JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 19, No. 2, March–April 1996

Methodology for Alerting-System Performance Evaluation

James K. Kuchar*
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

A probabilistic-analysis methodology is described that provides quantitative measures of alerting-system performance, including the probabilities of a false alarm and missed detection. As part of the approach, the alerting decision is recast as a signal-detection problem, and system operating-characteristic curves are introduced to describe the tradeoffs between alerting-threshold placement and system performance. The methodology fills the need for a means to determine appropriate alerting thresholds and to quantify the potential benefits that are possible through changes in the design of the system. Because the methodology is developed in a generalized manner, it can be used in a variety of vehicle, transportation-system, and process-control applications. The methodology is demonstrated through an example application to the Traffic Alert and Collision Avoidance System (TCAS). Recent changes in TCAS alerting thresholds are shown to reduce the probability of a false alarm in situations known to produce frequent nuisance alerts in actual operations.

Nomenclature

A, N, T	= probabilistic-state trajectories
E	= event of encountering a hazard
$f_x(x)$	= probability density function for the random variable x
h	= estimated relative altitude
$\dot{h}$	= estimated relative-altitude rate
I	= incident event
P_{malf}	= probability of alerting-system malfunction
$P(x)$	= probability of event x
$P_T(x)$	= probability of event x evaluated along trajectory T
r	= estimated relative range
$\dot{r}$	= estimated relative-range rate
S	= set of hazard situations
s	= hazard situation
$\boldsymbol{x}$	= true state vector
$\hat{\boldsymbol{x}}$	= estimated state vector
$\hat{\boldsymbol{x}}_{\text{thresh}}$	= alerting-threshold location
Z	= exclusion zone
Σ	= set of hazard extents
σ	= specific hazard extent, standard deviation
τ	= specific trajectory

Introduction

ALERTING-SYSTEM performance (e.g., accident and false-alarm rates) depends on a number of parameters, including sensor accuracy, alerting logic and thresholds, and the human response to an alert. These parameters can interact in complex ways, and it is not always clear at what point changes in design should be made to improve performance or where alerting thresholds should be placed. Because of this complexity, a generalized means of modeling and evaluating alerting systems has been lacking, and alerting-system design has traditionally followed an inefficient, evolutionary process. Typical analyses of alerting systems have used Monte Carlo methods to estimate performance under specific system designs.[1–3] Although successful for validating specific system configurations, current methods often do not provide a clear view of the complex design tradeoffs.

This paper presents a methodology for alerting-system evaluation based on a generalized, probabilistic approach and provides a means for examining the tradeoffs between system parameters and performance. With this methodology, alerting-system design efforts can be focused on those parameters having the greatest potential to improve performance.

Before the details of the methodology are described, it is necessary to consider a fundamental issue in alerting-system design: As the alerting system operates, it makes a discrete decision either to remain silent or to issue an alert to warn the human operator of a potential hazard to safety. Typically, this decision is based on whether certain state estimates exceed critical values that define an alerting threshold. Because of errors in measurements or limitations in design, however, faulty alerting decisions can occasionally occur. In particular, a system may fail to alert when necessary (which is termed a missed detection), or may issue an alert when one is not needed (which is termed a false alarm). Although both types of error are undesirable, they generally cannot be eliminated simultaneously. Rather, some compromise between false alarms and missed detections is typically made through the placement of the alerting threshold. For example, a conservative threshold results in early alerts, reducing the probability of missed detection but increasing the probability of false alarm. If the threshold is adjusted to compensate through a delay in the alert signal until more information about the hazard is available, the false-alarm rate will decrease, but the missed-detection rate will increase. Under current design methods, however, this tradeoff is difficult to visualize, and the threshold is often placed by means of an incremental process based on experience with the system in operation.

The methodology discussed here provides a set of generalized tools to predict system performance and to determine where alerting thresholds should be placed. The approach begins with a recasting of the alerting decision as a signal-detection problem, which allows methods from signal-detection theory (SDT) to be applied to the evaluation of system performance. Following a discussion of the mathematics behind the methodology, an example application is provided in which a recent change to a current aircraft alerting system is evaluated.

Analysis Methodology

The alerting decision, in which a binary decision is made on the basis of uncertain data, is analogous to SDT problems in which a threshold is used to decide whether a known signal is present in the background noise.[4–6] In particular, the alerting problem can be thought of as one in which a decision to alert is analogous to the SDT decision corresponding to signal plus noise, and the decision not to alert is analogous to a noise-only decision in SDT.

In SDT, a false alarm occurs when the decision signal plus noise is made when in fact only noise is present. For the alerting decision, the analogous false alarm is a case in which the decision to alert is made when the alert is not needed. An alert is considered to be needed when an undesirable incident is projected to occur. In general, alerts will be false alarms with some probability $P(\text{FA})$.

Received June 8, 1995; revision received Sept. 25, 1995; accepted for publication Oct. 20, 1995.

*Assistant Professor, Department of Aeronautics and Astronautics. Member AIAA.

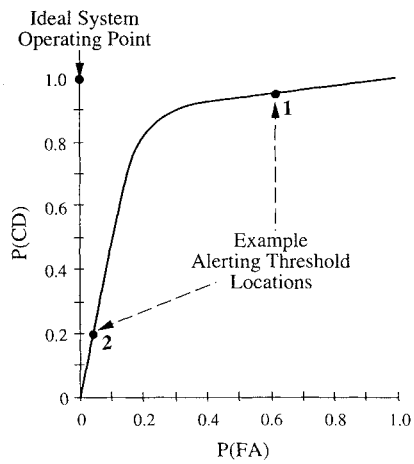

Fig. 1 Example of a SOC curve: P(CD) plotted against P(FA).

Similarly, a missed detection in SDT is a case in which the noise-only decision is made when in fact the signal is present. For alerting systems, the analogous probability of a missed detection P(MD) is the probability that the system fails to protect against an incident. Under this definition, missed detections occur when the system completely fails to alert or when an alert is issued but is too late to prevent an incident. Finally, the probability of a correct detection P(CD) is the probability that an incident is avoided given that an alert has been issued. Note that, with these definitions, it is possible that a given alert could be classified as both a correct detection and a false alarm. In such a case, the alert is not necessary to prevent an incident (thus the false alarm) but an incident is still avoided (thus the CD).

