backroom control tower. Controllers are exposed to the system by having them observe the simulated operation of the lights in response to actual operations that they can watch from the tower cab. This may be an effective approach because runway status lights can be seen to work with "every" operation, unlike alerts from a detection system that only occasionally notifies the controller.

9. Application of the Framework: Two Examples

The framework described in this document may be applied to a total runway safety system (e.g., one that contains a surface surveillance system, controller displays, controller alerts, and data link of surface traffic for flight deck display) or sub-systems as they are added to an existing system. Two examples of the application of this framework will be considered here: (1) the design of the controller alerting function to ASDE-X, and (2) the addition of a runway status lights function to ASDE-3 /AMASS. The detailed application of the framework is beyond the scope of this document; therefore, the purpose of these examples is to illustrate how the framework would be applied rather than to draw specific conclusions regarding the effectiveness of either of the example systems.

9.1 Example 1: Controller Alerting Function for ASDE-X

As described in Section 2, the first step in the process is to determine the operational objectives of adding controller alerts to ASDE-X. The basic ASDE-X system consists of a surface surveillance system and controller displays. These systems will provide controllers with a certain level of SA and will, by themselves, prevent a number of potential runway incursions. The operational objective of the alerting function, therefore, is to provide an additional layer of defense against a runway incursion by calling the attention of the controller to a hazardous situation and allowing the controller to make a timely intervention. As shown in Figure 1, it is essential to define the specific scenarios that the alerting function will be designed to address (e.g., a landing on an occupied runway). In the case of controller alerting, certain scenarios have been defined by the AMASS product team, but other scenarios such as those identified by the Runway Safety Team must also be analyzed to determine the general applicability of controller alerts to those scenarios.

Once the scenarios have been identified, the parameters that comprehensively define the alerting function design must be established. This will require modeling of the total system, including the response times associated with the surveillance, display, and alerting functions as well as the reaction times of the pilots and controllers and the response of the aircraft. The designer will select from among the modeling techniques (e.g., Monte Carlo simulation) to explore the safety logic parameter space and select the appropriate starting point. This is represented in the first cycle of Figure 1.

The field experience with AMASS has illustrated the importance of proper selection of the balance between missed detections and false alarms. The SDT analysis described in section 6 can be used to estimate an appropriate theoretical operating point to define an initial set of parameters. The second cycle of the framework of Figure 1 should focus on the measurement of the false alarm rate for a selected surface scenario. This can be done using Monte Carlo simulation of the safety logical parameters for specified incursions. The design parameters will need to be tested by simulating runway incursions under controlled conditions to validate the predicted detection performance.

What specific safety logic parameters should be examined? Two general components of safety logic are: (1) the state estimator and predictor, and (2) display generation. The state estimator translates the surveillance system track information (track history, position, velocity, size, ID, etc.) into an estimate of the vehicle state (e.g., a takeoff, arrival, runway crossing, etc.) and a prediction of its future state (including prediction of future track information). This information is passed to the display generation logic (often also called "safety logic") that turns this into a controller alert, runway status light state, pilot data link alert, etc., in whatever modality is chosen by the system designer. Each of these components has multiple parameters that can be set, including, for state estimator, the thresholds for state transition and the look-ahead time for trajectory prediction and for the display generation, the warning time for alerts, and the current extent of the "runway hot zone" for runway status lights. The design process should consist first of identification of initial values for each of these safety logic parameters.

The iterative process shown in Figure 1 should then be applied for each parameter, alone or in particular combinations.

Once a set of parameters that "passes" the initial criteria has been determined, a repeat of the modeling step must be conducted to validate that the total system performance is as expected. At this point, system designers must verify with the user group (controllers and pilots) that the operating point has been properly selected. While, shadow operations are quite useful in this step, in the case of controller alerts, they principally address only the false alarm performance, except when runway incursions are staged for demonstrations.