P(FA), P(MD), and P(CD) are functions of the uncertainties in the state estimates and in the location of the threshold. This relation can be examined by the use of an additional connection with SDT. Curves of the system-operating characteristics (SOCs) are plots of the P(CD) vs the P(FA) and are analogs of curves of the receiver operating characteristics used in SDT. Figure 1 shows an example of SOC curve for a hypothetical alerting system. Each choice of an alerting threshold maps onto a single point along the SOC curve. In Fig. 1, two example alerting-threshold locations are shown. If the threshold is placed at location 1, alerts are issued relatively early, resulting in frequent false alarms but also permitting enough time and space to avoid actual hazards successfully, with a high probability. As the threshold is moved to location 2, the alert is delayed, reducing false alarms but also decreasing the ability of the alerting system to provide enough time to avoid hazards.

The ideal operating point is shown in the upper-left corner of Fig. 1 and represents a system that always produces correct detections and does not issue false alarms. Because of uncertainties in measurements, the SOC curve generally does not pass through this point, and ideal performance is not obtainable. However, two methods are available through which system performance can be modified. First, improvements in sensors, displays, and training will shift the SOC curve toward the ideal operating point. Second, given a system configuration that defines the SOC curve, the threshold can be adjusted along that curve to compromise between the costs of false alarms and missed detections.

The SOC curve in alerting-system design is a novel application of SDT for use in analysis and is the foundation of the methodology presented here. Several steps are required to generate SOC curves. First, a model of the hazard-encounter situation must be developed through appropriate dynamical equations. These equations specify constraints such as the vehicle's maneuverability or changes in the severity and extent of the hazard over time. Probability density functions (PDFs) describing the uncertainties in the parameters that define the situation are also needed. For example, measurement uncertainties, the operator's response-time delay, and the aggressiveness of an avoidance maneuver can each be described with appropriate PDFs. The dynamic equations and PDFs are then used to determine performance metrics, such as the P(FA) and P(CD), that are used to generate SOC curves. Finally, given a SOC curve

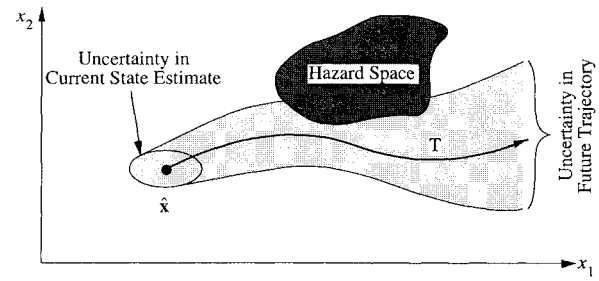

Fig. 2 Example of a probabilistic state estimate $\hat{x}$ and a projected trajectory T.

(or a set of SOC curves for different system configurations), the appropriate alerting threshold can be placed on the basis of the relative costs associated with false alarms and missed detections.

The PDFs describing the various parameters that affect the situation can be difficult to determine. Hardware specifications such as sensor errors can be used to build PDFs describing the uncertainties in the current state estimates. Statistical studies of the situation (from actual events and from simulations) can provide the data required to build PDFs describing the human operator's behavior that affect the outcome of an alert. However, even if the PDFs cannot be estimated accurately, the methodology can be used to examine the sensitivity of the system to different probability distributions.

Hazard Situation

To develop the SOC curve, it is first necessary to model the dynamics of a hazard situation. The methodology uses a state–space approach and places the observable states relating to the vehicle and the hazard in a vector $\boldsymbol{x}$. Figure 2 depicts an example state estimate $\hat{\boldsymbol{x}}$ that includes some uncertainty. An error ellipse is shown that describes the region in the state–space region in which $\boldsymbol{x}$, with some probability, truly lies. The hazard space denoting a region where an incident can occur is also shown. The definition of an incident (and therefore the definition of the hazard space) can be tailored to meet the particular problem under study. For example, an incident could be defined as the collision of two aircraft or, alternatively, as a near-miss event, when aircraft are less than 500 ft apart.

On the basis of the value of $\hat{\boldsymbol{x}}$, the alerting system must determine whether an alert is warranted. The need for an alert, however, depends on the trajectory that will be followed in the future. In Fig. 2, a set of potential future state trajectories T is also shown. T is probabilistic and depends on uncertainties in $\hat{\boldsymbol{x}}$, future control inputs, and knowledge of the system dynamics. The shaded area shown in Fig. 2 denotes the region where the true trajectory will, with some probability, be located.

With T defined and given a particular state estimate $\hat{\boldsymbol{x}}$, there exists some probability that an incident will occur in the future, denoted by $P_T(I \mid \hat{\boldsymbol{x}})$. This notation explicitly shows that the probability is evaluated by the use of the probabilistic trajectory T and that it is a function of $\hat{\boldsymbol{x}}$.

Whether an alert is needed at the time shown in Fig. 2 depends on the value of $P_T(I \mid \hat{\boldsymbol{x}})$. In general, the larger the probability that an incident will occur, the greater the need for an alert. An alert that is issued when $P_T(I \mid \hat{\boldsymbol{x}})$ is small may be considered a false alarm if the human operator is aware of the hazard or would have avoided the hazard without the alert. However, if the alert is delayed until $P_T(I \mid \hat{\boldsymbol{x}})$ is large, there may not be enough time or space in which to perform an avoidance maneuver and an incident may occur even if an alert is issued.

To determine whether an alert is warranted in a given situation, it is necessary to examine the hypothetical outcomes of the alert and do-not-alert decisions. Consider now the possible future state trajectories that will occur on the basis of a given alerting decision. If no alert is issued, the state continues along the projected nominal trajectory N. In response to an alert, there is a discrete change in the actions of the operator, and the state follows an avoidance trajectory A that may avoid an incident. Both N and A are, in general, probabilistic, just as T was in the above discussion. In particular, A may include some probability that no action is taken in response to the alert.

Following the notation used earlier, one can see that the probability that an incident will occur when evaluated at a particular value of $\hat{x}$ and with the assumption that the nominal trajectory is followed is denoted by $P_N(I \mid \hat{x})$. Similarly, $P_A(I \mid \hat{x})$ is the probability that an incident will occur along the avoidance trajectory (i.e., that the avoidance maneuver fails to avoid the hazard). Note that the avoidance trajectory may also result in incidents with other hazards that would not have been encountered had no alert been issued. It is therefore possible to examine both the positive and negative effects of an alert with this method.

Performance Metrics and SOC Curves

Consider now the performance of the alerting system at the instant that an alert is generated, that is, when $\hat{x} = \hat{x}_{\text{thresh}}$, where $\hat{x}_{\text{thresh}}$ defines the alerting threshold. Because an alert is issued, trajectory A will be followed, and $P_A(I \mid \hat{x} = \hat{x}_{\text{thresh}})$ represents the probability that an incident will still occur in the future when the threshold is crossed. Even though trajectory N will not be followed because the alert has been issued, $P_N(I \mid \hat{x} = \hat{x}_{\text{thresh}})$ is useful to consider as a representation of the hypothetical probability of an incident had there been no alert.

The difference in probability values between $P_N(I \mid \hat{x} = \hat{x}_{\text{thresh}})$ and $P_A(I \mid \hat{x} = \hat{x}_{\text{thresh}})$ when an alert is issued represents the effective system benefit. The system benefit is the decrease in the probability of an incident that is generated because an alert is issued.

Another measure of system performance is the safety ratio, which is the ratio between the probability that an incident would occur with the alerting system in operation to the probability that an incident would occur if the system was not in operation, evaluated at the threshold:

$$\text{safety ratio} = \frac{P_A(I \mid \hat{x} = \hat{x}_{\text{thresh}})}{P_N(I \mid \hat{x} = \hat{x}_{\text{thresh}})} \tag{1}$$

A safety ratio of 0 indicates a system that provides perfect protection against incidents. A safety ratio of 1 indicates a system that is of no additional benefit in protecting against hazards. Safety ratios above 1 are possible and indicate alerting systems that actually increase the probability of an incident. For example, an escape maneuver could help avoid one hazard but increase the probability of encountering a second.

Let us return to the connection between the alerting decision and SDT that was introduced above, as it is now possible to define $P(\text{FA})$ and $P(\text{CD})$. Whether the placement of the alerting threshold at $\hat{x}_{\text{thresh}}$ is appropriate can then be measured in terms of $P(\text{FA})$ and $P(\text{CD})$. Recall that a false alarm is an event for which the alert is not needed. Therefore, $P(\text{FA})$ corresponds to the probability that an incident will not occur along N:

$$P(\text{FA}) = 1 - P_N(I \mid \hat{x} = \hat{x}_{\text{thresh}}) \tag{2}$$

$P(\text{MD})$ is given by the probability that an incident occurs along the avoidance trajectory (i.e., the alerting threshold fails to protect against an incident, implying that the opportunity to alert at an earlier time was missed):

$$P(\text{MD}) = P_A(I \mid \hat{x} = \hat{x}_{\text{thresh}}) \tag{3}$$

Note that in general there may be some probability that, even though $\hat{x} = \hat{x}_{\text{thresh}}$, an alert is not issued. Such a situation could arise, for example, if the alerting system malfunctions or is turned off with some probability P_{malf}. In this case, Eq. (3) must be modified to include incidents that occur as a result of system malfunction:

$$P(\text{MD}) = (1 - P_{\text{malf}})P_A(I \mid \hat{x} = \hat{x}_{\text{thresh}}) + P_{\text{malf}}P_N(I \mid \hat{x} = \hat{x}_{\text{thresh}}) \tag{4}$$

For simplicity in the remainder of this paper, P_{malf} is assumed to be negligible and Eq. (3) is used.

$P(\text{CD})$ is the probability that an incident does not occur along trajectory A (i.e., the alert leads to the successful escape from the hazard) and is directly related to $P(\text{MD})$:

$$P(\text{CD}) = 1 - P_A(I \mid \hat{x} = \hat{x}_{\text{thresh}}) = 1 - P(\text{MD}) \tag{5}$$

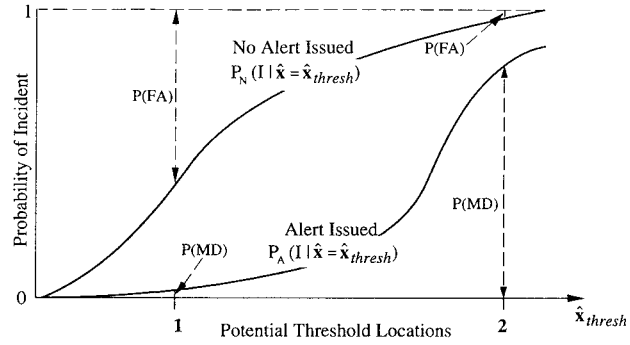

Fig. 3 Alerting threshold location $\hat{x}_{\text{thresh}}$, against the probability that an incident will occur. Also shown are the P(FA) and P(MD).

Note that $P(\text{FA})$, $P(\text{MD})$, and $P(\text{CD})$ are implicit functions of $\hat{x}_{\text{thresh}}$ and the assumed nominal and avoidance trajectories N and A. Also, $P(\text{FA})$, $P(\text{MD})$, and $P(\text{CD})$ are measures of performance when an alert is issued and indicate only how appropriate a given choice of $\hat{x}_{\text{thresh}}$ is. $P(\text{FA})$, for example, is not the probability that a false alarm will occur in the future given some operating condition, but rather the probability that, when an alert occurs, it is a false alarm. For determining the overall probability that a false alarm will occur, additional steps are required and are outlined below.

Figure 3 shows an example plot of $P_N(I \mid \hat{x} = \hat{x}_{\text{thresh}})$ and $P_A(I \mid \hat{x} = \hat{x}_{\text{thresh}})$ as functions of the alerting threshold $\hat{x}_{\text{thresh}}$ for the same system shown by the SOC curve in Fig. 1. If the alerting threshold is placed at location 1, alerts are issued when the value of $\hat{x}$ is relatively low. At this threshold location, there is a small probability of a missed detection because there is ample time and space in which to avoid the hazard successfully. The probability of a false alarm, however, is relatively large because the alert is issued before it is certain that an alert is absolutely necessary. The system benefit for this threshold corresponds to the difference between $P_N(I \mid \hat{x} = \hat{x}_{\text{thresh}})$ and $P_A(I \mid \hat{x} = \hat{x}_{\text{thresh}})$ at location 1.

As the threshold is adjusted toward location 2, $P(\text{FA})$ decreases but $P(\text{MD})$ increases. Somewhere within this range is an optimal threshold location that balances the relative costs of false alarms and missed detections. If a SOC curve (Fig. 1) is used, it is possible to visualize this tradeoff and to determine more easily an appropriate location for the alerting threshold.