If necessary, the whole design process can be iterated to achieve the desired performance. However, in each case, a repeat of the original effectiveness estimate against the operational objectives must be performed to ensure that, these have not been adversely impacted by “over-tuning” of the safety logic in an attempt, for example, to reduce false alarms at the expense of missed detections or drastically reduced warning times. The use of proper human performance models is required in this particular case since it is almost impossible for a user group to extrapolate the impact of alerts in an actual operational environment from the frequent “staged" alerts in a demonstration.

9.2 Example 2: Runway Status Lights with ASDE3/AMASS

AMASS is a tracking and safety logic function that has been added to the ASDE-3 surface surveillance system to provide controller alerts. As part of that system, imaginary “hold bars” are depicted on the intersection of taxiways with runways on the controller display when an aircraft is taking off or landing on the runway, indicating to the controller that if an aircraft is allowed to enter the runway, an AMASS alert will result. It has been hypothesized that this "hold bar" indication on the display could be used to drive a set of runway status lights on the airport surface to indicate to pilots that the runway was unsafe to enter.

Previous scenario analyses[3, 16, 24] can serve as a starting point to determine the operational objectives of this approach. In this particular case, the design philosophy is to use the existing safety logic parameters. Therefore the use of the framework will be to assess the operational suitability and effectiveness of the approach rather than an iterative design approach. In contrast to the controller alert example of Section 9.1, the application of the SDT analysis will focus on frequent, rather than rare events. That is, under ideal circumstances, a controller alerting system will produce an alert only when a runway incursion is in progress, runway status lights must operate every time an aircraft takes off or lands on the runway.

The framework approach, utilizing SDT, therefore, illustrates immediately an area of concern to the system evaluator: a set of safety logic parameters (in AMASS) that has been set to reduce false alarms for rare events is now to be used to produce detections of frequent events. For example, the ASDE-3/AMASS tracker produces a certain rate of missed detections on takeoff (that is, a certain number of aircraft will takeoff on a runway without the AMASS tracker having the aircraft under track). The impact on the controller alerting performance is the product of that missed detection probability and the probability that that operation will be part of a scenario that will require a controller alert (a runway incursion). If the second probability is very low, the missed (alert) detection probability will also be very low. However, the probability that the runway status light will not illuminate during a particular departure (missed detection) is equal to the missed detection probability of the AMASS tracker. Since, at busy

ASDE-3/AMASS airports, there would almost always be a flight crew in position to observe the operation of runway status lights, the potential exists for any significant missed detection rate of the AMASS tracker to produce a lack of confidence in the pilot community as they observe aircraft pass by on the runway without illumination of the runway status light.

Using the tools of the framework (e.g., shadow operations, data playback, human performance modeling, etc.) an assessment of the operational suitability of this approach can be made, preferably with field data (particularly since ASDE-3/AMASS is an operational system). It should be noted, however, that recorded track data alone cannot be used in the assessment since missed detections will, of course, not be recorded. Other means to measure missed detection rate (e.g., observers) must be employed when recording field data.

9.3 Summary of Examples

In both examples described above, the framework is used to provide a human factors assessment of a runway safety system. In the first, it is used as an iterative design tool. In the second, it is used to determine the operational effectiveness of a certain safety system implementation. In both cases, the framework provides a systematic approach to the complex problem of airport surface safety.

10. Conclusions

New technologies such as ASDE-X are being developed in efforts to reduce runway incursions and enhance surface safety at towered airports. The problem of preventing runway incursions requires a "layered" defense, in which different technologies are used so that if any one approach is insufficient, others can help mitigate the problem. These systems and the procedures for using them need to be designed taking human-system issues into consideration. This can be done using the framework outlined in this White Paper. The framework describes the best available methodologies required to assess three major human-system elements, detection system performance, total system performance, and controller performance (emphasizing SA and workload). The methodologies include: SDT and ROC analysis, Bayesian statistics, fuzzy SDT, SOC analysis, Monte Carlo simulation, fast-time simulation with human and system response times, computational cognitive modeling, and modeling of workload and SA. These methods can be applied as appropriate in an iterative manner and tested for different runway scenarios. Two examples are given, involving controller alerts with ASDE-X, and runway status lights with AMASS, to illustrate the use of the framework in an iterative design process.