Calculation of the Probability of an Incident

Now that the basic technique that is used to develop SOC curves has been described, it is necessary to develop a method by which $P_N(I \mid \hat{x})$ and $P_A(I \mid \hat{x})$ can be calculated for an arbitrary value of $\hat{x}$. First, the type of hazard in the situation must be considered. Some types of hazard space (e.g., severe weather) may exist such that an incident does not necessarily occur. Other types of hazard space (e.g., terrain) are such that an incident occurs with a probability of 1. To distinguish between the event of entering hazard space and the event of having an incident, the term encounter is used to indicate a situation in which the state vector is within the hazard space.

It is first necessary to calculate the probability that the vehicle will encounter hazard space along a probabilistic trajectory $P_T(E \mid \hat{x})$. To find the probability of an incident, one must include an additional term, $P(I \mid E)$, as a measure of the probability of an incident given that exposure to a hazard has occurred. For cases in which $P(I \mid E)$ is a constant, the probability of an incident is given simply by

$$P_T(I \mid \hat{x}) = P(I \mid E)P_T(E \mid \hat{x}) \tag{6}$$

When we deal with hazards for which $P(I \mid E)$ is not a constant, the time or distance that the trajectory remains in the hazard space must be considered by the use of additional methods, as discussed in Ref. 7.

Probability of Encounter and Exclusion Zones

As described earlier, a state estimate $\hat{x}$ is available, but because of measurement errors the state may actually be located at a different position x. The probability that the state is truly at some value x

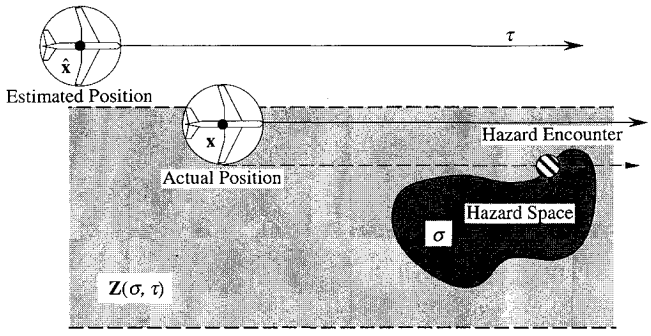

Fig. 4 Exclusion zone $Z(\sigma, \tau)$.

is given by the PDF, $f_x(\boldsymbol{x} - \hat{\boldsymbol{x}})$, and is typically based on sensor error distributions. Note that this PDF need not be Gaussian—any arbitrary distribution can be used. Assume initially that the vehicle's future trajectory, denoted by the vector τ, and the size and shape of the hazard space (or hazard extent), denoted by the vector σ, are known perfectly a priori. The extent of the vehicle must also be defined and may include an extra safety buffer zone if desired. After these parameters have been defined, the probability that the vehicle will encounter the hazard space along the known trajectory τ in a situation with the given hazard extent σ and the current state estimate $\hat{\boldsymbol{x}}$ is denoted as $P_T(E \mid \hat{\boldsymbol{x}}, \sigma, \tau)$.

For the calculation of $P_T(E \mid \hat{\boldsymbol{x}}, \sigma, \tau)$, an exclusion zone $Z(\sigma, \tau)$ is defined as the space in which $\boldsymbol{x}$ must be to cause a conflict along τ. Thus, if $\boldsymbol{x}$ is within $Z(\sigma, \tau)$, the extent of the vehicle will eventually intrude into the hazard space and an encounter will occur (Fig. 4).

The probability that $\boldsymbol{x}$ actually lies within $Z(\sigma, \tau)$ is then given by the integral of the state-estimate PDF over the space defined by $Z(\sigma, \tau)$:

$$P_T(E \mid \hat{\boldsymbol{x}}, \sigma, \tau) = \int_{Z(\sigma,\tau)} f_x(\boldsymbol{x} - \hat{\boldsymbol{x}})\, \mathrm{d}\boldsymbol{x} \tag{7}$$

Extensions to Uncertain Hazard Extent and Trajectory

Equation (7) can be modified to take into account any uncertainties in the hazard extent or relative trajectory. Let us assume next that σ is uncertain and may take on any value from a set of possible hazard extents, Σ, as described by the PDF $f_\sigma(\sigma)$. The probability that an encounter will occur is then given by

$$P_T(E \mid \hat{\boldsymbol{x}}, \tau) = \int_\Sigma P_T(E \mid \hat{\boldsymbol{x}}, \sigma, \tau) f_\sigma(\sigma)\, \mathrm{d}\sigma \tag{8}$$

The same method is also used to solve for $P_T(E \mid \hat{\boldsymbol{x}})$ when τ is uncertain. Assume that τ can take on any value from a set of possible trajectories T. A third PDF, $f_\tau(\tau)$, describes the probability that any particular trajectory occurs and is typically based on reaction time, operator actions, and uncertainties in the dynamics of the situation. The probability that an encounter will occur is then given by

$$P_T(E \mid \hat{\boldsymbol{x}}) = \int_T P_T(E \mid \hat{\boldsymbol{x}}, \tau) f_\tau(\tau)\, \mathrm{d}\tau$$

$$= \int_T \int_\Sigma P_T(E \mid \hat{\boldsymbol{x}}, \sigma, \tau) f_\sigma(\sigma) f_\tau(\tau) \mathrm{d}\sigma\, \mathrm{d}\tau \tag{9}$$

Equation (7) can now be substituted into Eq. (9) to obtain a final expression for $P_T(E \mid \hat{\boldsymbol{x}})$, given uncertain trajectories and hazard extent:

$$P_T(E \mid \hat{\boldsymbol{x}}) = \int_T \int_\Sigma \int_{Z(\sigma,\tau)} f_x(\boldsymbol{x} - \hat{\boldsymbol{x}}) f_\sigma(\sigma) f_\tau(\tau)\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}\sigma\, \mathrm{d}\tau \tag{10}$$

Provided that the PDFs are known or can be estimated, this expression can be solved with numerical-integration or Monte Carlo methods.[8,9] By the use of Monte Carlo or numerical-integration techniques, arbitrary PDFs can be handled as easily as Gaussian distributions. Finally, $P(I \mid E)$ is used [Eq. (6)] to determine the probability of an incident $P_T(I \mid \hat{\boldsymbol{x}})$. For the calculation of $P(\mathrm{FA})$ and $P(\mathrm{MD})$ for a given threshold location $\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}_{\mathrm{thresh}}$, the assumed nominal and avoidance trajectories are used in place of T and $f_\tau(\tau)$ in Eq. (10).

Exposure to Hazards

The above discussions regarding $P(\mathrm{FA})$, $P(\mathrm{MD})$, and $P(\mathrm{CD})$ were concerned with the probability that, when an alert is issued, the alert is a false alarm, is late (a missed detection), or is correct. Often, however, it is of interest to estimate the the overall probability that false alarms or missed detections will occur with an alerting system over some set of operational situations. In general, determining the overall false-alarm or missed-detection rate is a more difficult problem to solve and typically relies on a Monte Carlo simulation.

When one is interested in missed-detection or false-alarm rates over a range of hazard situations, it is necessary to consider the likely exposure to each situation. Each situation may result in an alert at a different threshold location, producing a different value of $P(\mathrm{FA})$ and $P(\mathrm{CD})$. Consider a set S of hazard situations s. Each situation describes a mutually exclusive set of initial conditions and maneuvering behavior. The probability that a certain situation occurs is given by the PDF $f_s(s)$. For each situation, the probability that an alert will occur, $P(\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}_{\mathrm{thresh}} \mid s)$, can be found through Monte Carlo simulation. For each alert that occurs, the probability of a missed detection $P_s(\mathrm{MD})$ can then be found with Eq. (3). The subscript s is added here to indicate the specific situation to which $P(\mathrm{MD})$ applies. The overall missed-detection rate (MDR) can then be found by the integration of $P_s(\mathrm{MD})$ and $P(\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}_{\mathrm{thresh}} \mid s)$ over all possible situations:

$$\mathrm{MDR} = \int_S P_s(\mathrm{MD}) P(\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}_{\mathrm{thresh}} \mid s) f_s(s)\, \mathrm{d}s \tag{11}$$

MDR represents the overall safety level of the system—the rate at which incidents will occur with the alerting system in operation. Similarly, the expected false-alarm rate (FAR) is given by

$$\mathrm{FAR} = \int_S P_s(\mathrm{FA}) P(\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}_{\mathrm{thresh}} \mid s) f_s(s)\, \mathrm{d}s \tag{12}$$

The correct detection rate (CDR) can also be defined:

$$\mathrm{CDR} = 1 - \mathrm{MDR} \tag{13}$$

Given that the MDR and FAR are functions of the alerting threshold, SOC curves may also be constructed to show the CDR vs the FAR to aid in alerting-threshold placement.

The set of situations over which Eqs.(11) and (12) are integrated depends on the application. For example, it may be of interest to find the FAR for a collision-alerting system based on the expected flight paths of aircraft arriving at a particular airport. Each potential flight path is considered to represent a single situation and the FAR would give a measure of the expected false-alarm rate per arrival. Alternatively, the FAR could be determined as a measure of alerts per year or other unit of time.

Example Application of the Methodology

An example application of the methodology is now provided through an evaluation of a recent change in the alerting threshold of the Traffic Alert and Collision Avoidance System (TCAS). TCAS was implemented in the early 1990s on jet transport aircraft and is currently mandated on all U.S. passenger aircraft with more than 30 seats.[10] The system monitors other aircraft and issues two types of alerts to the flight crew when collisions are predicted: traffic advisories are alerts that direct the crew's attention to a potential threat when the likelihood of a collision is relatively low and resolution advisories (RAs) are issued when immediate maneuvering is required and provide graphic and verbal avoidance commands such as climb or descend. To limit the scope of this paper, only RAs are considered here.

Problem Statement

The example involves a specific geometry between two aircraft that has been known to cause false alarms during actual operations.[11] The situation includes a TCAS aircraft that uses the alerting system

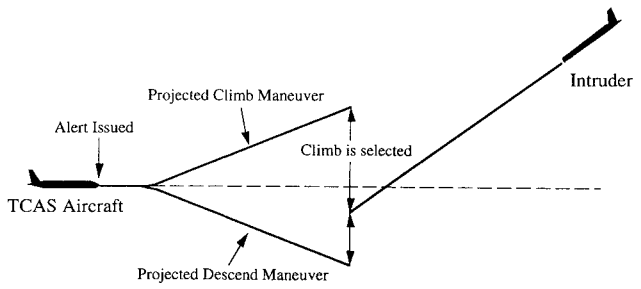

Fig. 5 Example of selection of a climb or descend RA.

under study, and an intruder aircraft that is assumed to not react to TCAS alerts. In this situation, the intruding aircraft is descending at a high vertical rate toward the TCAS aircraft (Fig. 5).

When an RA is issued, the TCAS examines several potential vertical avoidance maneuvers. In this situation, it assumes a 5-s response delay to the alert followed by a 0.25-g pitching maneuver to a vertical rate of 1500 ft/min. Both climbing and descending maneuvers are examined in relation to a straight-line extrapolation of the intruder's trajectory. The maneuvering advisory is chosen from the avoidance maneuver that provides a certain vertical separation at the point of closest approach, assuming that the intruder maintains a constant altitude rate. In Fig. 5, because the climb projected-vertical-miss distance is larger than the descend vertical-miss distance, a climb RA is chosen. However, in the meantime the intruder may level off above the TCAS aircraft, in which case an alert was not truly needed and may in fact increase the probability of a collision as the TCAS aircraft climbs.

Because of its tendency to produce false alarms when intruders level off above the TCAS aircraft, this situation has resulted in the recommendation that aircraft slow their rates of descent when approaching their final altitude.[11] In addition, the TCAS alerting thresholds have been modified several times during the past decade to reduce false alarms in this type of encounter.[12] The next example investigates how these threshold modifications have affected the performance of the alerting system. In particular, a comparison is made between the original version of TCAS (called here Version 1.0)[13] and the latest implementation of TCAS, called Version 6.04A, which was mandated on U.S. aircraft in December 1994.[1,12]

In the example discussed here, the TCAS aircraft is flying level at an altitude of 15,000 ft above mean sea level. The intruder aircraft is flying directly toward the TCAS aircraft on a reciprocal course, at a relative speed of 434 kn (733 ft/s), and with a descent rate of 2500 ft/min. These values were chosen as representative of aircraft maneuvering during the transition between cruise and approach, when false alarms with TCAS have been known to occur. The two aircraft are modeled as cylinders with a 100 ft radius and a height of 100 ft. If the cylinders intersect, a midair-collision incident has occurred.

For this example, the intruder's trajectory is specified such that the intruder is projected to pass 80 ft below the TCAS aircraft. This vertical-miss distance is used so that nominally a collision will occur, and, because the intruder is passing just below the TCAS aircraft, TCAS will issue a climb RA command. Furthermore, it is assumed that the intruder has been tracked by the TCAS aircraft long enough that the estimate uncertainties have reached steady-state values.