Application of the framework described in this White Paper may also allow for an assessment of the added safety value of implementing additional components of surface safety technologies. For example, one question that the FAA may pose is, what is the added benefit to surface safety of adding controller alerts to ASDE-X? More generally, how much "bang for the buck" will be gained by including additional surface safety technology components? This is consistent with the "layered defense" concept, but any added safety benefit must also be evaluated in economic terms. While cost-benefit analysis is beyond the scope of this paper, such analyses can be conducted in harness with the human-system evaluation methods described here. Comparing various combinations of technologies and procedures using the framework should provide for a more complete understanding of the effectiveness of future surface safety enhancements. Systematic comparison of technologies and procedures or their combination relative to human-system integration may provide additional information necessary for answering "bang for the buck" questions. Application of the framework can therefore identify viable solution sets from a human performance perspective that can be compared in the economic analysis component of investment analysis in order to reach decisions on the implementation of future surface safety technologies and procedures.

The framework outlined in this White Paper may also be useful in the *continuing* evaluation of safety improvements once new technologies and/or procedures are in place. There is a need not only to assess how much safety improvement occurs with a new surface system but also whether the gain is maintained once the system is fielded and becomes operational. The methods and associated human performance measures described in this paper could also be used for ongoing evaluation of fielded systems.

11. **Abbreviations**

ADS-B	Automatic Dependent Surveillance-Broadcast
AMASS	Airport Movement Area Safety System
ASDE-X	Airport Surface Detection Equipment—Model X
ATC	Air Traffic Control
ß	Decision criterion or threshold measure in signal detection theory
CDTI	Cockpit Display of Traffic Information
d'	Detection sensitivity measure in signal detection theory
FMS	Flight Management System
Fuzzy SDT	Fuzzy Signal Detection Theory
GPWS	Ground Proximity Warning System
ODP	Operator Directed Process
GUI	Graphical User Interface
HF	Human Factors
LAHSO	Land and Hold Short Order
ROC	Receiver Operating Characteristic
SDT	Signal Detection Theory
SA	Situation Awareness
SOC	System Operating Characteristic
URET	User Request Evaluation Tool

12. References

---

[1] White House Commission on Aviation Safety and Security. (1997). *Final report to President Clinton.* (Vice President Al Gore, Chair). Washington DC.

[2] US House Committee on Transportation and Infrastructure (2001, June 22). *Increase in runway collision risks to be focus of Congressional hearing. Committee to review technology needed to reduce number of runway incursions.* Washington DC: US House of Representatives. (www.house.gov/transportation/Press2001/release76.html).

[3] Federal Aviation Administration (2001, June). *FAA runway safety report: Runway incursion severity trends at towered airports in the United States 1997-2000.* Washington DC: FAA.

[4] Barnett, A., Paull, G., & Iadeluca, J. (undated). *Fatal US runway collisions over the next twenty years.* Cambridge, MA: Massachusetts Institute of Technology.

[5] Knott, B., Gannon, A., & Rench, M. (2000). *Human factors in runway incursions: An international review.* (Technical Report No. RA-00-002). HSIAC: WPAFB, OH.

[6] Rodgers, M. (2000). *Human factors and runway safety.* Presentation to the FAA National Summit on Runway Safety. Washington DC.

[7] Aviation Week and Space Technology (2001). Technology is key to Australia's ATC. October 22, 2001, 72-75.

[8] Federal Aviation Administration. (1996). *Report to the committee on the status and organization of human factors within the FAA.* Washington DC: Author.