To provide additional uncertainty in the behavior of the intruder, it is assumed that the intruder will level off 1000 ft above the TCAS aircraft, with a probability of 0.75, and that the intruder will continue to descend through the TCAS aircraft's altitude, with a probability of 0.25. A normally distributed window of potential altitudes at which the level-off maneuver begins is also defined. This window is centered on the nominal altitude at which a level-off maneuver would begin if a 0.1-g pull-up were assumed, and it has a standard deviation of 2-s flight time (approximately 80 ft at the descent rate of 2500 ft/min).

Architecture of the TCAS

The TCAS alerting criteria are based on four state estimates (r, $\dot{r}$, h, and $\dot{h}$) that are inferred from measurements of range and altitude. Radar pulses are sent once per second from the TCAS aircraft and are received by the transponder of the intruder aircraft. The intruder's transponder then sends a reply pulse to the TCAS aircraft, and the range is calculated with the round-trip time for the pulse. For altitude measurements, the intruding aircraft must have a transponder that sends its altitude encoded in the reply pulse.

Because the measurements of range contain noise, a linear estimator is used to produce a filtered estimate of the range and the range rate.[13] Range measurements are typically obtained with a standard deviation of approximately 30 ft and can be modeled by the use of a normal distribution. The filter equations were solved with this measurement error to determine the covariance matrix for the estimator, assuming a target moving with a constant range rate of 434 kn. The filter converges within approximately 17 s to a steady-state-error standard deviation of approximately 18 ft in range and 6 ft/s in range rate. Range and range-rate estimate errors are correlated with a correlation coefficient of 0.69.

A more complex, nonlinear tracker is used to estimate the altitude and the altitude rate. The minimum operational performance specifications for the TCAS require altimetry accuracies with a standard deviation of at most $\sigma = 58$ ft for this example situation.[13] In addition, altitude measurements obtained through the transponder are discretized into 100-ft bins. The nonlinear tracker used by the TCAS is designed to produce smooth estimates of altitude despite the coarse, discrete resolution of the altitude data.

Monte Carlo simulations were performed to determine the accuracy with which the TCAS nonlinear tracker estimates the altitude in this situation. In the simulation, the intruder's altimeter was assumed to produce a normally distributed error with a standard deviation of 58 ft. This error was assumed to be composed of a random-noise component of $\sigma = 6$ ft and a random constant bias with $\sigma = 57.7$ ft. The flight of the intruder aircraft was simulated at a constant descent rate of 2500 ft/min, and altitude reports from the intruder were quantized into 100-ft increments and passed to the altitude-tracker algorithm from Ref. 13. The standard deviation of the estimates converge to steady-state values within approximately 20 s. The steady-state altitude-error standard deviation is approximately 58.9 ft, and the altitude-rate-error standard deviation is 0.56 ft/s. Altitude- and altitude-rate-estimate errors are correlated with a correlation coefficient of 0.16.

The TCAS aircraft's altitude estimate also contains errors and is assumed to be estimated to a tolerance of $\sigma = 58$ ft. The total relative altitude error is then given by

$$\sigma_{\text{total}} = \sqrt{\sigma_{\text{TCAS}}^2 + \sigma_{\text{intruder}}^2} = \sqrt{58^2 + 58.9^2} = 82.6 \text{ ft} \qquad (14)$$

The estimation errors are summarized in Table 1.

Alerting Thresholds

The alerting thresholds are rather complex and are described in detail in the TCAS operational specifications.[13] However, for the simple traffic-encounter situation considered here, the alerting thresholds can be reduced to a dependence on two parameters, given in Table 2. The parameter TAU is a threshold on the estimated time to

Table 1 Steady-state estimate accuracies for the TCAS example

Parameter	Estimate standard deviation
r	18 ft
$\dot{r}$	6 ft/s
h	82.6 ft
$\dot{h}$	0.56 ft/s

Table 2 TCAS RA-threshold parameters[a]

Parameter	Version 1.0	Version 6.04A
TAU	30 s	22 s
DMOD	1.0 n mile	0.8 n mile

[a]Values taken from Refs. 12 and 13.

the closest point of approach, and DMOD is a safety-buffer distance around the TCAS aircraft. A range threshold is triggered when the intruder is projected to enter a sphere of radius DMOD around the TCAS aircraft within TAU seconds. An altitude threshold is triggered when the intruder has a projected time to co-altitude of less than TAU seconds. When both the range and the altitude thresholds are triggered, an RA is issued. As shown in Table 2, TCAS Version 6.04A uses a smaller buffer zone (DMOD) and delays alerting by 8 s beyond Version 1.0.

TCAS Performance Evaluation

The uncertainties listed in Table 1 can be used to develop a SOC curve for this situation by the use of the probabilistic methodology. Accordingly, two future trajectories for the TCAS aircraft are considered. In the nominal-trajectory case N, the TCAS aircraft is assumed to continue at its current altitude. In the avoidance-trajectory case A, the TCAS aircraft is assumed to fly a standard RA-avoidance maneuver: after a 5-s delay, the aircraft performs a 0.25-g pull-up maneuver until a 1500-ft/min climb rate has been achieved.

Because a collision threat is a hazard for which $P(I \mid E) = 1$, Eq. (10) is used to find $P_N(I \mid \hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{thresh}})$ directly, as discussed in the following. As shown in Fig. 6, the nominal-trajectory situation can be modeled as one in which the TCAS aircraft moves relative to a fixed intruder aircraft representing the hazard space. The direction of the relative trajectory followed by the TCAS aircraft, τ, is a function of $\dot{r}$ and $\dot{h}$. Given a particular trajectory, if the TCAS aircraft is located at a distance $h - \dot{h}(r/\dot{r})$ above the estimated relative altitude h, then a direct collision will occur. Because the two aircraft are modeled as cylinders 100 ft high, the exclusion zone is then defined as the approximate region between $h - \dot{h}(r/\dot{r}) - 100$ and $h - \dot{h}(r/\dot{r}) + 100$, as shown in Fig. 6.

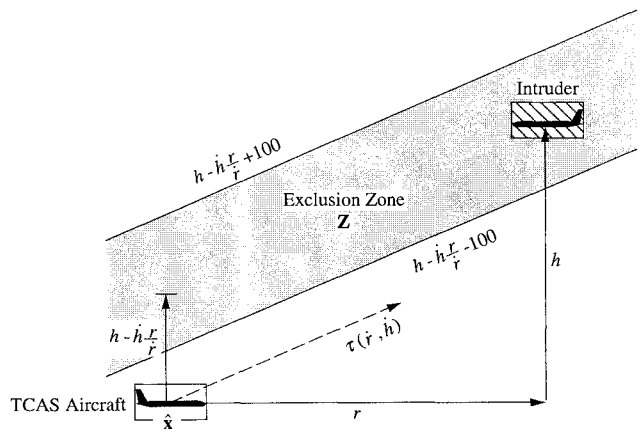

Fig. 6 Exclusion zone for the TCAS nominal-trajectory example.

Thus, the integral over T in Eq. (10) is a double integral over $\dot{h}$ and $\dot{r}$, and the integral over Z is a double integral over r and h:

$$P_N(I \mid \hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{thresh}}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{h-(\dot{h}r/\dot{r})-100}^{h-(\dot{h}r/\dot{r})+100} f_{h\dot{h}}(h,\dot{h})\, f_{r\dot{r}}(r,\dot{r})\, \mathrm{d}h\, \mathrm{d}r\mathrm{d}\dot{r}\, \mathrm{d}\dot{h} \tag{15}$$

where $f_{h\dot{h}}(h,\dot{h})$ and $f_{r\dot{r}}(r,\dot{r})$ are jointly normal PDFs with the standard deviations listed in Table 1. A slightly more complicated but similar method is used to calculate $P_A(I \mid \hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{thresh}})$ along the avoidance trajectory.

Equation (15) was solved with numerical integration as a function of the relative altitude h to produce curves of $P_N(I \mid \hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{thresh}})$ and $P_A(I \mid \hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{thresh}})$, as shown in Fig. 7. Overlaid on these curves are the locations of the alerting thresholds for TCAS Versions 1.0 and 6.04A. Equation (15) was also solved by means of a Monte Carlo integration to estimate the accuracy of the results. The values of $P(I \mid \hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{thresh}})$ shown in Fig. 7 are accurate to ± 0.01 or better.

The curve representing $P_N(I \mid \hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{thresh}})$ shows the probability that an incident will occur if the TCAS aircraft does not maneuver as a function of the relative altitude between the two aircraft at the time at which an alert is issued. At large relative altitudes (location 1 in the Fig. 7), $P_N(I \mid \hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{thresh}})$ is approximately 0.14: this is the probability that an alert will eventually be needed when the two aircraft are still far from one another. As the relative altitude decreases and the two aircraft come closer together, the probability that the intruder will not level off increases, resulting in a corresponding increase in $P_N(I \mid \hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{thresh}})$. At low relative altitudes (location 2), $P_N(I \mid \hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{thresh}})$ is approximately 0.57. Note that, even at a measured relative altitude of 0 ft, the probability that an incident will occur is less than 1 because of the relatively large uncertainties in the state estimates.

The curve representing $P_A(I \mid \hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{thresh}})$ shows two regions of local maximum. Between relative altitudes of approximately 2500 ft and 1300 ft (location 1), an avoidance maneuver risks a collision with an intruder that has leveled off. Below a relative altitude of 700 ft (location 2), $P_A(I \mid \hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{thresh}})$ rises because there is not enough time available to avoid a threat that has not leveled off. At relative altitudes above 2500 ft, an RA will allow the TCAS aircraft sufficient time to climb well above the intruder's trajectory, whether the intruder levels off or not.

Figure 8 shows the SOC curve for the data from Fig. 7. The unconventional shape of the SOC curve is a result of the fact that both high and low thresholds (locations 1 and 2 in Fig. 7) result in low values of $P(\text{CD})$ in this situation. Because the probability of a

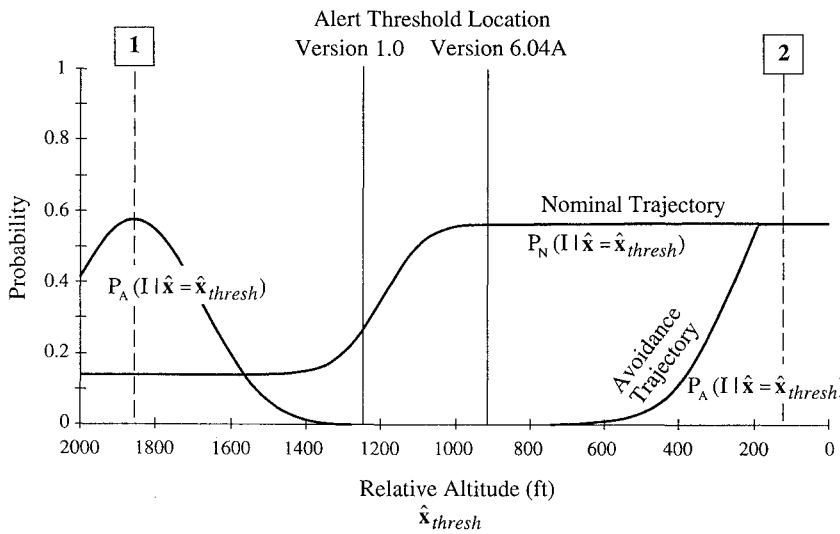

Fig. 7 Probability of incident vs relative altitude.

Table 3 Comparison of P(MD) and P(FA)

TCAS version	P(MD)	P(FA)
1.0	2×10^{-4}	0.63
6.04A	1×10^{-11}	0.44

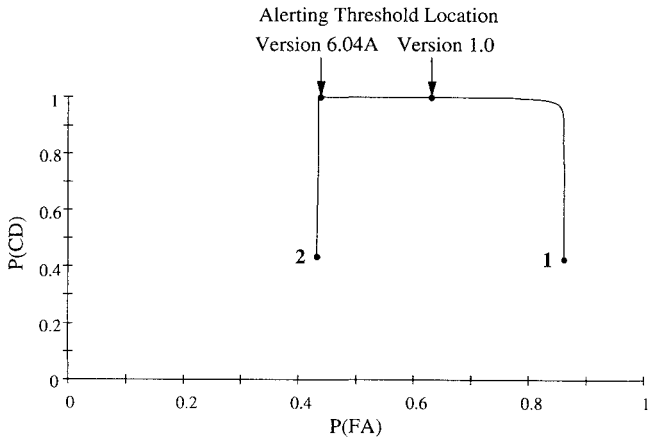

Fig. 8 SOC curves for TCAS Versions 1.0 and 6.04A.

false alarm does not drop below approximately 0.43, the SOC curve does not reach the y axis in the plot.