[9] Riley, V. (2000). Perceived complexity and mental models in human-computer interaction. In T. Samad & J. Weyrauch (Eds.) *Automation, control, and complexity: An integrated approach.* (pp. 59-73). New York: Wiley.

[10] Vakil, S. S., & Hansman, R. J. (2001). Approaches to mitigating complexity-driven issues in commercial autoflight systems. *Reliability Engineering and System Safety*, in press.

[11] Wilhelmsen, H. (1994). Preventing runway conflicts: The role of airport surveillance, tower-cab alerts, and runway status lights. *Massachusetts Institute of Technology Lincoln Laboratory Journal, 7* (2), 149-168.

[12] Swets, J., A., & Pickett, R. M. (1982). Evaluation of diagnostic systems: Methods from signal detection theory. New York: Academic Press.

[13] Edwards, W. (1961). Behavioral decision theory. *Annual Review of Psychology, 12,* 473-489.

[14] Parasuraman, R., Hancock, P.A., and Olofinboba, O. (1997). Alarm effectiveness in driver-centered collision-warning systems. *Ergonomics, 39*, 390-399.

[15] Parasuraman, R., Masalonis, A. J., & Hancock, P. A. (2000). Fuzzy signal detection theory: Basic postulates and formulas for analyzing human and machine performance. *Human Factors, 42*, 636-659.

[16] Thompson, S. & Eggert, J. (2001). *Surveillance performance requirements for runway incursion prevention systems.* (Project Report for ATC-300, FAA). Lexington MA: MIT Lincoln Laboratory.

[17] Federal Aviation Administration (2001, May 24). *Airport Surface Detection Equipment—Model X (ASDE-X). NAS Subsystem Level Specification. Version 1.1.* Washington DC: FAA.

---

[18] Kuchar, J. (2001). Markov model of terrain for evaluation of Ground Proximity Warning System thresholds. *Journal of Guidance, Control, and Dynamics, 24,* 428-435.

[19] Kuchar, J. (1996). Methodology for alerting-system performance evaluation. *Journal of Guidance, Control and Dynamics 19*, 438-444.

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

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

[22] Krois, P. (1999). *Human factors assessment of the URET conflict probe false alert rate.* White Paper, AAR-100, Federal Aviation Administration.

[23] Parasuraman, R., & Hancock, P. A. (1999). Designing automated warning systems: A signal detection and Bayesian analysis. In M. W. Scerbo & M. Mouloua (Eds.) *Automation technology and human performance: Current research and trends.* (pp. 63-67). Mahwah, NJ: Erlbaum.

[24] Surface Safety Framework Study Meeting with ASDE-X/AMASS Program Office. (2001, September 28). Washington DC: FAA.

[25] Reynolds, T. G., & Hansman, R. J. (undated). Fast-time simulation modeling of runway incursions. (Presentation slides). Cambridge, MA: International Center for Air Transportation, MIT.

[26] Corker, K., Pisanich, G., & Bunzo, M. (1997). Empirical and analytic studies of human/automation dynamics in airspace management for free flight. In *Proceedings of the 10th International Confederation of European Aerospace Societies Conference on Free Flight*. (pp. 115-120). Amsterdam: CEAS.

[27] Wickens, C. D., Mavor, A., Parasuraman, R., & McGee, J. (1998). *The future of air traffic control: Human operators and automation.* Washington, D.C.: National Academy Press.

[28] Wickens, C. D., & Holland, J. G. (2000). *Engineering psychology and human performance. Third Edition.* Upper Saddle Hall, NJ: Prentice Hall.

[29] Endsley, M. R., & Garland, D. J. (2000). *Situation awareness analysis and measurement.* Mahwah, NJ: Lawrence Erlbaum.

[30] Pew, R., & Mavor, A. (1998). *Modeling human and organizational behavior: Application to military simulations.* Washington DC: National Academy Press.