As shown in Fig. 8, the change in thresholds from TCAS Version 1.0 to Version 6.04A results in a reduction in the probability of a false alarm (the data are summarized in Table 3). Furthermore, P(MD) is also lower in Version 6.04A than Version 1.0: the probability that the TCAS aircraft will climb into a leveled-off intruder is lower because the alert is delayed under Version 6.04A. As can also be seen, Version 6.04A appears to have moved the threshold as far as possible before P(CD) drops sharply. Thus, delaying the alert much beyond Version 6.04A would not provide enough time to avoid a descending intruder. Use of a Monte Carlo evaluation and a change of the integration step size show that the probability values in Table 3 are estimated to be accurate to better than approximately 0.01 for P(FA) and are accurate to within a factor of 10 for P(MD).

In other situations, there may be more or less benefit derived from the TCAS threshold change than is evident from this example. A complete evaluation of the differences between Versions 1.0 and 6.04A would require an examination of all likely encounter situations and aircraft geometries. With the use of Eqs. (11–13), a comprehensive SOC curve could be developed to examine the overall performance of the TCAS.

Conclusions

Alerting-systems design must balance a tradeoff between alerting early to increase safety and delaying an alert to reduce unwanted false alarms. In complex alerting systems, these tradeoffs are often unclear under current methods. Therefore, a methodology to model and to evaluate alerting systems in a variety of applications is needed to provide insight into the issues that affect alerting-system performance.

This paper describes a generalized methodology that provides a means by which the performance of alerting systems can be defined and evaluated when the alerting decision is recast as a signal-detection problem. This approach accounts for uncertainties in measurements, the vehicle trajectory, and the size, shape, and severity of the hazard. These uncertainties are represented by appropriate probability density functions, which need not be Gaussian. Curves of the SOCs are used to illustrate the tradeoff between false alarms and missed detections. With the use of a SOC curve, an appropriate alerting-threshold location can be determined if the costs associated with false alarms and missed detections are weighed. Parametric studies can also be used to examine the impact of changes in sensors, dynamics, and the human response on the SOC curve and on alerting-system performance. The methodology has a wide applicability, including applications in transportation systems and process control.

To demonstrate the utility of the methodology, it is applied to the TCAS—an alerting system used on civil aircraft for the prevention of midair collisions. The basic properties of the TCAS are illustrated for an example situation that is known to produce false alarms in actual operations. With a probabilistic model of the state-estimate uncertainties, the methodology is used to determine the probability that a TCAS alert is a false alarm or a correct detection. Two versions of the TCAS are examined: the original Version 1.0, and the updated Version 6.04A.

As an example, the methodology is used to generate a SOC curve. By a comparison of the alerting thresholds along the SOC curve, it is shown that Version 6.04A can reduce the probability of a false alarm (as compared with Version 1.0) to the approximate minimum value that is possible before the probability of a missed detection increases rapidly. In addition, the probability of a missed detection is shown to decrease with TCAS Version 6.04A. Thus, the methodology provides an alternate, consistent analysis of the TCAS that predicts that the threshold changes in TCAS Version 6.04A will have the intended effect of a reduction in false alarms in the situation studied.

Acknowledgments

This work was supported by the Federal Aviation Administration under Grant 92-6-0001 and by the Boeing Commercial Airplane Company. Much appreciation is extended to R. John Hansman Jr. at the Massachusetts Institute of Technology for his valuable input during this research. The author thanks Don Bateman from AlliedSignal, Inc., George Boucek from the Boeing Commercial Airplane Group, and Andy Zeitlin from MITRE Corp. for providing information on current alerting systems. Bob Giuda and United Airlines also deserve recognition for their input and for providing jumpseat privileges so that the author could observe cockpit operations.

References

[1]Shank, E. M., and Hollister, K. M., "A Statistical Risk Assessment Model for the Precision Runway Monitor System," *Proceedings of the 37th Annual Air Traffic Control Association Conference* (Atlantic City, NJ), Air Traffic Control Association, Arlington, VA, 1992, pp. 686–692.

[2]Ebrahimi, Y. S., "Parallel Runway Requirement Analysis Study," NASA CR-191549, Dec. 1993.

[3]Miller, C. A., Williamson, T., Walsh, J. A., Nivert, L. J., Anderson, J. L., and Zeitlin, A. D., "Initiatives to Improve TCAS–ATC Compatibility," *Journal of ATC*, Vol. 36, No. 3, 1994, pp. 6–12.

[4]Barkat, M., *Signal Detection and Estimation*, Artech House, Boston, MA, 1991, pp. 115–174.

[5]Sheridan, T. B., and Ferrell, W. R., *Man–Machine Systems: Information, Control, and Decision Models of Human Performance*, MIT Press, Cambridge, MA, 1974, pp. 355–382.

[6]Sheridan, T. B., *Telerobotics, Automation, and Human Supervisory Control*, MIT Press, Cambridge, MA, 1992, pp. 49–60.

[7]Kuchar, J. K., and Hansman, R. J., "A Unified Methodology for the Evaluation of Hazard Alerting Systems," MIT Aeronautical Systems Lab., Rept. ASL-95-1, Massachusetts Inst. of Technology, Cambridge, MA, Jan. 1995.

[8]Edwards, C. H., Jr., and Penney, D. E., *Elementary Differential Equations with Applications*, Prentice–Hall, Englewood Cliffs, NJ, 1985, Chap. 6.

[9]Shreider, Y. A., *The Monte Carlo Method*, Pergamon, New York, 1966, Chap. 2.

[10]Federal Aviation Administration, *U.S. Federal Aviation Regulations (FAR)*, Pt. 121, Federal Aviation Administration, Washington, DC, 1994, Sec. 121.356.

[11]Mellone, V. J., and Frank, S. M., "Behavioral Impact of TCAS II on the National Air Traffic Control System," *Proceedings of the Seventh International Symposium on Aviation Psychology* (Ohio State Univ.), Ohio State Univ. Aviation Psychology Lab., Columbus, OH, 1993.

[12]Federal Aviation Administration, "TCAS II Collision Avoidance System (CAS): System Requirements Specification," Version 6.04A, Federal Aviation Administration, Washington, DC, 1993.

[13]Radio Technical Committee on Aeronautics (RTCA), "Minimum Performance Specifications for TCAS Airborne Equipment," Doc. RTCA/DO-185, Radio Technical Committee on Aeronautics, Washington, DC, Sept. 1983.